REVEALING CHOICE BRACKETING

ANDREW ELLIS AND DAVID J. FREEMAN

Abstract. In a decision problem comprised of multiple choices, a person may fail to take into account the interdependencies between her choices. To understand how people make decisions in such problems, we design a novel experiment and revealed preference tests that determine how each subject brackets her choices. In separate portfolio allocation under risk, social allocation, and induced payoff function shopping experiments, we find that 40-43% of our subjects are consistent with narrow bracketing while only 0-15% are consistent with broad bracketing. Classifying subjects while adjusting for models' predictive precision, 74% of subjects are best described by narrow bracketing, 13% by broad bracketing, and 6% by intermediate cases.

JEL codes: D01, D90

Keywords: choice bracketing, individual decision-making, revealed preference, experiment.

1. Introduction

Many decisions consist of multiple interdependent parts, and considering each in isolation may lead to different choices than when considering the whole. For instance, a person might invest differently in her retirement account if she evaluated the account's distribution of returns on its own instead of as part of her overall portfolio. How a person takes into

Date: January 2021.

Freeman thanks UCSB and QMUL for hosting him while parts of this paper were written, and UToronto for use of its TEEL lab for our experiments. We especially thank Johannes Hoelzemann for help in coordinating experiments at TEEL. We thank Ted Bergstrom, Ignacio Esponda, Erik Eyster, Yoram Halevy, Matt Levy, Marco Mariotti, Antony Millner, Ryan Oprea, Luba Petersen, Emanuel Vespa, Sevgi Yuksel, Lanny Zrill, and audiences at UCSB, Chapman, QMUL, UCL, D-TEA 2020,ESA 2020, RUD 2020, ESNAWM 2021 for helpful discussions and suggestions. We thank Alex Ballyk, Daphne Baldassari, Matheus Thompson Bandeira, En Hua Hu, Louise Song, and Johanathan Wang for research assistance. This paper was funded by a 2018 SFU/SSHRC Institutional Grant and SSHRC Insight Grant 435-2019-0658 and was conducted under SFU Research Ethics Study #2016s0380.

Ellis: Department of Economics, London School of Economics. e-mail: a.ellis@lse.ac.uk. Web: http://personal.lse.ac.uk/ellisa1/.

Freeman: Department of Economics, Simon Fraser University. e-mail: david_freeman@sfu.ca. Web: http://www.sfu.ca/~dfa19/.

account the whole of her problem when choosing in its individual parts, or how she *brackets* the decision, is of first-order importance for understanding, interpreting, and predicting her choices. For instance, bracketing determines which outcomes of a given choice are evaluated as gains and which as losses, whether an outcome is evaluated as fair, and the rate of intertemporal substitution and thus is a crucial determinant of behavior in many economic settings.

While economists have devoted considerable intellectual energy to studying preferences to explain choices in these and other domains, comparatively little is known about bracketing. Existing experiments (e.g. Tversky & Kahneman 1981) show that some people do not *broadly bracket* their choices, i.e. they fail to consider them jointly. However, there are other ways that people could bracket, such as *narrowly bracketing* by focusing only on the part of the problem at hand, or doing something in between these two extremes. Previous work has limited power to detect failures of broad bracketing and even less to detect failures of narrow bracketing, let alone other forms of behavior. Since preferences and choice bracketing jointly determine behavior, separating the two presents a challenging inference problem. This leaves open the empirical question: how do people bracket their choices?

To solve this inference problem, this paper develops a novel experimental design and revealed preference tools that allow us to uncover how a person brackets her choices. In our experiments, each decision is explicitly broken down into its component parts. The alternatives chosen in its constituent parts determine the overall outcome for the decision, and a person's payoff depends only on the overall outcome, not on how it arose. Since both bracketing and preferences jointly determine choice but neither are directly observed, we provide revealed-preference tests for broad and narrow bracketing that control for unobserved preferences. We conduct novel experiments that study a portfolio choice task under risk (Choi *et al.*, 2007) and a social allocation task (Andreoni & Miller, 2002; Fisman *et al.*, 2007), and then apply our revealed preference approach to test broad bracketing, narrow bracketing, and a model of partial-narrow bracketing (Barberis *et al.*, 2006; Rabin & Weizsäcker, 2009) while controlling for unobserved preferences. We also conduct a third experiment that directly induces a payoff function over fictitious goods, and thus measures choice bracketing directly.

Across the three experiments, we find that 40-43% of subjects are consistent with narrow but not broad bracketing, 0-15% are consistent with broad but not narrow bracketing, and the remainder are inconsistent with both. Because it places considerable restrictions on behavior, the substantial number of subjects whose behavior closely adheres to the testable content of narrow bracketing provides strong support for narrow bracketing as a model of how people make multiple interrelated choices. The existence of a notable minority of subjects whose behavior closely adheres to broad bracketing, even when most do not, suggests population-level heterogeneity in bracketing. Moreover, we find a lower fraction of broad bracketers in the Risk and Social Experiments than in the Shopping Experiment. This suggests that broad bracketing is more common in tasks that are more familiar or those where broad bracketing "objectively" improves payoffs.

All economic modeling makes an assumption, usually implicit, about how agents bracket their choices, so predictive accuracy crucially depends on whether assumptions about bracketing are behaviorally accurate. Our design allows us to compute how far each subject's behavior is from the set of behavior allowed by broad, narrow, and partial-narrow bracketing, and then assign them to the model with the best predictive success (Selten, 1991). We find that 74% of subjects are best described as narrow bracketers and 13% as broad bracketers, while only 6% are classified as partial-narrow bracketers, though these fractions differ across our three experiments.[1] The notable heterogeneity we find, and differences across domains, suggest that work assuming homogeneous bracketing compromises its predictive validity, regardless of whether narrow or broad is assumed. Moreover, the results suggest that partial-narrow bracketing is not a predictively-useful generalization of its polar cases.

To illustrate how the design of our Risk and Social Experiments enables individual-level tests of both narrow and broad bracketing that control for unobserved preferences, consider Decision 1 of the Risk Experiment. It consists of two parts, in each of which a subject allocates a fixed budget of tokens between two assets whose payoff depends on the result of a (single) roll of a six-sided die. In Part 1, the subject has 10 tokens to divide between Assets A and B; Asset A pays out $1 if the die rolls 1-3 and $0 otherwise, while Asset B

[1]The test does not classify the remaining 8%.

pays \$1.20 if the die rolls 4-6 and \$0 otherwise. In Part 2, the subject has 16 tokens to divide between Assets C and D; Asset C pays out \$1 if the die rolls 1-3 and \$0 otherwise, while Asset D pays \$1 if the die rolls 4-6 and \$0 otherwise. This decision permits direct tests of both narrow and broad bracketing that require only that the subject ranks lotteries in accord with first-order stochastic dominance. Moreover, if we assume that people are risk averse over 50-50 lotteries, then we additionally obtain a point prediction for broad bracketers and a distinct point prediction in Part 2 for narrow bracketers, since in each case the subject can maximize expected value while eliminating risk within their bracket. These point predictions enable powerful individual-level tests of choice bracketing based only on weak non-parametric assumptions about preferences. Collecting all such implications of narrow and broad bracketing across the five decisions in our Risk and Social Experiments, 0% and 10% of subjects respectively pass our test of broad bracketing while 43% pass our test of narrow bracketing in each experiment.

Our design also enables us to conduct novel revealed preference tests of broad and narrow bracketing. These tests compare allocations in pairs of decisions to control for unobserved preferences, and thus require even fewer assumptions about underlying preferences. In Decision 2 of the Risk Experiment, the set of feasible decision-level bundles over state-contingent payments contains that of Decision 1. Thus a broad bracketer who chooses a bundle in their intersection in Decision 2 must make choices that generate the same decision-level bundle over payments in Decision 1. Similarly, Decision 5 is identical to Part 1 of Decision 1, so a narrow bracketer should make the same choice in both. These revealed preference tests can be deployed in almost any other economic setting, and allow us to efficiently conduct individual-level tests of both broad and narrow bracketing with minimal assumptions about preferences. We nevertheless obtain only slightly higher pass rates for broad bracketing (20% in each experiment) and narrow bracketing (45% and 53%). These results establish that a plurality of subjects pass our most demanding tests of narrow bracketing, while 0-20% pass our tests of broad bracketing.

Our Shopping Experiment allows us to directly test how a person brackets by inducing their payoff function over fictitious goods. Decision 1 consists of two "stores" that sell fictitious apples and oranges at different relative prices, and the subject has fixed budget at each store. The subject's payoff is determined according to an induced payoff function from the final bundle of apples and oranges bought in all stores in the decision. In this design, narrow and broad bracketing each make direct and distinct predictions – and these predictions are the same for every subject who prefers more money to less. In fact, we also obtain point predictions for intermediate cases of partial-narrow bracketing (Barberis *et al.*, 2006; Rabin & Weizsäcker, 2009) given the degree of narrow bracketing; we can thus identify where a partial-narrow bracketer lies between the two extremes of broad and narrow bracketer from her choices. Yet our experiment suggests that broad and narrow bracketing explain most of the population, with 88% of subjects measured to parameter ranges consistent with one or the other (24% and 64%, respectively).

The powerful, individual-level tests of both narrow and broad bracketing in our experiments provide a comprehensive picture of how people bracket their choices. While the close adherence to narrow bracketing by a large fraction of our subjects supports the conventional wisdom among behavioral economists, we also find heterogeneity in bracketing that varies across economic settings. Existing work that attributes heterogeneity in choices solely to preferences may thus be misattributing some of this heterogeneity, leading to erroneous conclusions about preferences and welfare. Our results also provide the first evidence that intermediate degrees of partial-narrow bracketing are not useful for describing many people. Nonetheless, we find violations of both narrow and broad bracketing as well as some people consistent with neither of the two. The lack of evidence for intermediate cases suggests that narrow bracketing may arise due to unawareness of how to bracket broadly, lack of appreciation for the payoffs of doing so, or as a heuristic to simplify decision-making, rather than an intermediate bias that pulls a decision-maker away from the broad bracketing benchmark and towards narrow bracketing.

The paper proceeds as follows. Section 2 exposits formal models of broad, narrow, and partial-narrow bracketing, our revealed preference conditions for each model, and simple implications of these tests that we test in our experiments. We introduce our experimental designs in Section 3, the results of our experiments in Section 4, and discuss their implications in Section 5. We discuss other work on choice bracketing in Section 6 and also provide a detailed comparison of our design and identification strategy to Tversky & Kahneman (1981) and Rabin & Weizsäcker (2009).

2. Theory

We start by introducing formal tests of broad, narrow, and partial-narrow bracketing that we will apply to our experiment, but can also be deployed in wide variety of economic settings. A reader uninterested in formal theory may skip to Section 3, where we explain how our tests apply to our experiment.

We consider a data set describing the choices of a decision-maker (DM) who faces decision problems involving bundles of goods in $\mathbb{R}^n_+$. Each decision potentially consists of multiple concurrent parts, but consumption following a decision problem only depends on the final bundle of goods obtained from all parts in that decision. We index each decision by t and each part by k.

We observe a finite sequence of T choices from decision problems

$$\mathcal{D} = \left\{\left(x^{t,k}, B^{t,k}\right)\right\}_{(t,k)}$$

with $x^{t,k}$ denoting the alternative selected and $B^{t,k}$ indicating the feasible set from part k of decision t. The alternative consumed for a given decision t is labeled x^t and is the sum of the choices in each part k of t:

$$x^t = \sum_{\{(t',k'):t'=t\}} x^{t',k'}.$$

We assume that each decision problem is faced independently from and in the same set of economic circumstances as the others.[2] Thus there are no complementarities across decision problems – only among parts within the same decision problem.[3] For instance, in each of our experiments, one of T decisions is randomly selected and paid out. A real world analogue consists of a scanner data set with purchases at different stores (parts — different k) in a

[2]The order of the parts does not matter, and in our experiment it will be varied. We assume that the DM behaves identically in decisions t and s when $B^{t,k} = B^{s,\pi(k)}$ for any permutation π over all the parts corresponding to decision t.

[3]This is analogous to assumptions used in empirical revealed preference analysis Crawford & De Rock (2014) and in empirical tests of decision theoretical axioms and models that use more than one decision.

given time period (t), where the DM consumes these purchases after buying at all stores but before the next period.

We are interested in understanding how the DM takes into account the other parts in a decision when choosing in each part. In particular, does she take into account that good i in part k is a perfect substitute for good i in part k'? Following the behavioral economics literature, we consider three models of how a person brackets the parts of an underlying decision.

Neoclassical economics considers rational DMs who take the interaction between all the parts fully into account. We call this *broad bracketing.* Let B^t denote the feasible set of final bundles for the decision problem,

$$B^t = \left\{ \sum_{\{(t',k'):t'=t\}} x^{t',k'} : x^{t',k'} \in B^{t,k} \right\}.$$

Formally, the data set is *rationalized by broad bracketing* if there exists an increasing utility function $u : \mathbb{R}^n_+ \to \mathbb{R}$ so that x^t maximizes u over B^t,

$$x^t = \arg\max_{x \in B^t} u(x)$$

for every $t \in T$. That is, a broad bracketer recognizes that her parts define an underlying budget set, and she chooses in all parts to maximize some utility function over the underlying budget set.

A number of experiments have documented that subjects fail to broadly bracket. Instead, they propose that agents optimize within each part in isolation. That is, they ignore the possibility of substituting across parts. We follow the literature in calling this *narrow bracketing.* Formally, the data set is *rationalized by narrow bracketing* if there exists an increasing $u : \mathbb{R}^n_+ \to \mathbb{R}$ so that

$$x^{t,k} = \arg\max_{x \in B^{t,k}} u(x)$$

for every part (t,k). That is, the DM has an underlying preference over alternatives but maximizes this preference part-by-part.[4]

[4]A narrow bracketer acts as if she perceives alternatives correctly and maximizes a well-behaved preference over them, but misperceives the budget set. In contrast, a DM who misperceived correlation, e.g. Eyster

Finally, we consider the possibility of an intermediate model lying between the two extremes. We focus on one example, called *partial-narrow bracketing.* Given $\alpha \in [0,1]$, we say that the data set is *rationalized by α-partial-narrow bracketing* if there exists an increasing $u : \mathbb{R}^n_+ \to \mathbb{R}$ so that

$$x^{t,k} = \arg\max_{x \in B^{t,k}} \left[\alpha u\left(x\right) + (1-\alpha) u \left(\sum_{\{(t',k'):t'=t,k'\neq k\}} x^{t',k'} + x \right) \right]$$

for every part (t,k). The partial-narrow bracketing model is inspired by Barberis *et al.* (2006) and Barberis & Huang (2009), as adapted by Rabin & Weizsäcker (2009, p. 1513).[5] The DM takes into account her choices in the other parts but only partially. This is captured in the broadly bracketed utility term, which is weighted by $(1-\alpha)$. The remaining weight is attached to her payoff from her choice in the current part, as if there were no other parts. When $\alpha = 0$, the model coincides with broad bracketing, and when $\alpha = 1$, it coincides with narrow. The partial-narrow bracketing model defines a "personal equilibrium" (PE) of an intrapersonal game (Kőszegi & Rabin, 2006). We can strengthen the model's predictive power by assuming that the DM selects her "preferred personal equilibrium" (PPE): the PE that maximizes the α-weighted average of the sum of narrowly-bracketed utilities and the overall broadly bracketed utility (i.e. the partial-narrow objective function).[6]

In real world settings, we observe neither the utility function u nor the degree of partial-narrow bracketing α. Nonetheless, the three models have distinct testable implications for observable choices. We turn now to describing these implications.

& Weizsäcker (2016) or Ellis & Piccione (2017), perceives the choice set correctly but misperceives the alternatives themselves.

[5]The first two papers consider an average of the certainty equivalents, $u^{-1}(E[u(x^t)]) + b_0 \sum_{\{(t',k'):t'=t,k'\neq k\}} v^{-1}(E[v(x^{t,k})])$. Unlike our specification, they allow for the narrow utility function to differ from the broad utility function and assume expected utility. The algorithm in Theorem 1 can be adapted to allow for different narrow and broad utility functions ($u \neq v$) at the cost of less predictive power. Our specification is closest to Rabin & Weizsäcker (2009, p. 1513), although they also assume expected utility.

[6]In this problem, a collection of bundles from each choice set the that jointly maximizes the partial-narrow objective (known as a "choice acclimating personal equilibrium" in Kőszegi & Rabin (2007)) is also a PE and is thus a PPE. As a result, standard tools can be used to determine that a maximizer exists under standard assumptions on u and $\{B^{t,k}\}_k$.

Even though we do not observe u, observe that the subject's choice reveals her preference among a set of alternatives determined by how she brackets and that may or may not equal her actual budget set. We use this observation to design direct tests of bracketing that compare choices in pairs of decisions or parts of decisions. To conduct a test of narrow bracketing, we compare a part of a first decision problem to the exact same choices when it is part of another decision (for example, as its own standalone decision). To conduct a test of broad bracketing, we compare two decisions for which the feasible aggregate alternatives overlap. We formalize the above predictions as follows, specialized to the case where each decision has at most two parts.

Prediction 1 (NB-WARP)**.** Suppose $\mathcal{D}$ is rationalized by narrow bracketing.
If $(x^{t,k}, B^{t,k}), (x^{t',k'}, B^{t',k'}) \in \mathcal{D}$ and $x^{t',k'} \in B^{t,k} \subseteq B^{t',k'}$, then $x^{t,k} = x^{t',k'}$.

Prediction 2 (BB-WARP)**.** Suppose $\mathcal{D}$ is rationalized by broad bracketing.
For decisions t, t' in $\mathcal{D}$, if $x^{t'} \in B^t \subseteq B^{t'}$, then $x^t = x^{t'}$.

These two predictions, as the names suggest, reflect the appropriate manifestations of WARP in our setting. Narrow bracketing requires that WARP holds when comparing any pair of parts of decisions, even when they are part of economically different decisions. Broad bracketing implies that WARP holds at the decision problem level, comparing final bundles that are feasible in both aggregate budget sets.

