

I Choose For You: an Experimental Study*

Marina Agranov[†] Federico Echenique[‡] Kota Saito[§]

January 20, 2026

Abstract

We investigate whether risk and time preferences differ when individuals make decisions for others compared to making decisions for themselves. We introduce a novel “skin in the game” experimental design, where choices for others incur a direct cost to the decision-maker, ensuring a genuine trade-off between self-interest and surrogate allocation. The modal outcome is that participants are more risk-averse and impatient when choosing for others than for themselves. Our methodology reveals significant heterogeneity, successfully identifying selfish types often missed by the more standard “no skin in the game” approaches. The message is nuanced, as even non-selfish participants behave differently when they have skin in the game. Furthermore, our framework yields more consistent behavior and superior out-of-sample predictive power.

*We thank Ben Gillen for advice on econometric methodology.

[†]Division of the Humanities and Social Sciences, Caltech and NBER. Contact: maganov@caltech.edu

[‡]Department of Economics, UC Berkeley. Contact: fede@econ.berkeley.edu

[§]Division of the Humanities and Social Sciences, Caltech. Contact: saito@caltech.edu

Contents

1	Introduction	3
2	Model	7
2.1	No Skin In The Game Method	9
2.2	Skin In The Game Method	9
3	Experimental Design	10
3.1	Ordinal Comparison of Preferences: Identification	10
3.2	Experimental Protocol	13
3.3	Discussion of Experimental Design	16
4	Results	18
4.1	Identifying selfish types	19
4.2	SIG, non-parametric analysis	20
4.3	Comparing SIG and no-SIG, non-parametric analysis	22
4.4	SIG, structural parametric estimation	24
4.5	Predictive ability of SIG and no-SIG methods	27
A	Parameters	29
B	Additional Analysis	30

1 Introduction

Economists rely heavily on understanding how individuals navigate trade-offs involving risk and time. From household savings decisions to firm investment choices, these two dimensions — risk aversion and intertemporal substitution — are fundamental to modeling decision-making. Typically, preferences over risk and time are elicited and analyzed under the assumption that the decision-maker is also the beneficiary of the outcome. We observe choices made by the self, for the self, and recover the underlying parameters that govern these behaviors.

A substantial portion of economic decision-making involves choices made on behalf of others. Financial advisors manage roughly \$30 trillion in U.S. household assets, making risk and timing decisions for clients; physicians select treatments for patients who defer to their expertise; and parents open investment accounts for children who cannot yet decide for themselves. In these surrogate settings, one person’s preferences govern another person’s consumption. Yet virtually all economic analysis of such surrogate relationships implicitly assumes that decision-makers apply the same risk and time preferences they would use for their own consumption. Whether this assumption is correct is far from obvious. This is what we do in this paper. We study whether the fundamental parameters of risk and time preference remain stable across the self–other divide, or whether the act of choosing for another induces systematically different behavior.

Specifically, we aim to recover the preferences that one subject holds over the consumption of another to determine whether there is a divergence in risk attitudes and impatience. The central empirical questions are whether individuals are more risk-averse when making choices for a third party compared to themselves and whether they exhibit different levels of impatience when scheduling consumption for others. Understanding this “self-other wedge” is critical for delegation and agency relationships, as it suggests that misalignments may arise not just from differing information or incentives, but also from fundamental differences in how we process risk and time for others. In other words, in analyzing choices made on behalf of others, we are focused on misalignments that stem from differences in preferences and are more fundamental, or hard-wired, than the more common concern with conflicts of interest.

Our primary methodological contribution lies in the specific mechanism used to elicit these preferences. A standard experimental approach might attempt to study this wedge through what we call the “no-skin-in-the-game” design: a subject makes

choices that determine the consumption of a passive recipient, but without affecting their own payoff. While this setting isolates the decision-maker’s pure preference for the other’s welfare, it abstracts away from the material constraints and opportunity costs that characterize most real-world agency problems. Crucially, it is only incentivized to the extent that an experimental participant truly cares about the outcome consumed by others.