The next prediction reflects the appropriate manifestation of monotonicity in our setting.

Prediction 3 (BB-Mon)**.** Suppose $\mathcal{D}$ is rationalized by broad bracketing.
For any decision t in $\mathcal{D}$ and any $y \in B^t$, $y \geq x^t$ if and only if $y = x^t$.

BB-Mon requires that the subject chooses on the frontier of her aggregate budget set in a given decision. For a decision with two Walrasian budget sets with different price ratios, this implies that the DM makes an extreme allocation in at least one of the budgets. In the part where good i is relatively cheaper, she must exhaust the budget on good i, or consume none of good i in the other part. Otherwise, she forgoes the opportunity to consume more of each good.[7]

The above predictions are necessary but not sufficient conditions for the various types of bracketing. To obtain such conditions, we extend the logic of the above to include indirect implications, in the same manner that Strong Axiom of Realed Preference (SARP) extends the logic of the Weak Axiom (WARP). In our setting, a bundle x is *directly revealed preferred to* y in the data set $\mathcal{D}$, written $xP^{\mathcal{D}}y$, if there exists a budget set B so that $(x, B) \in \mathcal{D}$ and $y \in B\backslash\{x\}$, and the data set $\mathcal{D}$ *satisfies SARP* if the binary relation $P^{\mathcal{D}}$ is acyclic. Theorem 1 combines these indirect implications to provide necessary and sufficient conditions for a given data set to be consistent with a particular form of bracketing.

Broad bracketing implies that the DM maximizes her underlying preferences on a feasible set defined at the decision-problem level, i.e. over B^t rather than $B^{t,k}$. With this specification of budget sets, one can apply SARP to the ancillary data set

$$\mathcal{D}^{BB} = \left\{\left(x^t, B^t\right)\right\}_{t\in T}.$$

These tests may need to be adapted (as in Forges & Minelli 2009) because B^t is typically non-linear even if each $B^{t,k}$ is linear.

Similarly, narrow bracketing implies that the DM maximizes an underlying preference on a feasible set equal to each part, i.e. over $B^{t,k}$ rather than B^t. As above, one can apply SARP to an ancillary data set. In this case, the data set is

$$\mathcal{D}^{NB} = \left\{\left(x^i, B^i\right)\right\}_{i=1}^{N}$$

where each data point (x^i, B^i) corresponds to a data point from a part and vice versa; formally, there exists a bijection b from $\{1, \ldots, N\}$ to the pairs (t, k) of decisions and parts so that $(x^i, B^i) = \left(x^{b(i)}, B^{b(i)}\right)$ for every i. That is, $\mathcal{D}^{NB}$ is the original data set but where each part is treated as a separate, independent observation.

We clarify that applying SARP to the data sets $\mathcal{D}^{BB}$ and $\mathcal{D}^{NB}$ captures the complete implications for broad and narrow bracketing, respectively.

[7]For narrow bracketing, the equivalent condition requires that the subject consumes on the budget line. Since this is forced in our experimental implementation, we do not formally include it here.

Theorem 1. *The following are true:*

(i) The data set $\mathcal{D}^{BB}$ satisfies SARP if and only if $\mathcal{D}$ is rationalizable by broad bracketing,

(ii) The data set $\mathcal{D}^{NB}$ satisfies SARP if and only if $\mathcal{D}$ is rationalizable by narrow bracketing, and

(iii) There exists a decidable algorithm that outputs 1 if and only if $\mathcal{D}$ is rationalizable by α-partial-narrow bracketing.

Theorem 1 provides standard, albeit indirect, tests for broad and narrow bracketing. These conditions are readily applied and provide tests for the models of bracketing with only minimal restrictions on preferences. Moreover, they can be easily augmented to include additional structure on preferences. For instance, we impose that preferences are symmetric when we apply these tests to our experiments.

The algorithm that tests α-partial-narrow bracketing is described in Appendix A. It reduces the problem of determining whether $\mathcal{D}$ is consistent with α-partial-narrow bracketing to that of a standard linear programming problem. We elaborate on the ideas behind the partial-narrow test here, and provide a formal proof in Appendix A.

The algorithm starts by mapping the DM's choice in each part into both the narrow bundle and final (broad) bundle it generates. It then performs a revealed preference analysis on these bundle pairs that tests for the existence of a u such that utility of each pair is the additive sum of the utilities of the narrow and final bundles each evaluated according to u and weighted α and $1-\alpha$ respectively. The choices in the other parts determine the overall feasible set that applies for a part. This test is based on Clark's (1993) test of expected utility over lotteries but where utilities are weighted by α and $1-\alpha$ instead of by objectively-given probabilities. The PPE version of the test then checks for the existence of dominating combinations of choices across parts.

We can make tighter predictions by imposing more structure on the utility function that rationalizes the data. The main piece of structure we impose is that the DM treats each of the dimensions symmetrically, as induced by our experimental designs.[8] The revealed preference

[8]Formally, the rationalizing utility function u invariant to permutations: for any bijection $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$, $u(x_1, \ldots, x_n) = u\left(x_{\pi(1)}, \ldots, x_{\pi(n)}\right)$.

tests above are easily adapted to include symmetry. Specifically, symmetry implies that if $(x, y)\, R\, (x', y')$, then $(y, x)\, R\, (x', y')$, $(x, y)\, R\, (y', x')$, and $(y, x)\, R\, (y', x')$. These restrictions can be straightforwardly combined with each of NB- and BB-SARP and our partial-narrow bracketing tests (PNB Tests). In Appendix B, we derive some immediate implications of symmetry. For instance, a narrow bracketer should purchase at least as much of the cheaper good in each part (NB-Sym). Similarly, a broad bracketer should purchase at least as much of the good that has a cheaper overall cost in each part (BB-Sym). By inducing and assuming symmetric preferences, we will increase the predictive power of all models, especially partial-narrow bracketing.

3. Experimental Design

We design and conduct three experiments to test the models of bracketing in different domains of choice and to measure α in an induced-value experiment. In each experiment, a participant faces five decision rounds, each consisting of one or two parts.[9] Each part consists of all feasible integer-valued bundles of two goods obtained from a linear (or in one case, a piece-wise linear) budget set. At the end of the experiment, exactly one decision is randomly selected and all goods purchased in its parts determine the aggregate bundle that determines payments.[10] By design, there are no complementarities across decisions. We implement this experimental design to study choice bracketing in three domains of interest: portfolio choice under risk (Risk), a social allocation task (Social), and a consumer choice experiment in which we induced subjects' payoff function (Shopping).

In the Risk Experiment, each part of every decision asks the subject to choose an integer allocation of tokens between two assets. Each asset pays off on only one of two equally likely states: Asset A (or C) pays out only on a die roll of 1-3 whereas Asset B (or D) pays out only on a die roll of 4-6. The payoff of each asset varies across decision problems and across

[9]Compared to choice-from-budget-set experiments like Choi *et al.* (2007), each subject faces fewer rounds of decisions in our experiment and each budget set is coarser to allow us to conduct the experiment on paper. We made this design choice to minimize potential for learning to bracket, and also to slow down subjects' progression through the experiment to encourage slow and deliberate decision-making.

[10]In the Social Experiment we modify this procedure slightly: one person-decision pair per anonymous group of two subjects is randomly selected to determine payments of another anonymous group of two subjects.

parts. Because each decision problem uses assets with two equally likely states, preferences over portfolios of monetary payoffs for each state should be symmetric across states.

In the Social Experiment, each part of every decision asks the subject to choose an integer allocation of tokens between two anonymous other subjects, Person A and Person B. The value of each token to A and B varies across decision problems and across parts. Because the two recipients are anonymous, we expect preferences to be symmetric across money allocated to A versus B.[11]

In the Shopping Experiment, each part of every decision asks the subject to choose a bundle of integer quantities of fictitious "apples" and "oranges" subject to a budget constraint. The monetary payment for the experiment is calculated from the final bundle in the round that counts according to the function $\$\text{pay} = \frac{2}{5}\left(\sqrt{\#\text{apples}} + \sqrt{\#\text{oranges}}\right)^2$.[12] Any subject who prefers more money to less will wish to maximize this payoff function regardless of their underlying utility function, but how they attempt to do so will depend on how they bracket in each decision. As a result, this induced-value design allows us to test models of choice bracketing without having to estimate utility. Moreover, by assuming that their utility function over goods takes the known form generated by the payoff function, that is, $u(\#\text{apples}, \#\text{oranges}) = \frac{2}{5}\left(\sqrt{\#\text{apples}} + \sqrt{\#\text{oranges}}\right)^2$, α can be identified from any decision that consists of two parts with different price ratios (Theorem 2 in Appendix A).

Our experimental budget sets, summarized in Table 1, are designed to allow us to conduct our revealed preference tests in our Risk and Social Experiments, and to conduct analogous tests that make use of the induced payoff function in our Shopping Experiment. Throughout, we refer to part k of decision t as Dt.k, or Dt if a round has only one part.

We designed our experiment to enable the comparisons across decisions needed to perform our NB- and BB-WARP tests. A one-part round, D5, is identical to D1.1 and to D3.2 in Risk and Social Experiments, that is, $B^{1,1} = B^{3,2} = B^{5,1}$. Similarly, the one-part round Decision 4 is identical to D1.2, that is $B^{1,2} = B^{4,1}$. This allows us to test by NB-WARP

[11] Note that the budgets in the Risk and Social Experiments were designed to be identical up to the translation across domains.

[12] Subjects were provided with a payoff table (Appendix E, Figure 20) to calculate the earnings that would result from any possible final bundle, so could maximize earnings without having to manually compute this function.

		Risk			Social			Shopping		
Decision	Part	I	Asset A/C	Asset B/D	I	V_A	V_B	I	p_a	p_o
D1	1	10	$(\$1, \$0)$	$(\$0, \$1.20)$	10	\$1	\$1.20	8	2	1
	2	16	$(\$1, \$0)$	$(\$0, \$1)$	16	\$1	\$1	24	2	2
D2	1	14	$(\$2, \$0)$	$(\$0, \$2)$	14	\$2	\$2	32	2	1 (for 1st 8), 2
D3	1	10	$(\$1, \$0)$	$(\$0, \$1)$	10	\$1	\$1	30	3	3
	2	10	$(\$1, \$0)$	$(\$0, \$1.20)$	10	\$1	\$1.20	24	3	2
D4	1	16	$(\$1, \$0)$	$(\$0, \$1)$	16	\$1	\$1	12	1	1
D5	1	10	$(\$1, \$0)$	$(\$0, \$1.20)$	10	\$1	\$1.20	48	6	4

I: income for a part (in tokens)

$(\$x, \$y)$ indicates one unit of asset pays \$x/\$y if the die roll is 1-3/4-6

V_A: value/token to A
V_B: value/token to B

p_a: price/apple
p_o: price/orange

TABLE 1. Experimental Tasks

by comparing each of D1.1 and D3.2 to D5, and by comparing D1.2 to D4: NB-WARP requires $x^{1,1} = x^{3,2} = x^5$ and $x^{1,2} = x^4$. In order to test BB-WARP, Decision 2 is a one-part round that lies on top of the aggregate budget set of the two-budget-set round Decision 1, $B^{1,1} + B^{1,2} \subset B^{2,1}$. Given the nature of this overlap, BB-WARP requires that if $x_B^2 \geq 6$ then $1.2x_B^{1,1} + x_D^{1,2} = 2x_B^2$. These tests can falsify narrow and broad bracketing without assumptions on preferences other than completeness and transitivity.

Sessions took place in two experimental economics labs in Canada from June 2019 to February 2020. Subjects were recruited from the labs' student participant pools. The experiment was conducted on paper. After the instructions had been read aloud and subjects were given the opportunity to ask questions privately, participants completed a brief comprehension quiz, and the experimenter individually checked answers and explained any errors. Only one round was handed out to a subject at a time, and that round was collected before the next round was handed out. Subjects indicated each choice by highlighting the line corresponding to their choice for each part using a provided highlighter. Subjects were allowed no other aids at their desk when making choices. The order of decisions was varied across participants.[13] Instructions, experimental materials, and details of the experimental procedure are provided in Appendix E.[14]

4. Results

This section reports the results of our experimental tests of the models of bracketing. For each test, we also compute results allowing for one or two "errors" relative to its requirements. We define an error as how far we would need to move a subject's allocations for them to pass that test, measured in lines on the decision sheet(s).[15] For instance, in Risk and Social, a subject is within one error of passing a test if shifting one token from one asset/person to the other in a single part could lead them to pass. They are within two errors of passing a test if shifting two tokens from one asset/person to the other, either in the same part or in different parts of any decision, could lead the revised allocations to pass. For predictions that require Walrasian budget sets, we modify the tests to account for discreteness in our experiment. We visually represent all the data underlying our tests in Appendix B.

4.1. **Risk and Social Experiments: How subjects behave.** Consider Decision 1 in the Risk Experiment. Since Asset B in the first part is a perfect substitute for Asset D in the second, logic familiar from comparative advantage requires that the subject first purchases it in the part where it has a lower opportunity cost. Consequently, a broad bracketer necessarily allocates all of her wealth in the first part to Asset B before she allocates any to Asset D in the second (as required by BB-Mon). If preferences are symmetric across the two equally-likely states, then broad bracketing further requires that $x_A^{1,1} = 0$. With risk aversion, broad bracketing makes even stronger predictions. The allocation $x_A^{1,1} = 0$ and $x_C^{1,2} = 14$ obtains a risk-free return of \$14, and any other feasible allocation results in a first- or second-order

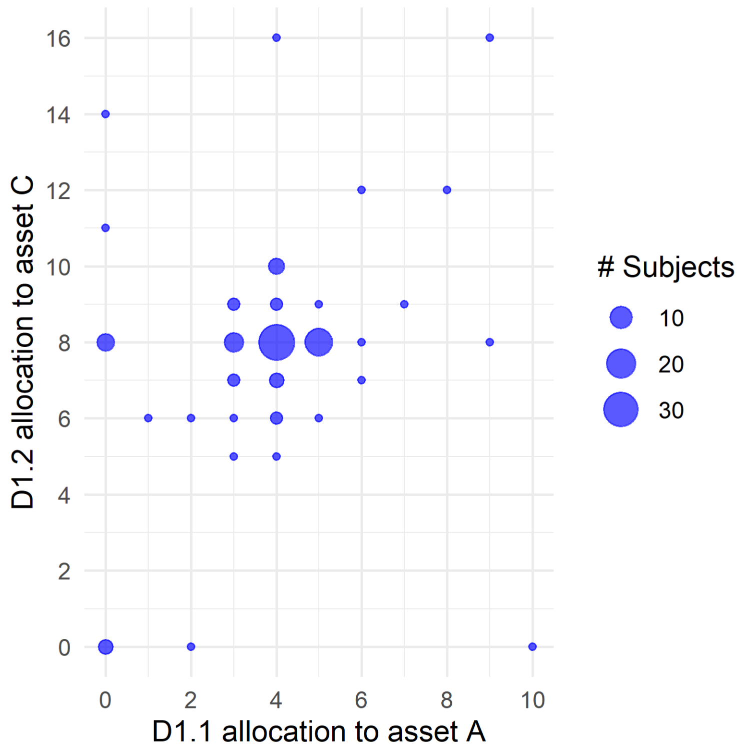
[13]We perform every pairwise test of the effect of the order in which decisions were faced separately for each experiment. Every such rank-sum test yields a p-value > 0.10 with the exceptions of D1.2 and D2 in the Shopping Experiment ($p = 0.07, 0.02$ respectively); we attribute the latter difference to sampling variation, while the former may reflect learning, which we discuss in 4.5.

[14]The three experiments were separately pre-registered through the Open Science Framework. Our analysis follows our plans with only minor modifications for expositional purposes that do not affect the interpretation of our results.

[15]Example decision sheets are provided in Figures 12, 13, 16, and 19 in Appendix E.

stochastically dominated distribution over returns at the decision-level. In contrast, a narrow bracketer facing Decision 1 allocates at least half of her budget to Asset B in Part 1. If she is also risk-averse, then she allocates exactly half her budget in Part 2 to each asset.

FIGURE 1. D1 allocations in Risk



To compare how our subjects perform to this benchmark, Figure 1 plots the joint distribution of their allocations in D1. The x-coordinate describes their allocation to Asset A in D1.1, and the y-coordinate their allocation to Asset C in D1.1. The above discussion shows that broad bracketers will have allocations on the y-axis, and a risk-averse broad bracketer will select the allocation $\left(x_A^{1,1}, x_C^{1,2}\right) = (0, 14)$. Narrow bracketers will select allocations with $x_A^{1,1} \leq 5$, and risk-averse narrow bracketers will also select $x_C^{1,2} = 8$. The plot shows that few subjects are close to consistent with broad bracketing. Indeed, only a single subject makes an allocation close to the prediction of risk-averse broad bracketing. Narrow bracketing, even with risk aversion, does much better and is consistent with the modal allocation of $\left(x_A^{1,1}, x_C^{1,2}\right) = (4, 8)$. Three subjects select $\left(x_A^{1,1}, x_C^{1,2}\right) = (0, 0)$, as a risk-seeking agent would do under either narrow or broad bracketing.

FIGURE 2. D3.2 vs. D5 allocations in Risk

In our design, narrow bracketing makes strong predictions across decisions that have a part in common. For instance, D3.2 and D5 both ask subjects to allocate 10 tokens between identical assets with identical prices. A narrow bracketer would make the same allocation in each of the two parts. To show this, Figure 2 plots each subject's allocation in D3.2 against their allocation in D5. The x-coordinate describes their allocation to Asset C in D3.2, and the y-coordinate their allocation to its counterpart, Asset A in D5. A narrow bracketer's choices will fall on the 45-degree line as required by NB-WARP, and the farther the allocations are from that line the farther the subject is from narrow bracketing. In the plot, we can see that this prediction of narrow bracketing holds exactly for 54 of the 99 subjects.

4.2. **Risk and Social Experiments: Revealed Preference Tests of NB-WARP, BB-WARP, and BB-Mon.** We begin by performing the simple, direct tests of bracketing developed in Section 3: NB-WARP, BB-WARP, and BB-Mon (Table 2).

First, we show that very few subjects are broad bracketers. There is only a single pair of decisions (1 and 2) where choices could directly violate BB-WARP. For that pair, we test BB-WARP by comparing the final bundle for Decision 1 to the final bundle for Decision 2: for any choice of x^2 in Decision 2 with $x^2_A \leq 8$, the same final bundle can be achieved in

# errors	Risk 0	1	2	Social 0	1	2
NB-WARP (D1.1 and D5)	56	76	89	45	70	77
NB-WARP (D1.2 and D4))	56	74	81	63	78	82
NB-WARP (D3.2 and D5)	54	76	83	49	75	80
NB-WARP (D1.1 and D3.2)	49	76	85	51	83	87
NB-WARP (all)	29	44	61	28	54	64
BB-WARP (D1 and D2)	13	20	87	16	20	94
BB-Mon (D1)	12	13	15	14	14	14
BB-Mon (D3)	14	16	18	17	17	18
BB-Mon (both)	7	8	10	12	12	12
# subjects		99			102	

Entries count the # of subjects who pass test at the listed error allowance.

Table 2. Tests of NB-WARP and BB-WARP

Decision 1. In each of Risk and Social, only 20% of subjects are within one error of passing BB-WARP.[16]

We next test BB-Mon in Decisions 1 and 3. Since good i in the first part is a perfect substitute for good i in the second, logic familiar from comparative advantage requires that the subject first purchases the good in the part where it has a lower opportunity cost. As a consequence, she must make at least one extreme allocation among the two parts. For instance in Decision 1, it requires that either $x_A^{1,1} = 0$ or $x_D^{1,2} = 0$. Even fewer subjects are consistent with BB-Mon than with BB-WARP. In Risk and Social respectively, we find that 8% and 12% of subjects are within one error of passing BB-Mon in both decisions. Looking separately at Decisions 1 and 3, between 13% - 17% of subjects are within one error of passing BB-Mon, so economically-significant bracketing-related dominance violations are the norm rather than the exception among our subjects.

[16]Allowing for two errors substantially raises pass rates of BB-WARP in Risk and Social. While the vast majority of subjects, 66% in Risk and 84% in Social, allocate equally between assets/people A and B in Decision 2, if a subject allocates $x_A^2 > 8$, then BB-WARP is trivially satisfied.

All told, the BB-WARP and BB-Mon tests provide evidence showing that 80%-92% of subjects are not broad bracketers. These rates of violations of broad bracketing are qualitatively similar but higher to those found by Tversky & Kahneman (1981) (73%) and Rabin & Weizsäcker (2009) (28%-66%), and very close to the structural estimates of the latter (89%). In these prior experiments, each part consisted of a pairwise choice, so failures of broad bracketing are detected only for a particular range of risk preferences. In contrast, there are many ways a subject could reveal their failure to bracket broadly in our experiments, which give us more power to detect failures.

While previous work only falsifies broad bracketing, our design allows us to test narrow bracketing as well. We test NB-WARP by comparing the allocation in each of the two parts that appear multiple times. Specifically, D1.1, D3.2, and D5 are all the same, as are D1.2 and D4. NB-WARP requires that she makes the same choice in a given part whenever it appears. Figure 2 illustrates the behavior underlying the NB-WARP test comparing D3.2 vs. D5 test in Risk. The large mass of subjects on the 45-degree line shows that many but not all subjects pass NB-WARP.

Our tests find that NB-WARP does much better than either BB-WARP or BB-Mon. Between 75-77% of subjects in Risk and 69-81% of subjects in Social are within one error of passing each of the pairwise NB-WARP tests. If we jointly conduct all possible NB-WARP tests and allow for one error, 44% and 53% pass in Risk and Social respectively.[17] While narrow bracketing does much better, some subjects are inconsistent with both NB-WARP and BB-WARP, even after allowing for one error. This could result from some other form of bracketing, such as partial-narrow, or simply noisy behavior.

Result 1. *When allowing for one error, 45% and 53% of subjects pass NB-WARP, 20% and 20% pass BB-WARP, and 16% and 17% pass BB-Mon in the Risk and Social Experiments, respectively.*

By design, symmetry is a natural restriction in our experiments – the two states are equally likely in Risk, and the two other individuals are anonymous in Social. Thus, we can conduct

[17] In contrast, randomly-generated choices have only a 0.4% chance of being within one error of passing all NB-WARP tests.

similar decision-by-decision tests of narrow, broad, and partial-narrow bracketing while assuming symmetric preferences – using the restrictions implied by symmetry to eliminate the need to compare across decisions. We call these tests NB-Sym, BB-Sym, and PNB-Sym, and derive and perform them in Appendix B. Unsurprisingly, we observe very similar pass rates for our NB-Sym tests of narrow bracketing as we did for NB-WARP. However, we observe lower pass rates of BB-Sym – no subjects in Risk pass it when allowing for one error, and only 10% of subjects pass it in Social. Pass rates are much higher for PNB-Sym – when allowing for one error across both Decisions 1 and 3, 89% and 75% pass in Risk and Social respectively – but this includes only 23% and 21% additional subjects beyond those that pass either NB-Sym or BB-Sym.

We next conduct our revealed preference tests of symmetric versions of the three models considered using the entire set of decisions for each subject. These tests are demanding. For example, all three tests make point predictions in Decisions 2 and 4. Random behavior has less than a 10^{-5} chance of passing either BB- or NB-SARP with two errors, and less than a 10^{-2} chance of passing the PNB Algorithm with two errors (see Appendix D). Table 3 shows how many subjects pass each test.

	Risk			Social		
# errors	0	1	2	0	1	2
NB-SARP	23	34	43	15	36	44
BB-SARP	0	0	0	8	10	10
PNB-PPE	49	59	71	31	58	69
PNB-PE	51	63	75	39	67	76
# subjects		99			102	

Entries count the # of subjects who pass each test at the listed error allowance.

Table 3. Full Tests of Symmetric Models

We compare results of the tests that allow for up to two errors. Strikingly, we find that no subjects pass BB-SARP in Risk while 10% of subjects in Social pass it. Substantially more pass NB-SARP – 43% of subjects in each experiment. While notably fewer subjects pass BB-SARP than either BB-WARP or BB-Mon, a similar number of subjects pass all NB-WARP restrictions with one error as pass the more demanding NB-SARP with two errors.

These results strongly support that a plurality of our subjects are well-described as narrow bracketers and reinforce the relative rarity of broad bracketing.

Even allowing for two errors, only about half of subjects pass either BB- or NB-SARP. Our test of partial-narrow bracketing diagnoses how many of these behave consistently with intermediate degrees of bracketing. We find that 28% of subjects in Risk and 15% of subjects in Social pass the PNB-PPE test but neither BB-SARP nor NB-SARP.[18] Only 4% of subjects in Risk and 7% of subjects in Social pass the PNB-PE test but not the PNB-PPE test. This leaves 22% and 26% of subjects in Risk and Social inconsistent with all models of bracketing considered.

Result 2. *When allowing for two errors relative to each test, 43% of subjects in both settings pass NB-SARP. Only 0% and 10% of subjects respectively pass BB-SARP, while an additional 28% and 15% of subjects pass the PNB PPE tests but neither NB- nor BB-SARP, in the Risk and Social Experiments, respectively.*

The tests reported in Result 2 require both that preferences are symmetric and that choices reveal strict preferences. Consequently, subjects with linear preferences, such as those who maximize expected or total payoffs, may fail our tests.[19] A subject with linear preferences makes the same choices in any data set regardless of how she brackets, so the presence of linear preferences does not bias our results in favor of a given type of bracketing.[20] Five subjects in Risk and two subjects in Social make choices consistent with linearity throughout the experiment. However, they have limited influence on our overall results. Out of these seven subjects, one each in Risk and in Social pass the tests for narrow bracketing. The other subject in Social is within one error of passing all tests of broad bracketing. Two other subjects in Risk always allocated their entire portfolio to Assets B/D, consistent with either risk-seeking or risk-neutral expected utility preferences. Consequently, the two pass

[18]Broad and narrow bracketing are both special cases of partial-narrow bracketing, and thus any subject who passes BB- or NB-SARP will also pass this test.
[19]Risk- or inequity-seeking preferences also predict extreme choices that are also consistent with linear preferences.
[20]Indeed, linear preferences are the unique class of preferences that can be simultaneously consistent with both narrow and broad endowment bracketing (Appendix A, Theorem 3)

BB-WARP and NB-WARP but fail NB-Sym and BB-Sym. The remaining three subjects are not within two errors of passing any of the tests.

4.3. **Shopping Experiment: Induced-Value Tests and Estimates of α.** The tests of our predictions thus far assume that utility is not observed. When utility is known, as in our Shopping Experiment, narrow and broad bracketing each make unique predictions in each decision.[21] To test the models, we compare how far each subject's choices are from each model's predictions (Table 4).

	D1			D3			Both			Full		
# errors	0	1	2	0	1	2	0	1	2	0	1	2
NB	23	23	60	53	65	69	20	21	49	15	16	40
BB	20	21	27	23	24	24	12	14	17	11	13	15
PNB-PPE	44	45	95	76	91	98	32	36	68	26	29	56
PNB-PE	44	45	95	76	93	98	32	36	70	26	29	57
# subjects	101											

Entries count the # of subjects who pass each test at the listed error allowance.

TABLE 4. Shopping Tests

Testing all implications of narrow bracketing in each of Decisions 1 and 3, 23% and 64% of subjects are respectively within one error of the predictions of narrow-bracketed maximization, while 21% are consistent in both. Allowing two errors raises pass rates to 59% for Decision 1 and 49% for both.[22] Testing the full set of implications of narrow-bracketed maximization on all choices made in the experiment, 40% of subjects are within two errors of being consistent with all implications of narrow-bracketed maximization.

In contrast, 21%, 24%, and 14% of subjects are within one error of being consistent with broadly-bracketed maximization in Decisions 1, 3, and both, respectively. When allowing for

[21]Narrow-bracketed maximization implies $x_a^{1,1} = 1$, $x_a^{1,2} = 6$, $x_a^{3,1} = 5$, and $x_a^{3,2} = 4$. Broad-bracketed maximization implies $x_a^{1,1} = 0, x_a^{1,2} = 10, x_a^{3,1} = 10$, and $x_a^{3,2} = 0$.

[22]The sharp difference between one- and two-error tests in Decision 1 but not Decision 3 results from the fact that 40 subjects selected $x_a^{1,1} = 2, x_o^{1,1} = 4$, which is the second-best available bundle from a narrow bracketer's perspective but is two lines away from the best bundle of $x_a^{1,1} = 1, x_o^{1,1} = 6$ – due to the discreteness of the budget set, one apple and five oranges is between these bundles. We note that this deviation is in the opposite direction of that predicted by broad or partial-narrow bracketing, and 32 of these subjects also selected $x_a^{1,2} = x_o^{1,2} = 6$.

two errors, those numbers remain similar. Using all decisions in the experiment, only 15% of subjects are within two errors of being consistent with all implications of broad-bracketed maximization.

Partial-narrow bracketing appears to describe few additional subjects beyond those consistent with narrow and broad bracketing in each test. In Decisions 1, 3, and both respectively, when allowing for one error, only 1%, 2%, and 1% of subjects are consistent with partial-narrow bracketing but neither narrow nor broad bracketing. When testing the full implications of each model, partial-narrow bracketing (PPE version)[23] accommodates at most one additional subject whether allowing for zero, one, or two errors.

Result 3. *In the Shopping Experiment, allowing for two errors, 40% of subjects pass narrow bracketing, 15% pass broad bracketing, and one additional subject passes partial-narrow bracketing.*

Since we induce the payoff function, we are able to measure α exactly (up to limits imposed by discretization of the budget sets) under the assumption that the induced payoff function acts as their utility function.[24] To that end, we compute the point predictions of the partial-narrow bracketing model for each $\alpha \in \{0, .01, .02, \ldots, .99, 1\}$ for Decisions 1 and 3, and obtain distinct predictions for nine distinct ranges of α. We assign each subject to the range of α for which their choices exhibit the fewest errors relative to that range's predictions.

We find that 64% of subjects are classified to a range that includes full narrow bracketing, $\alpha = 1$, and 24% are classified to a range that includes full broad bracketing, $\alpha = 0$ (Figure 3). Strikingly, no subjects are classified to the range $\alpha \in [.25, .72)$, in spite of this range comprising 47% of the parameter space considered.

[23]The PE is always unique in Decision 1, and is unique in Decision 3 except for $\alpha \in [.14, .18]$, in which case there are two PE.

[24]If instead they are a partial-narrow bracketer with $\alpha \in (0, 1)$ but apply a concave utility-for-money function on top of their induced payoff function before maximizing, then the procedure we used applied to each decision separately will recover bounds on a transformation of their true α, with the transformation depending on the shape of their utility-for-money function. However, this procedure will correctly classify subjects who are either full broad bracketers ($\alpha = 0$) or full narrow bracketers ($\alpha = 1$), and will assign intermediate values of α to subjects whose values are between the two extremes.

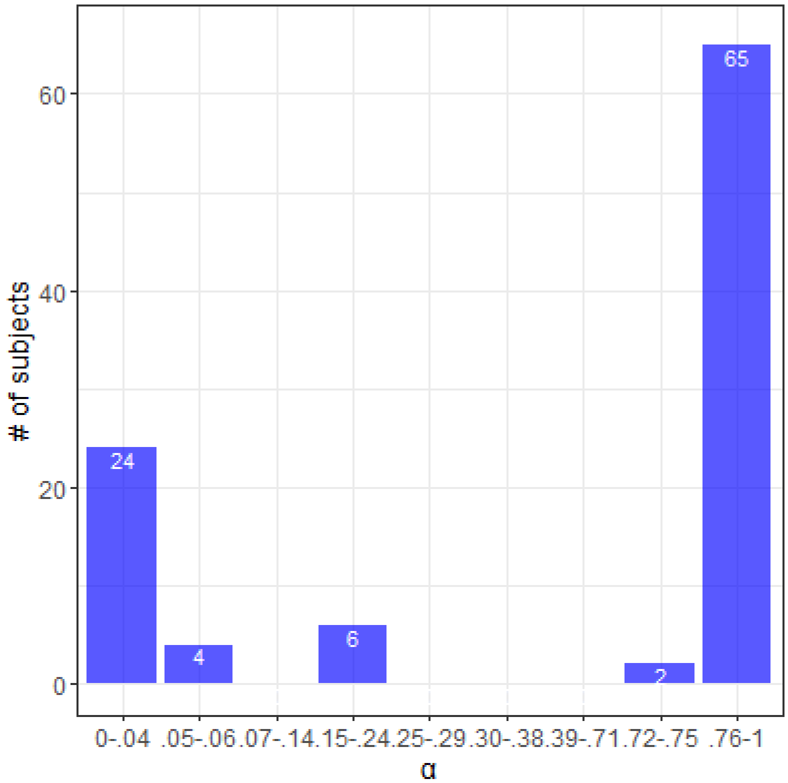


FIGURE 3. Distribution of estimated values α

Result 4. *In the Shopping Experiment, we find that 64% of subjects are best fit by* $\alpha = 1$*, 24% by* $\alpha = 0$*, and none by any* $\alpha \in (.25, .72]$.

4.4. **Classifying subjects to models.** The tests thus far do not make any adjustment for the fact that partial-narrow bracketing nests narrow and broad bracketing as polar cases, and thus can accommodate more behavior. To compare the predictive success of each model at the subject level, we use a subject-level implementation of the Selten score (Selten, 1991; Beatty & Crawford, 2011). For each subject and each model (symmetric versions for Risk and Social, using the induced payoff function for Shopping), we calculate the number of errors the subject exhibits relative to that model. Then, we calculate the number of possible choice combinations in the experiment that are consistent with that model and that number of errors, which we divide by the total number of possible combinations of choices in the experiment to compute the measure for each subject i and model $m \in \{\text{broad,narrow,partial}\}$

as $\text{predictive\_success}_{i,m} = 1 - \frac{\#\text{predictive\_area\_for\_}i,m}{\#\text{all\_possible\_choices}}$.[25] We use all choices made in the experiment to assign each subject to the model with the highest predictive success; in cases where every rationalizing model-error pair for a subject would rationalize more than one million possible combinations of choices in our experiment, we categorize them as "Unclassified".

	# Selten Score Maximized		
	Risk	Social	Shopping
Broad Bracketing	3	10	27
Narrow Bracketing	77	77	68
Partial-Narrow (PPE)	6	2	4
Partial-Narrow (PE)	0	4	1
Unclassified	13	9	1

Table 5. Classification of subjects

We find that across the three experiments, 67-78% of subjects are classified as narrow bracketers (Table 5). In contrast, 3%, 10%, and 27% of subjects are classified as broad bracketers in the Risk, Social, and Shopping Experiments respectively. However, almost no one was classified as a partial-narrow bracketer – 6%, 6%, and 5% in the Risk, Social, and Shopping Experiments were so classified, even after considering both the more predictively-precise PPE and the more permissive PE versions of partial-narrow bracketing. This suggests that partial-narrow bracketing does not help explain many subjects' behavior beyond what can be explained by its two polar cases.

Result 5. *Judging each model's fit by its predictive success, 67-78% of subjects are classified as narrow bracketers, 3-27% are classified as broad bracketers, and 5-6% are classified as partial-narrow bracketers across the three experiments.*

[25]In the Risk and Social Experiments, there are $(11 \times 17) \times 15 \times (11 \times 11) \times 17 \times 11 = 63,468,735$ possible combinations of choices. Symmetric narrow bracketing allows 6, 87, and 606 possible combinations of choices when allowing for zero, one, and two errors, respectively, whereas symmetric broad bracketing allows 12, 116, and 585 combinations of choices, symmetric partial-narrow bracketing (PPE) allows 35,797, 200,828, and 597,728 combinations, and symmetric partial-narrow bracketing (PE) allows 116,267, 619,375, and 1,725,466 combinations. Thus, a subject whose choices are consistent with partial-narrow bracketing will be classified as a partial-narrow bracketer if and only if they are sufficiently far from being consistent with both broad and narrow bracketing.

4.5. **Secondary analyses.**

Aggregate statistical analysis. Our preceding analysis focuses on subject-level tests, which reveal substantial heterogeneity in how people bracket. In contrast, an aggregate analysis masks this heterogeneity. While a notable number of subjects fail each NB-WARP test, the data fail to reject the null hypothesis that average subject behavior satisfies NB-WARP. Matched-pairs t-tests that separately compare D1.1 to D5, D3.2 to D5, and D1.2 to D4 in Risk and Social ($p = 0.84, 0.16, 0.97$ respectively for Risk, $p = 0.38, 0.13, 0.19$ for Social) fail to reject. Similarly, we strongly reject the null hypothesis that average subject behavior satisfies BB-WARP despite a substantial minority of subjects satisfying it at the individual-level, at least in the Social Experiment. A matched-pairs t-test comparing aggregate bundles in D1 to D2 in Risk and Social ($p < 0.001$ for Risk, $p < 0.001$ for Social) strongly rejects the null hypothesis.[26] These results are generally consistent with our finding from individual-level analysis that most subjects are best described by narrow bracketing. Nonetheless, the aggregate tests fail to uncover heterogeneity in behavior.

For the Shopping Experiment, we separately compare allocations in each of D1.1, D1.2, D3.1, and D3.2 to the point predictions of each of narrow and broad bracketing (eight t-tests total, $p = 0.06$ for D1.1 test of narrow bracketing, $p < 0.001$ for all other tests), and strongly reject both narrow and broad bracketing. If we ended here, we might incorrectly find in favor of partial-narrow bracketing, even though we only classify 5 subjects in the Shopping Experiment to that model. These tests further reinforce that our main individual-level analysis detects meaningful heterogeneity across subjects, whereas aggregate tests can produce misleading conclusions.

Differences across experiments. One unexpected finding, apparent from the results provided thus far, is that the rate of broad bracketing varies widely across experiments: we classify as broad bracketers 27% of subjects in the Shopping, 10% in Social, and 3% in Risk. We find the difference between Social and Risk is particularly noteworthy since the two experiments are designed to be exactly the same except for the underlying domain. The probability that

[26]This test uses only those subjects who selected a bundle in D2 that would be feasible in the decision-level D1 choice set.

each observed difference in broad bracketing rates between experiments arose purely from sampling variation is small ($p = .08$ for Risk vs. Social, $p \leq 0.01$ for each of Shopping vs. Social, Shopping vs. Risk, Fisher's exact tests).[27]

Learning to broadly bracket? We find some weak evidence from the Shopping Experiment that subjects tend to make decisions that are more consistent with broad bracketing in the second two-budget-set round they face. We compute the number of errors a subject makes in each of Decisions 1 and 3 relative to optimal decisions implied by broad bracketing with the induced payoff function. 52 subjects deviate less severely from broad bracketing in the second two-part round than in the first, while 35 exhibit the opposite pattern ($p = .09$, sign test); the average difference is .99 fewer errors in the second two-part round ($p = .01$, paired t-test).

However, we do not find analogous evidence in the Risk and Social Experiments. In these experiments, the difference in the deviation from the broad bracketing benchmark implied by BB-SARP with symmetric preferences is approximately zero from the first to the second two-budget round. The average difference is -.16 in Risk ($p = .70$, matched-pairs t-test) and 0.53 in Social ($p = .25$). In Risk, 48 subjects exhibit a smaller deviation from broad bracketing in the first two-budget round compared to the second whereas 44 exhibit the opposite pattern ($p = .75$, sign test); the analogous numbers are 45 and 43 in Social 36 ($p = .92$).

5. Discussion

Narrow bracketing is a crucial ingredient of applications of prospect theory to explain behavior, both in experiments (Kahneman & Tversky, 1979) and outside of them (Camerer 2004, Table 5.1). Since it is impossible to avoid all other risks, a broad bracketer will not exhibit noticeable loss aversion in a moderate-stakes risk-taking opportunity; outcomes of larger-stakes risks will primarily determine what counts as a gain or loss outcome rather than the risk being evaluated (Barberis *et al.*, 2006). In other domains, bracketing plays an equally central role in understanding and modeling behavior, but assumptions about

[27]In addition, in Social, a subject's choices determine the payment of other subjects, which, a priori, we worried may bias subjects in this experiment against broad bracketing as compared to the Risk experiment.

it are usually implicit. For example, Sobel (2005, p. 400) remarks that in experiments studying social decisions, a subject should "maximize her monetary payoff in the laboratory and then redistribute her earnings to deserving people later... if the concern for inequity was 'broadly bracketed.'" However, the literature he surveys finds ample evidence consistent with subjects caring about narrowly-bracketed equity (e.g. Charness & Rabin, 2002) and inconsistent with such a form of broadly-bracketed equity. The individual-level measures of choice bracketing obtained from our laboratory experiments provide an empirical basis for judging the descriptive realism of assumptions made about bracketing in practice, and our experimental approach provides the convenient tool that will enable future experiments to obtain useful individual-level measures of choice bracketing.

We found that 40-43% of our subjects pass our tests of narrow bracketing and 0-15% of our subjects pass our tests of broad bracketing, a fraction that varies across experiments. After adjusting for predictive power, 67-78% of our subjects are classified as narrow bracketers, only 3-27% are classified as broad bracketers, and 4-6% are classified as partial-narrow bracketers. At least two-and-a-third times (in Shopping) and up to 26 times (in Risk) as many subjects are classified as narrow bracketers than as broad bracketers. Taken together, our results suggest that a majority of people tend to narrowly bracket, while a noticeable minority broadly bracket.

We find a notable minority of subjects who are consistent with neither broad nor narrow bracketing. While some of these subjects pass our tests of partial-narrow bracketing, few are classified as partial-narrow bracketers after adjusting for predictive success. Our results thus suggest that this model of partial-narrow bracketing is not an empirically useful generalization of its polar cases. By ruling out partial-narrow bracketing as a common mode of decision-making, our findings suggest that failures of broad bracketing do not come from incorrectly combining their choices in different parts of the problem. Potential explanations include inability or unwillingness to analyze the entirety of their choice, unawareness that fully appreciating the bundles available leads to higher payoffs, or a rational choice to employ narrow bracketing to simplify their choice. Our study has focused on measurement and does not attempt to disentangle these explanations. However, we note that our experiment is the first to test partial-narrow bracketing and to measure α. Further, there may be the opportunity for new descriptive models to better capture the behavior of the subjects not well-described by polar cases. We hope that future work will build on our design to either corroborate or refine this point.

We found surprising differences in broad bracketing rates across domains. Specifically, more subjects broadly bracketed in the induced payoff Shopping Experiment than in Social, and more in Social than in Risk. This was not hypothesized ex-ante, but taking it at face value has a number of implications. It suggests, for instance, that the narrow or broad classification is not absolute, i.e. some people bracket differently in different situations. Evidence of some learning to broadly bracket from the Shopping Experiment further supports this interpretation.

In Risk and Social, decisions involve a subjective element that may differ depending on one's preferences, unlike in Shopping. This subjective, as opposed to objective, task distinction may mediate how people bracket. Perhaps broad bracketing requires a form of reasoning that people are more likely to engage in when facing a "problem-solving task" than a "choose what you like" task. We think this is an interesting direction for future work.

Our framework for revealing choice bracketing takes the breakdown of decisions into parts as observed or pre-specified by the researcher and our experiment implemented a particular such breakdown. Moreover, we purposefully limit the complementarities between parts by, for instance, ruling out transfering wealth between budgets. This simplifies analysis but the distinction may not be obvious in more complicated real-world settings where choice data is insufficiently rich. We hope future experiments will build on our design to understand how people bracket in richer decision environments and thereby provide more direct guidance for both modeling and improving choice bracketing in applications.

6. Related Literature

Most existing direct evidence on how people bracket their choices is derived from Tversky & Kahneman (1981), in which subjects face the decision problem described in Table 6. This design can rule out broad bracketing and does so for the 73% of their subjects who choose

A and D (which generates a first-order stochastically-dominated distribution over outcomes compared to the pair of choices B and C). It cannot, however, falsify narrow bracketing since any pattern of choices except A and D is consistent with expected utility under either narrow or broad bracketing.[28] Hence without further restrictions, their results are uninformative about the choice bracketing of subjects who do not choose A and D. Rabin & Weizsäcker (2009) revisit this design with incentivized choices, and only 28% of their subjects make the A and D choice. This raises the question of how to classify the remaining 72% of subjects: the design lacks the power to distinguish how they bracket their choices. In contrast, in our design, both narrow and broad bracketing are testable from a single decision without parametric assumptions on utility.

TABLE 6. Tversky and Kahneman's Decision problem

Imagine that you face the following pair of concurrent decisions. First, examine both decisions, then indicate the options you prefer.

Decision (i). Choose between:

A. a sure gain of $240 [85%]	B. 25% chance to gain $1000, and 75% chance to gain nothing [16%]

Decision (ii). Choose between:

C. a sure loss of $750 [13%]	D. 75% chance to lose $1000, and 25% chance to lose nothing [87%]

Our classification exercise is individual-level and non-parametric, and it allows heterogeneity in preferences and bracketing at the subject level. Rabin & Weizsäcker (2009) estimate a parametric structural model from unincentivized survey data based on versions of Table 6 to estimate the fraction of narrow and broad bracketers. Under the assumption that subjects have homogeneous risk preferences but may differ in their bracketing, they find that 11% of their population brackets broadly. Despite the considerable difference in experimental designs and identification strategies, we obtain a similar fraction of broad bracketers. But whereas they classify the remaining 89% as narrow bracketers, our design allows a richer space of possible behavior for each subject, and our analysis suggests that 14% of the population is not well-described by either broad or narrow bracketing.

[28]The combination of strictly concave utility-for-money and narrow bracketing rules out AD and BD, while strictly convex utility-for-money and narrow bracketing requires BD.

Since Tversky and Kahneman's work, a number of related experiments and attendant behavioral concepts have been proposed. Read *et al.* (1999) introduce the umbrella term "choice bracketing" to refer to how an individual groups "individual choices together into sets" and Thaler (1999) defines "mental accounting" as "the set of cognitive operations used by individuals and households to organize, evaluate, and keep track of financial activities". In particular, the concepts of endowment bracketing (Kahneman & Tversky, 1979), mental budgeting (Thaler, 1985), and temporal bracketing (Benartzi & Thaler, 1995; Gneezy & Potters, 1997) have been particularly relevant within behavioral and experimental economics.[29]

Kahneman & Tversky (1979) find that, when subjects are endowed with experimental income and then given a choice between lotteries, they make almost identical choices regardless of the experimentally-endowed income. However, they make different choices when the amount of experimental endowment is instead explicitly incorporated into the prizes of available lotteries to create an economically-identical choice problem. We refer to this as "endowment bracketing". Recent work by Exley & Kessler (2018) arrives at a similar finding in a social allocation task: respondents tend to allocate equal numbers of money-valued tokens from a given account between two anonymous others, regardless of their initial endowments.[30] Imas (2016) provides evidence consistent with broader bracketing of paper losses from past gambles together with a current decision, but more narrowly bracketing when past losses have been realized. Endowment bracketing emerges as a special case of our framework where all but one part of an underlying decision are degenerate.

In the context of repeated decision-making under risk, the question arises as to how risks that resolve at different times are temporally bracketed, together or apart. Gneezy & Potters (1997) show that subjects invest less in a risky asset when decisions are made and returns shown period-by-period than when they must make their investment choices for the next three periods in advance. They explain this behavior in terms of "myopic loss aversion" (Benartzi & Thaler, 1995), that is, loss aversion applied myopically at the horizon of a given

[29]Koch & Nafziger (2019) study the empirical relationship between experimental measures of mental budgeting, endowment bracketing, and choice bracketing in the Tversky-Kahneman 1981 task.

[30]Relatedly, (Thaler, 1985) provides some evidence that most people evaluate outcomes from different sources as if the sources constitute separate "mental accounts" evaluated narrowly and then added up. Evers & Imas (2019) use this mental accounting approach to derive and test for preferences over the timing of outcomes.

decision. Their experimental designs can be modeled in our framework, and if underlying preferences are loss averse, then our definition of narrow bracketing implies myopic loss aversion. In a rather different design, recent evidence from Heimer *et al.* (2020) suggests that some subjects behave as if they do not narrowly bracket future risk-taking opportunities when making a current risk-taking decision; however, their between-subjects design is unable to measure the prevalence of different types of bracketing.

Thaler (1985; 1999) introduces the idea of mental budgeting to capture a consumer who brackets broadly within categories of goods, but narrowly across categories. This provides a particular approach to specifying choice brackets. The mental budgeting approach has been useful in applications to finance (e.g. as in the disposition effect of Shefrin & Statman 1985 and Odean 1998), the study of flypaper effects (Hines & Thaler, 1995; Abeler & Marklein, 2017), and some consumption choices (Hastings and Shapiro 2013; 2018). Recent work derives mental accounts endogenously: Kőszegi & Matějka (2020) derive mental accounts as a response to attention costs in a rational inattention framework, and Lian (2019) does so in a framework in which a decision-maker has different imperfect information in each part.[31,32]

Our approach to bracketing allows different parts to involve the same goods, and our experimental tests rely on this feature. In contrast to mental budgeting, a part (e.g. one store) may have goods from many categories, and a category may be involved in many parts. Nevertheless if a specification of categories is given, then our theoretical approach could be modified to test the hypotheses of (i) broad bracketing within each category, and (ii) narrow bracketing across categories. We see this as a fruitful direction for future work; Blow & Crawford's (2018) theoretical revealed preference characterization of unobserved mental accounts is a step in this direction.

[31]However, Lian models each decision variable as a choice in $\mathbb{R}$ whereas we study parts in $\mathbb{R}^n$. His model can be applied to our setting by mapping each of Lian's decision variables to a relevant "part", and assuming that information is the same for all decision variables within each part. One substantive consequence is that we assume that good i in part k ought to be a perfect substitute for good i in part k', which cannot be accommodated in Lian's model under his assumption of strictly concave utility. In Lian's model, a narrow thinker will best respond to her conditional expectation about her behavior in other parts – which will only generate what we term "narrow bracketing" if a decision-maker expects that, on average, other parts in a decision involve buying zero of each good.

[32]Related work by Galperti (2019) endogenously derives mental budgets as an optimal response to allow a person to counteract present bias while allowing flexibility to respond to taste shocks that affect intratemporal trade-offs.

References

Abeler, Johannes, & Marklein, Felix. 2017. Fungibility, labels, and consumption. *Journal of the European Economic Association*, **15**(1), 99–127.

Andreoni, James, & Miller, John. 2002. Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica*, **70**(2), 737–753.

Barberis, Nicholas, & Huang, Ming. 2009. Preferences with frames: a new utility specification that allows for the framing of risks. *Journal of Economic Dynamics and Control*, **33**(8), 1555–1576.

Barberis, Nicholas, Huang, Ming, & Thaler, Richard H. 2006. Individual preferences, monetary gambles, and stock market participation: A case for narrow framing. *American Economic Review*, **96**(4), 1069–1090.

Beatty, Timothy KM, & Crawford, Ian A. 2011. How demanding is the revealed preference approach to demand? *American Economic Review*, **101**(6), 2782–95.

Benartzi, Shlomo, & Thaler, Richard H. 1995. Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics*, **110**(1), 73–92.

Blow, Laura, & Crawford, Ian. 2018. *Observable Consequences of Mental Accounting*. Working Paper.

Bronars, Stephen G. 1987. The power of nonparametric tests of preference maximization. *Econometrica*, **55**(3), 693–698.

Camerer, Colin F. 2004. Prospect Theory in the Wild: Evidence from the Field. *Pages 148–162 of:* Camerer, Colin F, Loewenstein, George, & Rabin, Matthew (eds), *Advances in Behavioral Economics*. Princeton University Press.

Charness, Gary, & Rabin, Matthew. 2002. Understanding social preferences with simple tests. *Quarterly Journal of Economics*, **117**(3), 817–869.

Choi, Syngjoo, Fisman, Raymond, Gale, Douglas, & Kariv, Shachar. 2007. Consistency and heterogeneity of individual behavior under uncertainty. *American Economic Review*, **97**(5), 1921–1938.

Clark, Stephen A. 1993. Revealed preference and linear utility. *Theory and Decision*, **34**(1), 21–45.

Crawford, Ian, & De Rock, Bram. 2014. Empirical revealed preference. *Annual Review of Economics*, **6**(1), 503–524.

Ellis, Andrew, & Piccione, Michele. 2017. Correlation Misperception in Choice. *American Economic Review*, **107**(4), 1264–92.

Evers, Ellen, & Imas, Alex. 2019. *Mental accounting, similarity, and preferences over the timing of outcomes*. Tech. rept.

Exley, Christine L, & Kessler, Judd B. 2018. *Equity concerns are narrowly framed*. Tech. rept. National Bureau of Economic Research.

Eyster, Erik, & Weizsäcker, Georg. 2016. Correlation neglect in portfolio choice: Lab evidence. *SSRN Working Paper 2914526*.

Fisman, Raymond, Kariv, Shachar, & Markovits, Daniel. 2007. Individual preferences for giving. *American Economic Review*, **97**(5), 1858–1876.

Forges, Francoise, & Minelli, Enrico. 2009. Afriat's theorem for general budget sets. *Journal of Economic Theory*, **144**(1), 135–145.

Galperti, Simone. 2019. A theory of personal budgeting. *Theoretical Economics*, **14**(1), 173–210.

Gneezy, Uri, & Potters, Jan. 1997. An experiment on risk taking and evaluation periods. *Quarterly Journal of Economics*, **112**(2), 631–645.

Hastings, Justine, & Shapiro, Jesse M. 2018. How are SNAP benefits spent? Evidence from a retail panel. *American Economic Review*, **108**(12), 3493–3540.

Hastings, Justine S, & Shapiro, Jesse M. 2013. Fungibility and consumer choice: Evidence from commodity price shocks. *Quarterly Journal of Economics*, **128**(4), 1449–1498.

Heimer, Rawley, Iliewa, Zwetelina, Imas, Alex, & Weber, Martin. 2020. *Dynamic Inconsistency in Risky Choice: Evidence from the Lab and Field*. Tech. rept. SSRN Working Paper 3600583.

Hines, James R, & Thaler, Richard H. 1995. The flypaper effect. *Journal of Economic Perspectives*, **9**(4), 217–226.

Imas, Alex. 2016. The realization effect: Risk-taking after realized versus paper losses. *American Economic Review*, **106**(8), 2086–2109.

Kahneman, Daniel, & Tversky, Amos. 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, **47**(2), 263–291.

Kassambara, Alboukadel. 2020. *ggpubr: 'ggplot2' Based Publication Ready Plots*. R package version 0.4.0.

Koch, Alexander K, & Nafziger, Julia. 2019. Correlates of narrow bracketing. *Scandinavian Journal of Economics*, **121**(4), 1441–1472.

Kőszegi, Botond, & Matějka, Filip. 2020. Choice simplification: A theory of mental budgeting and naive diversification. *Quarterly Journal of Economics*, **135**(2), 1153–1207.

Kőszegi, Botond, & Rabin, Matthew. 2006. A model of reference-dependent preferences. *Quarterly Journal of Economics*, **121**(4), 1133–1165.

Kőszegi, Botond, & Rabin, Matthew. 2007. Reference-dependent risk attitudes. *American Economic Review*, **97**(4), 1047–1073.

Lian, Chen. 2019. *A theory of narrow thinking*. Tech. rept. MIT mimeo. Working Paper.

Odean, Terrance. 1998. Are investors reluctant to realize their losses? *Journal of Finance*, **53**(5), 1775–1798.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rabin, Matthew, & Weizsäcker, Georg. 2009. Narrow bracketing and dominated choices. *American Economic Review*, **99**(4), 1508–43.

Read, Daniel, Loewenstein, George, & Rabin, Matthew. 1999. Choice Bracketing. *Journal of Risk & Uncertainty*, **19**(1-3), 171–197.

Selten, Reinhard. 1991. Properties of a measure of predictive success. *Mathematical Social Sciences*, **21**(2), 153–167.

Shefrin, Hersh, & Statman, Meir. 1985. The disposition to sell winners too early and ride losers too long: Theory and evidence. *Journal of Finance*, **40**(3), 777–790.

Sobel, Joel. 2005. Interdependent preferences and reciprocity. *Journal of Economic Literature*, **43**(2), 392–436.

Thaler, Richard. 1985. Mental accounting and consumer choice. *Marketing Science*, **4**(3), 199–214.

Thaler, Richard H. 1999. Mental accounting matters. *Journal of Behavioral Decision Making*, **12**(3), 183–206.

Tversky, Amos, & Kahneman, Daniel. 1981. The framing of decisions and the psychology of choice. *Science*, **211**(4481), 453–458.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wickham, Hadley, & Henry, Lionel. 2020. *tidyr: Tidy Messy Data*. R package version 1.1.0.

Appendix A. Theoretical Appendix: Proofs, Derivations, and PNB Algorithm

Proof of Theorem 1. Parts (i) and (ii) follow the usual proof that SARP holds if and only if a data set is rationalizable; see e.g. Forges & Minelli (2009).

To see (iii), fix α and $\mathcal{D}$. Our approach will be to map the choices in $\mathcal{D}$ to an ancillary data set $\mathcal{D}^\alpha$ of choices from lotteries over a finite subset $Y \subseteq \mathbb{R}^n$. Then, it will be shown that $\mathcal{D}^\alpha$ is consistent with expected utility if and only if $\mathcal{D}$ is rationalizable by α-partial-narrow bracketing.

We first introduce some notation. For each pair (t,k), let $x^{t,-k} = \sum_{j \neq k} x^{t,j}$, $p^{t,k} = (\alpha, x^t; (1-\alpha), x^{t,k})$, a lottery over $\mathbb{R}^n$. From these lotteries, let $Y \subseteq \mathbb{R}^n$ be a finite set that includes the union of the supports of the lotteries $p^{t,k}$, as well as the vector 0; it will be convenient to take Y equal to this set but that is inessential. Denote by ΔY the finite support lotteries over Y. The budget set

$$Q^{t,k} = \Big\{(\alpha, y; (1-\alpha), y') \in \Delta Y : \text{there exists } x \in B^{t,k} \text{ so that } y' \leq x + x^{t,-k}, y \leq x\Big\}$$

is a finite set of lotteries over Y that includes $p^{t,k}$and any others that are affordable. The ancillary data set is

$$\mathcal{D}^\alpha = \Big\{\Big(p^{t,k}, Q^{t,k}\Big)\Big\}_{(t,k)}$$

where any other interdependencies between the lotteries are ignored.

Observe that if the data is rationalized by α-PNB, then there exists an EU preference V over ΔY so that

$$V\left(p^{t,k}\right) > V\left(q^{t,k}\right) \; \forall q^{t,k} \in Q^{t,k} \backslash p^{t,k}.$$

Let u be the utility index of V and extend to $\mathbb{R}^n_+$ by

$$u^*(z) = \max_{y \leq z, y \in Y} u(y).$$

Then for any $x \in B^{t,k} \backslash x^{t,k}$,

$$\alpha u(x^{t,k}) + (1-\alpha) u(x^{t,k} + x^{t,-k}) > \alpha u^*(x) + (1-\alpha) u^*(x + x^{t,-k}).$$

Conditions under which $\mathcal{D}^\alpha$ is rationalized by expected utility are well-known. Here, we follow Clark (1993) to get a utility index u over Y.

As standard, p is directly revealed preferred to q, written $p \succ q$, if $p = p^{t,k}$ and $q \in Q^{t,k}$ for some index (t,k). Define $p \succsim q$ if $p \succ q$ or $p = q$. We say that p is indirectly reveal preferred via independence to q, written $p \tilde{\succsim} q$, whenever there exist $r, p_1, \ldots, p_n, q_1, \ldots, q_n \in \Delta Y$, $\alpha \in (0,1]$, and $\lambda_1, \ldots, \lambda_n > 0$ so that

$$\begin{aligned} \alpha p + (1-\alpha) r &= \sum_{i=1}^n \lambda_i p_i \\ \alpha q + (1-\alpha) r &= \sum_{i=1}^n \lambda_i q_i \\ p_i &\succsim q_i \text{ for all } i = 1, \ldots, n \\ \sum_{i=1}^n \lambda_i &= 1 \end{aligned}$$

and $p \tilde{\succ} q$ holds whenever at $p_i \succ q_i$ for some i. The Linear Axiom of Revealed Preference (LARP) is that $q \tilde{\succsim} p$ implies that $p \tilde{\succ} q$ does not hold. Theorem 3 of Clark (1993), combined with the finiteness of $\mathcal{D}^\alpha$, implies that $\tilde{\succsim}$ has an expected utility rationalization. Conclude $\mathcal{D}^\alpha$ satisfies LARP if and only if $\mathcal{D}$ is rationalized by α-partial-narrow bracketing.

The following algorithm inputs $\mathcal{D}^\alpha$ and outputs 1 if and only $\mathcal{D}$ is rationalized by α-partial-narrow bracketing for any rational number $\alpha \in [0,1]$. For concreteness, set $Y = \{y_1, y_2, \ldots, y_m\}$. We can identify $p \in \Delta Y$ with $p \in \mathbb{R}^m$ so that $p_i = p(y_i)$.

Define the matrix A with rows given by $p^{t,k} - q^{t,k}$ for each $q^{t,k} \in Q^{t,k} \backslash p^{t,k}$. A is an $m \times D$ matrix interpreted as strict preference, where D is the number of data points extracted. Similarly, let B be an $m \times D'$ matrix with rows given by $p - q$ where p and q represent comparisons including weak preference, such as implications of symmetry.

Let $\vec{1}$ be a $1 \times D$ matrix of ones, $\vec{0}$ a $1 \times D'$ matrix of zeroes and

$$\hat{A} = \begin{bmatrix} A & B \\ \vec{1} & \vec{0} \end{bmatrix}.$$

Let $\hat{b}$ be an $m+1$ vector with the first m components 0 and the last 1. Solve the system

$$\min_{y \in \mathbb{R}^{m+1}} y^T \hat{b}$$
$$s.t. \hat{A}^T y \geq 0$$

for y. The Ellipsoid algorithm provides a polynomial-time solution to this. Let y^* be the solution, if one exists. There are two cases:

There is no solution to $\hat{A}^T y \geq 0$ or y^* satisfies $y^{*T} \hat{b} \geq 0$. Then, Farkas's Lemma implies there exists $\phi \geq 0$ so that $\hat{A}\phi = \hat{b}$. Defining $\lambda_i = \frac{\phi_i}{\sum \phi_i}$, we see that LARP fails since $\lambda_i > 0$ for some $p_i \succ q_i$ and $\sum_{i=1}^{m+1} \lambda_i p_i = \sum_{i=1}^{m+1} \lambda_i q_i$. So return 0. Observe $[A, B]\, \phi = 0$ and $\hat{b}^T \phi = 1$.

The program is unbounded below or $y^{*T} \hat{b} < 0$. Then, Farkas's Lemma implies there does not exist any $\phi \geq 0$ so that $\hat{A}\phi = \hat{b}$. If there existed $\lambda_1, \ldots, \lambda_{m+1}$ that led to a failure of LARP, taking $\phi_i = \lambda_i$ would give $\phi \geq 0$ so that $\hat{A}\phi = \hat{b}$. Conclude LARP is satisfied and return 1.

To see why this algorithm works, note that if $p \tilde{\succsim} q$, then there exists $\lambda > 0$ and $a_1, \ldots, a_n \in A$ so that

$$\lambda(p - q) = \sum \phi^i a_i$$

for $\phi^i > 0$. This is equivalent to $p - q \in cone(co(A))$. LARP holds if $p - q \in cone(co(A))$, then $q - p \notin cone(co(A))$, or equivalently $0 \notin cone(co(A))$. That is, there does not exist $\phi^1, \ldots, \phi^D \geq 0$ so that $\sum \phi^i = 1$ and $\lambda > 0$ so that

$$\lambda \sum_{i=1}^{D} \phi^i A_i = 0;$$

here A_i is the ith row of A. The λ is redundant, so this can be rewritten as

$$A\phi = 0$$

$$\vec{1}\phi = 1$$

for some vector $\phi \in \mathbb{R}^D$and $\vec{1}$ is $1 \times D$ matrix of ones.

To specialize the above algorithm to preferred personal equilibrium partial-narrow bracketing, we perform the analysis decision by decision. Letting $n(t)$ be the number of parts in decision t, we consider the lottery

$$p^t = \alpha \left(\frac{1}{n}, x^{t,k}\right)_{k=1}^{n(t)} + (1-\alpha)\left(1, x^t\right)$$

to be chosen from the budget set

$$Q^t = \left\{ \alpha \left(\frac{1}{n}, y^k\right)_{k=1}^{n(t)} + (1-\alpha)(1, y) \in \Delta Y : \text{there exists } z^k \in B^{t,k} \text{ so that } y^k \leq z^k \, and \, y \leq \sum_{k=1}^{n(t)} z^k \right\}$$

is a finite set of lotteries over Y that includes $p^{t,k}$and any others that are affordable. The ancillary data set

$$\mathcal{D}^{PPE,\alpha} = \left\{\left(p^t, Q^t\right)\right\}_{(t)}$$

replaces $\mathcal{D}^\alpha$, and the rest of the proof goes through. □

Theorem 2 (Identification of α.). *Suppose that u is a known differentiable and strictly quasi-concave function and $\mathcal{D}$ is rationalized by partial-narrow bracketing. Further suppose that for some t, k, $B^{t,k}$is a Walrasian budget set for prices $p^{t,k}$, and $n = 2$, $\frac{p_1^{t,1}}{p_2^{t,1}} \neq \frac{p_1^{t,2}}{p_2^{t,2}}$, $\lim\limits_{x_1 \to 0^+} \frac{\partial u(x_1,x_2)}{\partial x_1} = +\infty$ for all $x_2 > 0$, and $\lim\limits_{x_2 \to 0^+} \frac{\partial u(x_1,x_2)}{\partial x_2} = +\infty$ for all $x_1 > 0$.*
Then, there exists a unique α that rationalizes choices.

Proof of Theorem 2. Suppose the assumptions of the Theorem are satisfied. In the partial-narrow bracketing model, the marginal utility per dollar of spending on good j in decision k is given by:

$$\frac{1}{p_j^{t,k}} \left((1-\alpha) \frac{\partial u(x^{t,1} + x^{t,2})}{\partial x_j} + \alpha \frac{\partial u(x^{t,k})}{\partial x_j} \right)$$

for $k = 1, 2$. By the limit conditions, if $\alpha > 0$, then $x_j^{t,k} > 0$ for $j = 1, 2$ and $k = 1, 2$; thus $x_1^{t,k} = 0$ or $x_2^{t,k} = 0$ implies $\alpha = 0$. If $\frac{1}{p_2^{t,k}} \partial u(x^{t,k})/\partial x_2 = \frac{1}{p_1^{t,k}} \partial u(x^{t,k})/\partial x_1$ for both k, then set $\alpha = 1$. Now suppose that $\frac{1}{p_2^{t,k}} \partial u(x^{t,k})/\partial x_2 \neq \frac{1}{p_1^{t,k}} \partial u(x^{t,k})/\partial x_1$ and an interior allocation is chosen in budget set k. The first-order condition for an interior maximizer equates the marginal utility per dollar across the two goods; rearranging this condition, we obtain

$$\alpha = \frac{\frac{1}{p_1^{t,k}} \partial u(x^{t,1} + x^{t,2})/\partial x_1 - \frac{1}{p_2^{t,k}} \partial u(x^{t,1} + x^{t,2})/\partial x_2}{\frac{1}{p_1^{t,k}} \partial u(x^{t,1} + x^{t,2})/\partial x_1 - \frac{1}{p_2^{t,k}} \partial u(x^{t,1} + x^{t,2})/\partial x_2 + \frac{1}{p_2^{t,k}} \partial u(x^{t,k})/\partial x_2 - \frac{1}{p_1^{t,k}} \partial u(x^{t,k})/\partial x_1}$$

for $k = 1, 2$. It remains to verify that the expression on the right-hand side always lies in the interval $(0, 1)$. Since $\frac{1}{p_2^{t,k}} \partial u(x^{t,k})/\partial x_2 \neq \frac{1}{p_1^{t,k}} \partial u(x^{t,k})/\partial x_1$, the solution satisfies $(1 - \alpha)\left(\frac{1}{p_1^{t,k}} \frac{\partial u(x^{t,1}+x^{t,2})}{\partial x_1} - \frac{1}{p_2^{t,k}} \frac{\partial u(x^{t,1}+x^{t,2})}{\partial x_2}\right) = \alpha \left(\frac{1}{p_2^{t,k}} \frac{\partial u(x^{t,k})}{\partial x_2} - \frac{1}{p_1^{t,k}} \frac{\partial u(x^{t,k})}{\partial x_1}\right)$. Thus, if $\alpha \in (0, 1)$, the denominator and numerator have the same sign, and the denominator is strictly larger in absolute value – therefore the expression is well-defined and thus α is identified from choices. □

.

Joint implications of narrow and broad bracketing. To study the joint implications of broad and narrow bracketing we consider endowment bracketing in a rich domain. We thus suppose that we observe an infinite data set on all decisions that each consist of a binary part and a second degenerate part summarized by its only available bundle z. Suppose that when a person has bundle $z \in \mathbb{R}_+^n$ in their second part and must make a binary choice, she applies a complete, transitive, continuous, and strictly increasing binary relation $\succsim_z$ (call such a binary relation a *well-behaved preference*). In this domain, we say that a person *broadly brackets* if $x \succsim_z y \iff x + z \succsim_{\underline{0}} y + z$; a person *narrowly brackets* if $x \succsim_z y \iff x \succsim_{z'} y$ for all $z' \in \mathbb{R}_+^n$.

Theorem 3. *The family of well-behaved preferences $\{\succsim_z\}_{z \in \mathbb{R}_+^n}$ satisfies both narrow and broad bracketing if and only if there exists a vector $u \in \mathbb{R}_{++}^n$ such that $x \succsim_z y$ if and only if $x \cdot u \geq y \cdot u$.*

Proof. The "if" direction is immediate. So now consider the "only if" direction.

We do so by constructing a utility representation for $\succsim_{\underline{0}}$. Since $\succsim_{\underline{0}}$ is a well-behaved preference relation, for each $x \in \mathbb{R}^n_+$ there exists a unique number $\kappa \in \mathbb{R}_+$ such that $x \sim_{\underline{0}} \kappa\underline{1}$; define $u(x) = \kappa$ for each x and the corresponding κ. By strict monotonicity, u represents $\succsim_{\underline{0}}$.

Now suppose $x \sim_{\underline{0}} y$. Then, by narrow bracketing, $x \sim_x y$ and $x \sim_y y$. By broad bracketing, $2x \sim_{\underline{0}} y + x$ and $x + y \sim_{\underline{0}} 2y$. By transitivity, $2x \sim_{\underline{0}} 2y$. But since for each $x, y \in \mathbb{R}^n_+$, $.5x, .5y \in \mathbb{R}^n_+$, it follows that $x \sim_{\underline{0}} y$ implies $.5x \sim_{\underline{0}} .5y$ as well. Now suppose that for every natural number of at most a, we have shown that $x \sim_{\underline{0}} y \implies ax \sim_{\underline{0}} (a - b)x + by \sim_{\underline{0}} ay$ for every natural number $b \leq a$. Fix $\beta \in (0, 1]$. Then, by narrow bracketing, $ax \sim_{\beta x} (a - b)x + by$ and $(a - c)x + cy \sim_{\beta y} ay$ for all $b, c \leq a$. Thus by broad bracketing, $(a+\beta)x \sim_{\underline{0}} (a-b+\beta)x+by$ and $(a-c)x+(c+\beta)y \sim_{\underline{0}} (a+\beta)y$. Setting $b = c+\beta$, transitivity implies $(a + \beta)x \sim_{\underline{0}} (a + \beta)y$. By induction, it follows that $x \sim_{\underline{0}} y$ implies $\beta x \sim_{\underline{0}} \beta y$ for all $\kappa \in \mathbb{R}_+$. Thus, $u(x) = u(y) = u(\kappa\underline{1}) = \kappa \implies u(\lambda x) = u(\lambda y) = u(\lambda\kappa\underline{1}) = \lambda\kappa$ for all $\lambda > 1$. Since the utility of arbitrarily small bundles is well-defined, the preceding holds for $\lambda \in (0, 1)$ as well. Thus u is homogeneous of degree 1.

Next, take $x, y, z \in \Delta^n$ (the n-dimensional unit simplex) – we will establish additive linearity on that subspace, and then by the homogeneity, the result follows. Suppose $x \sim_{\underline{0}} y$. By the homogeneity of u, $\lambda x \sim_{\underline{0}} \lambda y$; by narrow bracketing, $\lambda x \sim_{(1-\lambda)z} \lambda y$, thus by broad bracketing, $(1-\lambda)x+\lambda z \sim_{\underline{0}} (1-\lambda)y+\lambda z$. Thus $\succsim_0$ satisfies the Independence Axiom on Δ^n, so there exists a vector $u \in \mathbb{R}^n \backslash \{\underline{0}\}$ such that preferences can be represented by $u(x) = x \cdot u$. By homogeneity of degree 1, , we extend this representation to all of $\mathbb{R}^n_+$. □

Derivations of predictions and their experimental implementations. First, observe that if relation P is acyclic, it must be anti-symmetric.

Derivation of Prediction 1 (NB-WARP). Suppose $x^{t',k'} \in B^{t,k} \subseteq B^{t',k'}$. By the definition of P in NB-SARP, $x^{t',k'}Px$ for all $x \in B^{t',k'} \backslash \{x^{t',k'}\}$. Since $B^{t,k} \subseteq B^{t',k'}$, it follows that $x^{t',k'}Px$ for all $x \in B^{t,k} \backslash \{x^{t',k'}\}$. But if $x^{t,k} \neq x^{t',k'}$, the definition of P would imply $x^{t,k}Px^{t',k'}$, which would violate acyclicity. Thus $x^{t,k} = x^{t',k'}$.

Discretized implementation in our experiment. In all of our tests of NB-WARP, we have $B^{t,k} = B^{t',k'}$ exactly. We thus test NB-WARP exactly, without needing to adjust for discreteness.

Derivation of Prediction 2 (BB-WARP). Suppose $x^{t'} \in B^t \subseteq B^{t'}$. By the definition of P in BB-SARP, $x^{t',k'} P x$ for all $x \in B^{t'} \backslash \{x^{t'}\}$. Since $B^t \subseteq B^{t'}$, it follows that $x^{t'} P x$ for all $x \in B^t \backslash \{x^{t'}\}$. But if $x^t \neq x^{t'}$, the definition of P would imply $x^t P x^{t'}$, which would violate acyclicity. Thus $x^t = x^{t'}$.

Discretized implementation in our experiment. In the Risk and Social experiments, $\text{co}B^{1,1} + \text{co}B^{1,2} \subseteq \text{co}B^{2,1}$. In particular, $B^{2,1} = \{(\$0, \$28)\,; (\$2, \$26)\,; \ldots ; (\$28, \$0)\}\}$, whereas $B^{1,1} + B^{1,2} = \{(\$0, \$28)\,; (\$1, \$27)\,; \ldots ; (\$16, \$12)\,; (\$17, \$10.80)\,; \ldots ; (\$26, \$0)\}$. Consider two cases.

If $x_A^{2,1} \leq \$16$, the bundle chosen in $B^{2,1}$ is exactly affordable in $B^{1,1} + B^{1,2}$. The caveat is that the set $B^{1,1} + B^{1,2}$ has a higher resolution in this range. We do not adjust for this, implicitly interpreting their choice from $B^{2,1}$ as revealing their preferences over the $\text{co}B^{2,1}$. However, if this leads a subject to fail BB-WARP, if their preferences are well-behaved, they would choose in $B^{2,1}$ one of the closest bundles to their preferred bundle in $\text{co}B^{2,1}$ and be within one error away from passing BB-WARP.

However, if $x_A^2 > \$16$, then the person reveals a sufficiently strong preference for person/state A over B that their desired bundle in B^2 is not affordable in $\text{co}B^{1,1} + \text{co}B^{1,2}$ and thus they trivially pass BB-WARP.

Derivation of Prediction 3 (BB-Mon). Suppose $x_1^{t,2} > 0$, $x_2^{t,1} > 0$, and $p_1^{t,1} \leq p_2^{t,1}$, $p_1^{t,2} > p_2^{t,2}$. Let $\epsilon = \min\{x_1^{t,2}, x_2^{t,1}\}$. Consider the alternative pair of choices $(y^{t,1}, y^{t,2})$ given by $y^{t,1} = \left(x_1^{t,1} + \frac{p_2^{t,1}}{p_1^{t,1}}\epsilon, x_2^{t,1} - \epsilon\right)$ and $y^{t,2} = \left(x_1^{t,2} - \epsilon, x_2^{t,2} + \frac{p_1^{t,2}}{p_2^{t,2}}\epsilon\right)$. By construction, $(y^{t,1}, y^{t,2})$ is affordable. We have that

$y_1^{t,1} + y_1^{t,2} = x_1^{t,1} + x_1^{t,2} + \frac{p_2^{t,1} - p_1^{t,1}}{p_1^{t,1}}\epsilon \geq x_1^{t,1} + x_1^{t,2}$, and

$y_2^{t,1} + y_2^{t,2} = x_2^{t,1} + x_2^{t,2} + \frac{p_1^{t,2} - p_2^{t,2}}{p_2^{t,1}}\epsilon > x_2^{t,1} + x_2^{t,2}$

where the inequalities respectively follow from $p_1^{t,1} \leq p_2^{t,1}$ and $p_1^{t,2} > p_2^{t,2}$. But $x^t P y^t$ would violate monotonicity; thus such an $x^{t,1}, x^{t,2}$ pair could not pass BB-SARP.

Discretized implementation in our experiment. The argument in the proof applies with minimal modification to our discretized experiment, and thus we apply the conditions directly.

Online Appendix

Appendix B. Supplementary Data Visualizations

Risk and Social Experiments. First, we plot histograms of allocations in each part in Risk and Social Experiments (Figures 4 and 5). Notice that while the two experiments have a similar number of subjects (99 and 102 respectively), the y-axis scale differs across the two histograms. This reflects greater variation in behavior in the Risk than in the Social Experiment in D2 and D4. In both experiments, the vast majority of subjects make an equal allocation between assets/persons A and B, consistent with strict symmetric preferences. In Risk, 69 subjects make equal allocations in D4 (79 are within one token of doing so) while in Social the corresponding number is 90 subjects (97 within one token). While some variation in behavior is the norm in experiments, prior normative, behavioral, and experimental work on risk preferences generally suggests that risk aversion over 50-50 lotteries is not just strongly modal but a nearly universal. This illustrates the value of including D2, D4, and D5 in our Risk Experiment to enable us to apply WARP-style tests which do not require assumptions about preferences.

Figure 4. Allocations in Risk by part

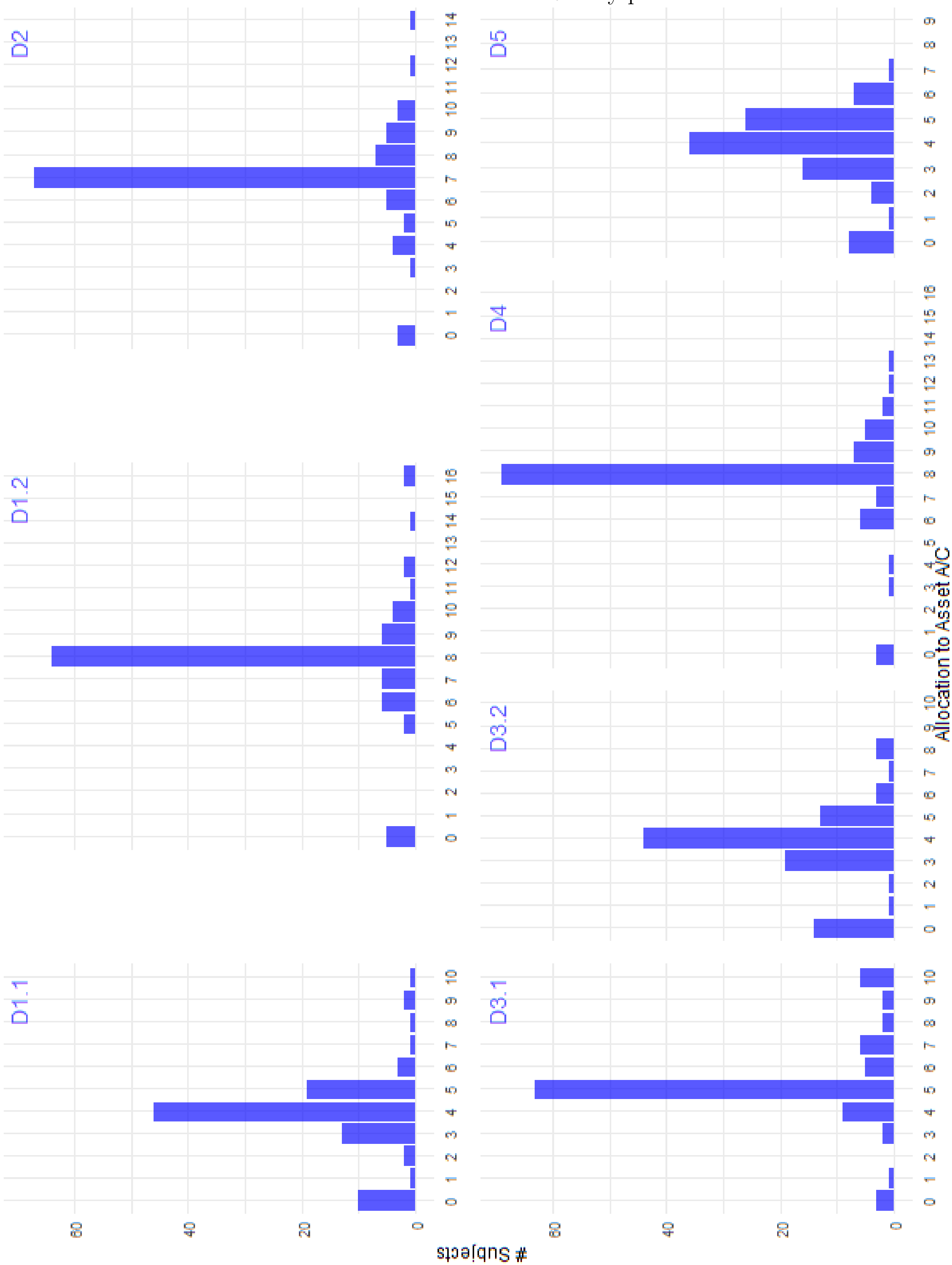


FIGURE 5. Allocations in Social by part

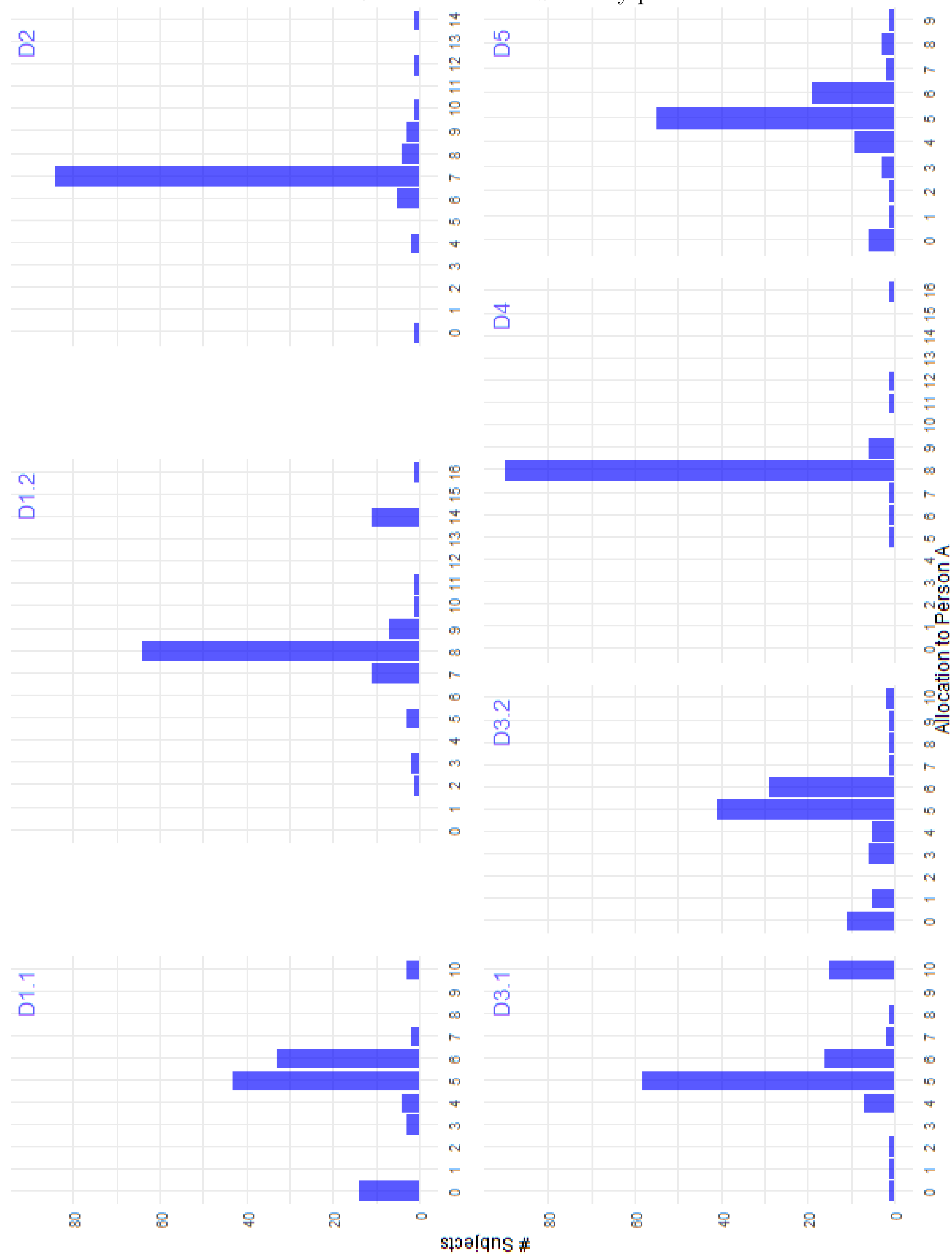


Figure 6 plots the aggregate (i.e. decision-level) allocations to die rolls 1-3/4-6 (Risk) and Person A/B (Social) in Decision 1 (shown in red) and in Decision 2 (shown in blue); the size of each circle is proportional to the number of subjects making an allocation. Our BB-WARP is based on a comparison of these allocations for each individual. While the blue dots (D2) lie on a line with slope -1 extending from (0, 28) to (28,0), almost all of the mass in red (D1) lies below that line – indicating subjects who made D1 choices that resulted in a dominated overall bundle. Since it is impossible in Decision 1 to obtain an aggregate bundle that allocates strictly more than \$16 to die roll 1-3/Person A, the very small number of subjects (10 in Risk, 6 in Social) who make such an allocation in Decision 2 automatically pass BB-WARP.

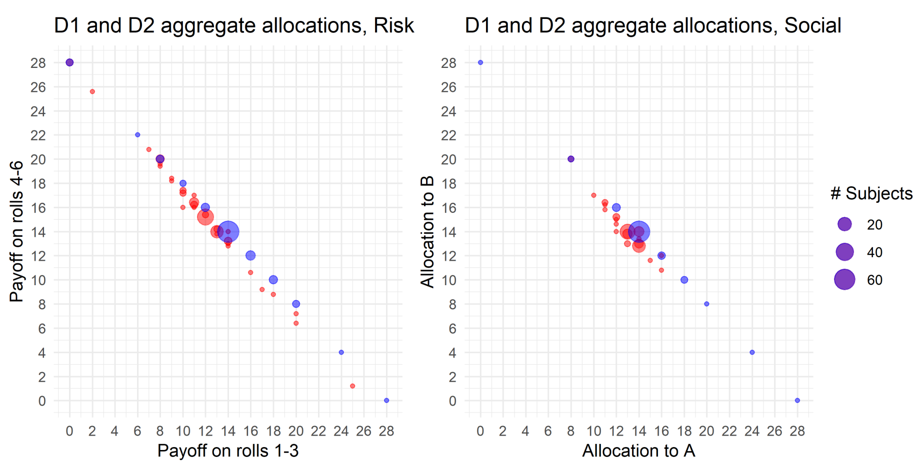


Figure 6. Aggregate allocations in D1 (red) and D2 (blue)

Figure 7 plots the individual-level allocations to Asset A/C in D1.1, D3.2, and D1.2 against their allocation in D5, D5, and D4 respectively. NB-WARP tests that the person makes the same allocation in each pair of parts, and is satisfied by all dots on line through (0,0) with a slope of 1.

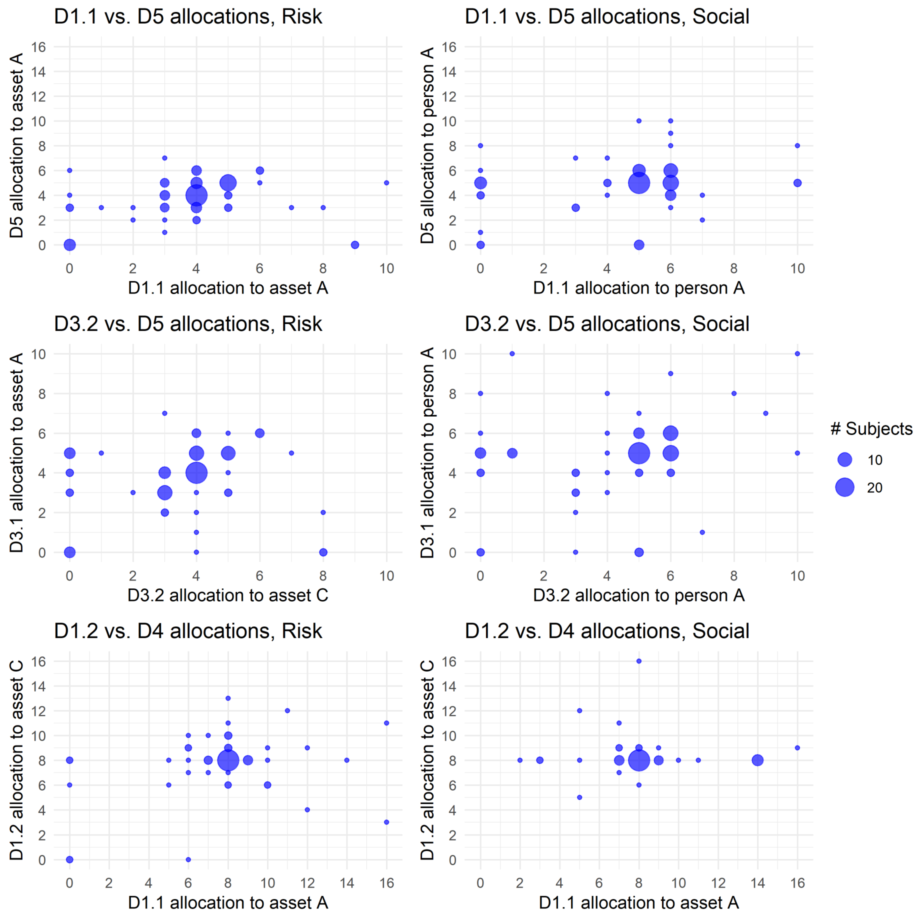
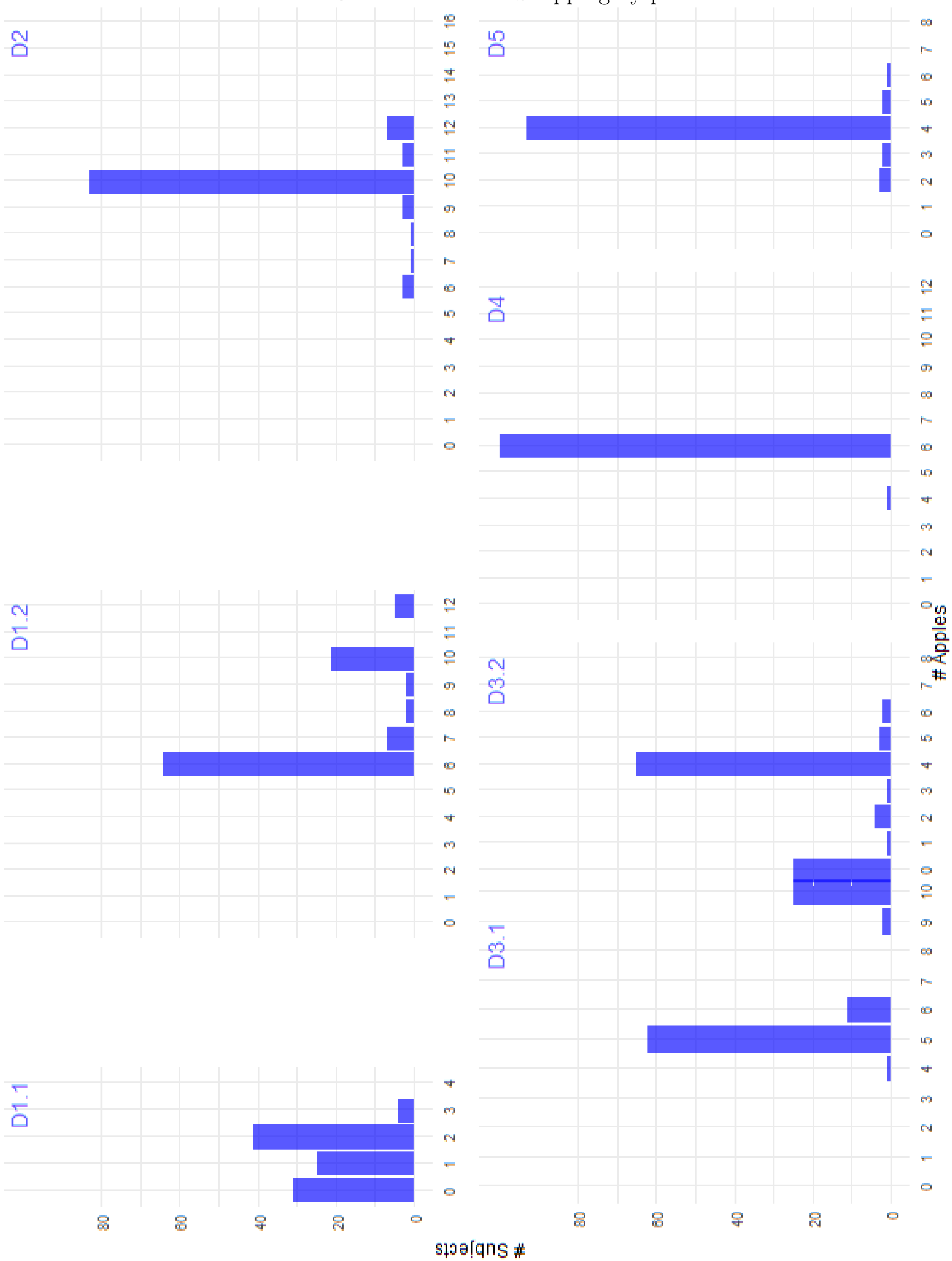


FIGURE 7. Comparison of identical parts

Shopping Experiment. Figure 8 plots histograms of purchases in each part. The histograms for D2, D4, and D5, which each have a single part, show very good adherence to value maximization: 91% of these allocations maximize payoffs exactly, and there are only 4 allocations (all in D2) in which subjects are more than two errors from the optimal allocation. The histograms for D1.1, D1.2, D3.1, and D3.2 illustrate close adherence to the predictions of full narrow bracketing with the exception that many subjects purchase two apples in D1.1 when narrow bracketing predicts they should buy only one,[33] and broad bracketing implies they should buy no apples.

[33]Note that since the price of oranges is half that of apples in D1.1 and lines correspond to numbers of oranges here, this counts as two errors relative to the narrow bracketing prediction. For a narrow bracketer, this is a very "small" mistake in terms of narrowly-assessed payoffs.

Figure 8. Allocations in Shopping by part



Further notes. With the exception of the PNB Algorithm, results in this paper were computed in R (R Core Team, 2020), and made use of the tidyr package (Wickham & Henry, 2020). Plots generated using ggplot2 (Wickham, 2016) and ggpubr (Kassambara, 2020).

Appendix C. Tests Assuming Symmetry

C.1. **Testable conditions.** We present the following testable implications of narrow, broad, and partial-narrow bracketing when underlying preferences are required to be symmetric.

Prediction 4 (NB-Sym)**.** Suppose $\mathcal{D}$ is rationalized by symmetric narrow bracketing and each $B^{t,k}$ is a Walrasian budget set for prices $p^{t,k}$.
If $p_j^{t,k} \geq p_i^{t,k}$, then $x_i^{t,k} \geq x_j^{t,k}$.

With narrow bracketing, symmetry's implications are straightforward. The subject should purchase at least as much of the cheaper good. With broad and partial-narrow bracketing, the implications are more subtle.

Prediction 5 (BB-Sym)**.** Suppose $\mathcal{D}$ is rationalized by symmetric broad bracketing, and each $B^{t,k}$ is a Walrasian budget set for prices $p^{t,k}$ and income $I^{t,k}$. Then,
(i) if $B^t = B^{t,1} + B^{t,2}$,$p_1^{t,1} = p_2^{t,1}$, and $p_i^{t,2} > p_j^{t,2}$, then $x_i^{t,1} + x_i^{t,2} \leq x_j^{t,1} + x_j^{t,2}$; if in addition $\frac{I^{t,2}}{p_j^{t,2}} \leq \frac{I^{t,1}}{p_i^{t,1}}$, then $x_i^{t,2} = 0$, and
(ii) if $B^t = B^{t,1}$ and $p_i^{t,1} \geq p_j^{t,1}$, then $x_j^{t,1} \geq x_i^{t,1}$.

Prediction 6 (PNB-Sym)**.** Suppose $\mathcal{D}$ is rationalized by symmetric partial-narrow bracketing and each $B^{t,k}$ is a Walrasian budget set for prices $p^{t,k}$. Then,
(i) if $B^t = B^{t,1} + B^{t,2}$, $p_1^{t,1} = p_2^{t,1}$ and $p_i^{t,2} > p_j^{t,2}$, then $x_j^{t,2} \geq x_i^{t,2}$ and $x_j^t \geq x_i^t$, and
(ii) if $B^t = B^{t,1}$ and $p_1^{t,1} \geq p_2^{t,1}$, then $x_2^{t,1} \geq x_1^{t,1}$.

For an intuition, recall the logic of comparative advantage. The opportunity cost of consuming good i from $B^{t,1}$ is lower than in $B^{t,2}$, and vice versa for consuming good j. A DM should purchase at least as much of the good in the part where it is cheapest relative to the other good. With broad bracketing, this specialization is extreme: if a subject purchases positive amounts of good i from the second part, then she must exhaust the budget of the first on good i. Otherwise, she forgoes the opportunity to consume more of each. The DM purchases the good only where it is cheapest, until switching to purchasing the other good. With partial-narrow, the specialization is less extreme as the DM weighs the effect of less consumption of the more expensive good in the part as well as in the decision overall.

C.2. **Tests of NB-, BB-, and PNB-Sym.** Symmetric preferences are natural in our experimental setup, and in the decisions with equal prices, such as Decision 2, the majority of subjects allocate evenly. By testing this property's implications in other choices, we obtain more powerful tests that do not rely on cross-decision comparisons. This power comes at the cost of having to jointly test bracketing and that underlying preferences are symmetric. To the extent that these properties are compelling in our environment, the next set of tests distinguish the three models of bracketing.

	Risk			Social		
# errors	0	1	2	0	1	2
NB-Sym (D1)	62	73	84	41	72	82
NB-Sym (D3)	60	74	82	34	76	82
NB-Sym (both)	46	57	65	22	46	67
BB-Sym (D1)	1	1	1	11	11	12
BB-Sym (D3)	6	7	8	14	14	14
BB-Sym (both)	0	0	0	10	10	10
PNB-Sym (D1)	90	92	93	63	95	99
PNB-Sym (D3)	91	94	96	67	96	97
PNB-Sym (both)	86	88	92	46	77	93
# subjects	99			102		

Entries count the # of subjects who pass each test at the listed error allowance.

TABLE 7. Tests of NB-, BB-, and PNB-Sym

We thus test NB-Sym, BB-Sym, and PNB-Sym, using all of each condition's implications for a given decision. In Decision 1 tests, we find that 63% and 40% of subjects pass NB-Sym in the Risk and Social Experiments respectively. These pass rates increase to 74% and 71% if we allow for one error. We find quantitatively similar results for Decision 3. Pass rates go down to 46% and 22% when we require a subject to pass NB-Sym in both Decisions 1 and 3 – but these numbers increase to 58% and 45% when we allow for one error across both decisions. Thus most, but not all, subjects pass this test of the conjunction of narrow bracketing with symmetric preferences.

Turning to tests of broad bracketing, 1% and 11% of subjects pass BB-Sym in Decision 1 of the Risk and Social Experiments respectively. Allowing for one error does not increase pass rates at all, though one additional subject is within two errors of passing. We obtain slightly higher pass rates (6% and 14%) in Decision 3. This discrepancy can be traced to the relative strength of BB-Sym in these decisions. It makes a point prediction in Decision 1 but allows two possible choices in Decision 3. No subjects pass BB-Sym in both Decisions 1 and 3 for Risk, even when allowing for two errors. In Social, 10% of subjects pass, which does not change when allowing for up to two errors. Thus these tests suggest that only a small minority of subjects are close to consistent with broad bracketing.

Our PNB-Sym tests of partial-narrow bracketing have higher pass rates, as expected – 91% and 62% for Risk and Social Experiments respectively in Decision 1. Allowing one error raises pass rates to include the vast majority of subjects – 93% and 93% respectively. We obtain quantitatively similar results for Decision 3. However, allowing one error, only 18% and 12% pass PNB-Sym and neither BB-Sym nor NB-Sym in Decision 1 in the two experiments. We find that 87% and 45% pass PNB-Sym in both decisions, while 89% and 75% do so when allowing for one error. However, this means that, when allowing for one error of tolerance, only 31% and 20% of subjects pass PNB-Sym in both Decision 1 and 3 who pass neither BB-Sym nor NB-Sym. Thus, partial-narrow bracketing only somewhat helps to account for behavior in our experiments, in spite of the model's relatively weak implications in our experiment without parametric assumptions about utility.

Result 6. *When allowing for one error in the Risk and Social Experiments, 58% and 45% of subjects pass NB-Sym and 0% and 10% pass BB-Sym; only 31% and 20% pass PNB-Sym but not NB- nor BB-Sym.*

Appendix D. Power

For each test presented in Tables 2, 3, and 4 we compute the probability that randomly-generated choices would pass each. This approach to analyzing the power of revealed preference tests follows Bronars (1987), and has been used in Andreoni & Miller (2002, p. 744) and Choi *et al.* (2007, p. 1927), among other papers.

	Risk/Social		
# errors	0	1	2
NB-WARP (D1.1 and D5)	0.091	0.256	0.405
NB-WARP (D1.2 and D4))	0.059	0.170	0.273
NB-WARP (D3.2 and D5)	0.091	0.256	0.405
NB-WARP (D1.1 and D3.2)	0.091	0.256	0.405
NB-WARP (all)	0.0005	0.004	0.014
BB-WARP (D1 and D2)	0.427	0.517	0.591
BB-Mon (D1)	0.144	0.278	0.401
BB-Mon (D3)	0.174	0.331	0.471
BB-Mon (both)	0.025	0.071	0.134

Table 8. Probability of Random Choice Passing NB-/BB- WARP tests

		Risk/Social	
# errors	0	1	2
NB-SARP	$1{\times}10^{-7}$	$1{\times}10^{-6}$	$8{\times}10^{-6}$
BB-SARP	$2{\times}10^{-7}$	$2{\times}10^{-6}$	$9{\times}10^{-6}$
PNB-PPE	$6{\times}10^{-4}$	$3{\times}10^{-3}$	$9{\times}10^{-3}$
PNB-PE	2×10^{-3}	9×10^{-3}	3×10^{-2}

Table 9. Probability of Random Choice Passing Full Tests

	D1			D3		
# errors	0	1	2	0	1	2
NB	0.009	0.043	0.111	0.007	0.035	0.091
BB	0.009	0.034	0.077	0.007	0.021	0.042
PNB-PPE	0.043	0.179	0.333	0.035	0.126	0.224
PNB-PE	0.043	0.179	0.333	0.041	0.147	0.238

	Both			Full		
# errors	0	1	2	0	1	2
NB	6×10^{-5}	5×10^{-4}	0.002	2×10^{-8}	3×10^{-7}	2×10^{-6}
BB	6×10^{-5}	4×10^{-4}	0.001	2×10^{-8}	2×10^{-7}	2×10^{-6}
PNB-PPE	5×10^{-4}	0.003	0.011	2×10^{-7}	2×10^{-6}	1×10^{-5}
PNB-PE	6×10^{-4}	0.004	0.012	2×10^{-7}	3×10^{-6}	2×10^{-5}

TABLE 10. Probability of Random Choice Passing each Shopping Test

APPENDIX E. ADDITIONAL DETAILS OF THE EXPERIMENTS

We provide a list of the different orders, instructions and quizzes for all experiments, sample decision sheets for each, and the payoff table for the Shopping Experiment. An experimental round of choices was stapled together with the cover sheet on the first page.

TABLE 11. Orders

Order		Risk	Social	Shopping
(D1.1,D1.2), D2, (D3.1,D3.2), D4, D5	1M	14	24	15
(D1.2,D1.1), D2, (D3.1,D3.2), D4, D5	1L	16	14	15
(D1.1,D1.2), D2, (D3.2,D3.1), D4, D5	1H	19	15	17
(D3.2,D3.1), D2, (D1.1,D1.2), D4, D5	2X[34]	0	6	0
D4, (D3.2,D3.1), D2, (D1.2,D1.1), D5	2F	15	14	16
D4, (D3.1,D3.2), D2, (D1.2,D1.1), D5	2L	18	16	20
D4, (D3.2,D3.1), D2, (D1.1,D1.2), D5	2H	17	13	18
Total		99	102	101

[34]Note that order 2X was unintended: it was printed, and run, due to a copy-and-paste error. However, we saw no reason to exclude it from our analysis.

Figure 9. Risk: Instructions

Investment task

There will be five rounds of the investment task. The first page of each round will announce the number of accounts in that round. At the end of each round, raise your hand so that the experimenter can collect your decisions and give you the decision sheet for the next round. At the end of all rounds, one round will be randomly selected to be the "round that counts". You will be paid your earnings from the round that counts based on (and only based on) your decisions in that round. Since any round could be the round that counts, you should behave in each round as if it is the round that counts.

In each round of this task, you will buy risky investments in up to two different "investment accounts". Each investment generates a return that depends on a roll of a six-sided dice. You have a separate budget for each account that can be spent only in that account. The dice will be rolled once, and you receive the returns from all your investments in all accounts in that round.

Example

As an example, suppose that in the round-that-counts you have two accounts.

You have 20 ECU in Account 1, which has two investments available; each investment costs 1 ECU per unit.

One unit in Asset A pays

$0.40 if the dice roll is 1, 2, or 3;

$0.10 if the dice roll is 4, 5, or 6.

One unit in Asset B pays

$0.25 if the dice roll is 1, 2, or 3;

$0.25 if the dice roll is 4, 5, or 6.

You have 15 ECU in Account 2, which has two investments available; each investment costs 1 ECU per unit.

One unit in Asset C pays

$0.60 if the dice roll is 1 or 2;

$0.00 if the dice roll is 3, 4, 5, or 6.

One unit in Asset D pays

$0.30 if the dice roll is 1 or 2;

$0.30 if the dice roll is 3, 4, 5, or 6.

Suppose that

In Account 1: you allocate 5 ECU to Asset A and 15 ECU to Asset B;

In Account 2: you allocate 8 ECU to Asset C and 7 ECU to Asset D.

Then, if the dice roll is **2**, you will be paid:

$$5 \times \$0.40 + 15 \times \$0.25 + 8 \times \$0.60 + 7 \times \$0.30 = \$12.65.$$

Figure 10. Risk: Quiz

Please answer the following questions and raise your hand after you have done so.

Question.

Suppose that a round has two accounts. Do your purchases in Account 1 affect what items you can afford to purchase in Account 2?

YES / NO (highlight one)

Question.

Suppose that in a round of the experiment has two accounts. Account 1 has two assets available, A and B. Account 2 has two different assets available, C and D.

Each unit of Asset A pays $0.50 if the dice roll is 1 or 2 and $1.00 if the dice roll is 3, 4, 5, or 6;

Each unit of Asset B pays $1.00 if the dice roll is 1 or 2 and $0.50 if the dice roll is 3, 4, 5, or 6.

Each unit of Asset C pays $0.50 if the dice roll is 1 or 2 and $0.00 if the dice roll is 3, 4, 5, or 6;

Each unit of Asset D pays $0.00 if the dice roll 1 or 2 and $1.00 if the dice roll is 3, 4, 5, or 6.

Suppose that you invest as follows:

in Account 1, you invest 2 ECU in Asset A and 6 ECU in Asset B;

in Account 2, you invest 4 ECU in Asset C and 2 ECU in Asset D.

1. If this round determines your payment, then how much will you earn if the dice roll is 2?

   ____

2. If this round determines your payment, then how much will you earn if the dice roll is 6?

   ____

Figure 11. Risk: Cover Sheet for Round 1 (order 1M)

Subject # Session

Round 1

In round 1, you have 2 investment accounts.

Figure 12. Risk: D1.1 Decision Sheet

Investment Account 1

You have **10 ECU** available in Account 1. Two assets are available for purchase, Asset A and Asset B.

The price of Asset A is **1 ECU per unit**.

The price of Asset B is **1 ECU per unit**.

One unit in Asset A pays

$1.00 if the dice roll is 1, 2, or 3;

$0.00 if the dice roll is 4, 5, or 6.

One unit in Asset B pays

$0.00 if the dice roll is 1, 2, or 3;

$1.20 if the dice roll is 4, 5, or 6.

Please highlight a feasible combination of purchases of Asset A and Asset B from the list below.

0 units of Asset A and 10 units of Asset B.

1 unit of Asset A and 9 units of Asset B.

2 units of Asset A and 8 units of Asset B.

3 units of Asset A and 7 units of Asset B.

4 units of Asset A and 6 units of Asset B.

5 units of Asset A and 5 units of Asset B.

6 units of Asset A and 4 units of Asset B.

7 units of Asset A and 3 units of Asset B.

8 units of Asset A and 2 units of Asset B.

9 units of Asset A and 1 unit of Asset B.

10 units of Asset A and 0 units of Asset B.

FIGURE 13. Risk: D1.1 Decision Sheet

Investment Account 2

You have **16 ECU** available in Account 2. Two assets are available for purchase, Asset C and Asset D.

The price of Asset C is **1 ECU per unit**.

The price of Asset D is **1 ECU per unit**.

One unit in Asset C pays

$1.00 if the dice roll is 1, 2, or 3;

$0.00 if the dice roll is 4, 5, or 6.

One unit in Asset D pays

$0.00 if the dice roll is 1, 2, or 3;

$1.00 if the dice roll is 4, 5, or 6.

Please highlight a feasible combination of purchases of Asset C and Asset D from the list below.

0 units of Asset C and 16 units of Asset D.

1 unit of Asset C and 15 units of Asset D.

2 units of Asset C and 14 units of Asset D.

3 units of Asset C and 13 units of Asset D.

4 units of Asset C and 12 units of Asset D.

5 units of Asset C and 11 units of Asset D.

6 units of Asset C and 10 units of Asset D.

7 units of Asset C and 9 units of Asset D.

8 units of Asset C and 8 units of Asset D.

9 units of Asset C and 7 units of Asset D.

10 units of Asset C and 6 units of Asset D.

11 units of Asset C and 5 units of Asset D.

12 units of Asset C and 4 units of Asset D.

13 units of Asset C and 3 units of Asset D.

14 units of Asset C and 2 units of Asset D.

15 units of Asset C and 1 unit of Asset D.

16 units of Asset C and 0 units of Asset D.

FIGURE 14. Social: Instructions

Division task

There will be five rounds of a task where you will asked to allocate tokens between two other participants who will herein be labelled “person A” and “person B”. They will not be told your identity, and you will not be told their identities. That is, you will remain completely anonymous to each other.

In each round of this task, you will have tokens in up to two different accounts. You decide how to allocate tokens between person A and person B in each account. The value per token allocated to each of A and B may vary across rounds and across accounts. You have a separate budget of tokens for each account that can be allocated only in that account. Payments for a given round will be determined by the sum of the value of all tokens allocated in all accounts in that round.

The first page of each round will announce the number of accounts in that round. At the end of each round, raise your hand so that the experimenter can collect your decisions and give you the decision sheet for the next round.

You and every other participant has numbered a sealed envelope at the beginning of the experiment. Each participant has been randomly allocated to a group and role (A or B); this is recorded in the envelope. The round that counts to determine your payment has also been randomly selected and recorded in each envelope. Your group has been randomly and anonymously matched to determine the payment of another group and one round of your choices will determine the earnings of person A and person B in that group. Since each round could be the round that counts and actually determines a two other subjects’ payments, you should treat each round as if it is the round that counts.

Example

As an example, suppose that in the round-that-counts there are two accounts.

There are 10 tokens in Account 1.

One token is worth $0.80 to A and $0.60 to B.

There are 12 tokens in Account 2.

One token pays $1.00 to A and $0.20 to B.

Suppose that

In Account 1: you allocate 4 tokens to A and 6 tokens to B.

In Account 2: you allocate 2 tokens to A and 10 tokens to B.

Then,

A’s earnings are $4 \times \$0.80 + 2 \times \$1.00 = \$5.20$;

B’s earnings are $6 \times \$0.60 + 10 \times \$0.20 = \$5.60$.

Figure 15. Social: Quiz

Please answer the following questions and raise your hand after you have done so.

Question.

Suppose that a round has two accounts. Does your allocation in Account 1 affect what you have available to allocate in Account 2?

YES / NO (highlight one)

Question.

Suppose that a round of the experiment has two accounts.

In Account 1, each token pays $0.40 to A and $0.60 to B.

In Account 2, each token pays $0.30 to A and $0.40 to B.

Suppose that you invest as follows:

in Account 1, you allocate 2 tokens to A and 4 tokens to B;

in Account 2, you allocate 6 tokens to A and 1 token to B.

1. If this is the round that counts for this group, then how much will person A receive?

   ____

2. If this is the round that counts for this group, then how much will person B receive?

   ____

FIGURE 16. Social: D5 Decision Sheet

Account 1

You have **10 tokens** available in Account 1.

Each token allocated to A is worth \$1.00.

Each token allocated to B is worth \$1.20.

Please highlight a feasible allocation of tokens between A and B.

0 tokens for A and 10 tokens for B.

1 token for A and 9 tokens for B.

2 tokens for A and 8 tokens for B.

3 tokens for A and 7 tokens for B.

4 tokens for A and 6 tokens for B.

5 tokens for A and 5 tokens for B.

6 tokens for A and 4 tokens for B.

7 tokens for A and 3 tokens for B.

8 tokens for A and 2 tokens for B.

9 tokens for A and 1 token for B.

10 tokens for A and 0 tokens for B.

Figure 17. Shopping: Instructions

Shopping Task

There will be five rounds of the shopping task. At the end of all rounds of the experiment, one round will be randomly selected to be the "round that counts". You will be paid your earnings from the round that counts based on (and only based on) your decisions in that round. Since any round could be the round that counts, you should behave in each round as if it is the round that counts.

In each round of this task, you will buy up to two different fictitious "fruits" at up to two "stores". You have a separate gift certificate (denominated in experimental currency units – ECUs) at each store that can be spent only at that store. However, your monetary earnings for the experiment are based on the total amount of each fruit in your final bundle for a round after you have completed your shopping at all stores.

The first page of each round will announce the number of stores in that round. At the end of each round, raise your hand so that the experimenter can collect your decisions and give you the decision sheet for the next round.

How Your Payment is Determined

Your monetary payment will be calculated from your final bundle in the round that counts according to the function

$$Payment = \frac{2}{5}\left(\sqrt{\#apples} + \sqrt{\#oranges}\right)^2.$$

To help you calculate the payment you would receive for a final bundle, we have provided tables at the end of the experiment that indicates the payment that would result from all possible final bundles (and some impossible ones).

As an example of how your payment will be calculated, suppose you buy:

1 apple and 5 oranges at Store 1,

2 apples and 6 oranges at Store 2.

Then your final bundle is

3 apples and 11 oranges.

To calculate your payment locate the entry in the "3 apples" column and the "11 oranges" row of the payment table.

Notice three features of the payment table:

(i) A final bundle with more of every fruit earns a higher payment.
(ii) A mix of fruits earns a higher payment: a final bundle with 5 apples and 5 oranges earns you a higher payment than a final bundle with 8 apples and 2 oranges, which in turn earns a higher payment than a final bundle with 10 apples and 0 oranges.
(iii) A final bundle with 7 apples and 3 oranges earns the same final payment as a final bundle with 3 apples and 7 oranges.

If the prices of apples and oranges are not the same, you thus face a trade-off between buying as many units of fruit as possible versus buying a mix that includes both fruits.

FIGURE 18. Shopping: Quiz

How to Shop in each Store

You will have a separate gift certificate at each store denominated in Experimental Currency Units (ECUs). The page for each store will present you with the prices of the fruits in that store. You must highlight one of the feasible apple-orange-watermelon combinations at each store to spend your gift certificate. Feasible combinations will be denoted in a list. If that combination does not appear in the, then it is not affordable with your gift certificate at that store.

To illustrate how you make your decision in each store, consider the following hypothetical store; you have a 6 ECU gift certificate for this store, and apples and oranges each cost 1 ECU per unit of fruit. Then your store page will be laid out as follows.

Store

You have a **6 ECU** gift certificate to spend.

The price of apples is **1 ECU per apple**.

The price of oranges is **1 ECU per orange**.

Please highlight a feasible combination of apples and oranges from the list below to make your purchase from this store.

0 apples and 6 oranges.

1 apple and 5 oranges.

2 apples and 4 oranges.

3 apples and 3 oranges.

4 apples and 2 oranges.

5 apples and 1 orange.

6 apples and 0 oranges.

Question 1.

How much would you earn if a round with only the store above was the round that counts, and you had chosen the bundle you indicated above?

Question 2.

Suppose that the round that counts had two stores. In Store 1, you bought 1 apple and 4 oranges. In store 2, you bought 3 apples and 5 oranges. What would your earnings be for the experiment?

Figure 19. Shopping: D3.2 Decision Sheet

Store 2

You have a **24 ECU** gift certificate at Store 2.

The price of apples is **3 ECU per apple**.

The price of oranges is **2 ECU per orange**.

Please highlight a feasible combination of apples and oranges from the list below to make your purchase from this store.

0 apples and 12 oranges.

0 apples and 11 oranges.

1 apple and 10 oranges.

2 apples and 9 oranges.

2 apples and 8 oranges.

3 apples and 7 oranges.

4 apples and 6 oranges.

4 apples and 5 oranges.

5 apples and 4 oranges.

6 apples and 3 oranges.

6 apples and 2 oranges.

7 apples and 1 orange.

8 apples and 0 oranges.

FIGURE 20. Shopping Payoff Table

	0 apples	1 apple	2 apples	3 apples	4 apples	5 apples	6 apples	7 apples	8 apples	9 apples	10 apples	11 apples	12 apples	13 apples	14 apples	15 apples	16 apples	17 apples	18 apples	19 apples	20 apples	21 apples	22 apples
0 oranges	0.00	0.40	0.80	1.20	1.60	2.00	2.40	2.80	3.20	3.60	4.00	4.40	4.80	5.20	5.60	6.00	6.40	6.80	7.20	7.60	8.00	8.40	8.80
1 orange	0.40	1.60	2.33	2.99	3.60	4.19	4.76	5.32	5.86	6.40	6.93	7.45	7.97	8.48	8.99	9.50	10.00	10.50	10.99	11.49	11.98	12.47	12.95
2 oranges	0.80	2.33	3.20	3.96	4.66	5.33	5.97	6.59	7.20	7.79	8.38	8.95	9.52	10.08	10.63	11.18	11.73	12.26	12.80	13.33	13.86	14.38	14.91
3 oranges	1.20	2.99	3.96	4.80	5.57	6.30	6.99	7.67	8.32	8.96	9.58	10.20	10.80	11.40	11.98	12.57	13.14	13.71	14.28	14.84	15.40	15.95	16.50
4 oranges	1.60	3.60	4.66	5.57	6.40	7.18	7.92	8.63	9.33	10.00	10.66	11.31	11.94	12.57	13.19	13.80	14.40	15.00	15.59	16.17	16.76	17.33	17.90
5 oranges	2.00	4.19	5.33	6.30	7.18	8.00	8.78	9.53	10.26	10.97	11.66	12.33	13.00	13.65	14.29	14.93	15.56	16.18	16.79	17.40	18.00	18.60	19.19
6 oranges	2.40	4.76	5.97	6.99	7.92	8.78	9.60	10.38	11.14	11.88	12.60	13.30	13.99	14.67	15.33	15.99	16.64	17.28	17.91	18.54	19.16	19.78	20.39
7 oranges	2.80	5.32	6.59	7.67	8.63	9.53	10.38	11.20	11.99	12.75	13.49	14.22	14.93	15.63	16.32	17.00	17.67	18.33	18.98	19.63	20.27	20.90	21.53
8 oranges	3.20	5.86	7.20	8.32	9.33	10.26	11.14	11.99	12.80	13.59	14.36	15.10	15.84	16.56	17.27	17.96	18.65	19.33	20.00	20.66	21.32	21.97	22.61
9 oranges	3.60	6.40	7.79	8.96	10.00	10.97	11.88	12.75	13.59	14.40	15.19	15.96	16.71	17.45	18.18	18.90	19.60	20.30	20.98	21.66	22.33	23.00	23.66
10 oranges	4.00	6.93	8.38	9.58	10.66	11.66	12.60	13.49	14.36	15.19	16.00	16.79	17.56	18.32	19.07	19.80	20.52	21.23	21.93	22.63	23.31	23.99	24.67
11 oranges	4.40	7.45	8.95	10.20	11.31	12.33	13.30	14.22	15.10	15.96	16.79	17.60	18.39	19.17	19.93	20.68	21.41	22.14	22.86	23.57	24.27	24.96	25.65
12 oranges	4.80	7.97	9.52	10.80	11.94	13.00	13.99	14.93	15.84	16.71	17.56	18.39	19.20	19.99	20.77	21.53	22.29	23.03	23.76	24.48	25.19	25.90	26.60
13 oranges	5.20	8.48	10.08	11.40	12.57	13.65	14.67	15.63	16.56	17.45	18.32	19.17	19.99	20.80	21.59	22.37	23.14	23.89	24.64	25.37	26.10	26.82	27.53
14 oranges	5.60	8.99	10.63	11.98	13.19	14.29	15.33	16.32	17.27	18.18	19.07	19.93	20.77	21.59	22.40	23.19	23.97	24.74	25.50	26.25	26.99	27.72	28.44
15 oranges	6.00	9.50	11.18	12.57	13.80	14.93	15.99	17.00	17.96	18.90	19.80	20.68	21.53	22.37	23.19	24.00	24.79	25.57	26.35	27.11	27.86	28.60	29.33
16 oranges	6.40	10.00	11.73	13.14	14.40	15.56	16.64	17.67	18.65	19.60	20.52	21.41	22.29	23.14	23.97	24.79	25.60	26.39	27.18	27.95	28.71	29.46	30.21
17 oranges	6.80	10.50	12.26	13.71	15.00	16.18	17.28	18.33	19.33	20.30	21.23	22.14	23.03	23.89	24.74	25.57	26.39	27.20	27.99	28.78	29.55	30.32	31.07
18 oranges	7.20	10.99	12.80	14.28	15.59	16.79	17.91	18.98	20.00	20.98	21.93	22.86	23.76	24.64	25.50	26.35	27.18	27.99	28.80	29.59	30.38	31.15	31.92
19 oranges	7.60	11.49	13.33	14.84	16.17	17.40	18.54	19.63	20.66	21.66	22.63	23.57	24.48	25.37	26.25	27.11	27.95	28.78	29.59	30.40	31.19	31.98	32.76
20 oranges	8.00	11.98	13.86	15.40	16.76	18.00	19.16	20.27	21.32	22.33	23.31	24.27	25.19	26.10	26.99	27.86	28.71	29.55	30.38	31.19	32.00	32.80	33.58
21 oranges	8.40	12.47	14.38	15.95	17.33	18.60	19.78	20.90	21.97	23.00	23.99	24.96	25.90	26.82	27.72	28.60	29.46	30.32	31.15	31.98	32.80	33.60	34.40
22 oranges	8.80	12.95	14.91	16.50	17.90	19.19	20.39	21.53	22.61	23.66	24.67	25.65	26.60	27.53	28.44	29.33	30.21	31.07	31.92	32.76	33.58	34.40	35.20