In contrast, we introduce a novel “skin-in-the-game” framework. In our design, the act of making choices on behalf of another is not costless; rather, it comes at the direct cost of reduced consumption for the decision-maker. This ensures that the decision-maker faces a genuine trade-off: providing a specific risk profile or time schedule for the other agent requires a sacrifice of their own resources. By linking the choices made for the beneficiary to a reduction in the decision-maker’s own potential consumption, we move beyond a costless allocation. The design is explained in detail in Section 3. Our approach allows us to observe how surrogate risk and time preferences manifest when the decision-maker is personally invested in the allocation process and must weigh the beneficiary’s interests against their own material well-being.

Our experimental results show significant individual heterogeneity, but the most common pattern is that participants are more risk averse and more impatient with respect to others’ consumption than they are with their own consumption. See Section 4 for details. The skin-in-the-game methodology successfully identifies “selfish types,” which constitute a significant fraction of participants and allow for a more nuanced understanding of agents’ preferences over others’ consumption than the simple no-skin-in the-game approach.

The comparison with the no-skin method yields some interesting conclusions. The skin in the game results are consistent across tasks and parameterizations, while the no-skin results are much less consistent. Our skin in the game estimations perform better when evaluated out of sample (one question is left out of the analysis and is used as a hold-out for predictions) than the same exercise using no-skin data. So, we conclude that the skin in the game method has some clear advantages compared to having no skin in the game. This finding is significant, as the relative attitudes towards one’s own and others’ consumption differ across the two methodologies, with the no-skin method underestimating how many participants are more risk averse for others than for themselves. It also underestimates how many participants have similar attitudes towards risk for themselves and others.

The difference in results between skin and no-skin cannot be explained by the mere presence of selfish agents who remain undetected in the no-skin treatment. Even non-selfish agents behave differently in the no-skin tasks compared to the skin in the game method. The incentives provided in the skin approach matter for the conclusions we draw about agents' choices on behalf of others. Given the apparent statistical benefits of the skin in the game results (stability across tasks and out-of-sample performance), we are inclined to put more weight on the conclusions from the skin, than the no-skin, treatments.

The remainder of the paper is organized as follows. We conclude this section with a review of the related literature. Section 2 introduces the conceptual framework and outlines two approaches to studying how people make decisions for others: the skin-in-the-game and no-skin-in-the-game methods. Section 3 describes the experimental design, focusing on the identification strategy (Section 3.1) and the main features of the experimental protocol (Sections 3.2 and 3.3). Section 4 presents the experimental results, beginning with non-parametric analyses (Sections 4.2 and 4.3), followed by structural estimation (Section 4.4) and out-of-sample predictions (Section 4.5).

Related literature The skin in the game design is inspired by the work of [Andreoni and Miller \[2002\]](#), who developed the modified dictator game in which subjects share surplus with another participant and do so for a variety of costs of sharing and amounts available for sharing. [Andreoni and Miller \[2002\]](#) use this task to study the axioms of revealed preference in the domain of altruism. [Fisman et al. \[2007\]](#) provide a graphical representation of the same task and recovers the underlying preferences for giving, taking special care to distinguish such preferences from the social preferences. [Fisman et al. \[2007\]](#) find that about 25% of people are selfish, which is in the same ballpark as our results. We extend this idea to allocating the available budget between inputs in a richer set of objects, such as the high prize in the lottery, consumed either by oneself or by others, or the delayed payments.

Following what we have called the no-skin design, [Chakravarty et al. \[2011\]](#) find that people are less risk averse when choosing on behalf of others than when they choose for themselves. The authors present two experimental designs. The first were multiple-price lists (MPLs) involving risky outcomes for the decision makers, and the second were identical MPLs for which the outcomes would be enjoyed by a randomly-drawn set of passive participants in the experiment. The subject pool consisted

of management students in India. The paper documents a significantly more risk-averse behavior when choosing for self than when choosing for others. The second experimental design in the paper involves choices in first-price sealed-bid auctions and confirms qualitatively the findings from the first experiment. To quote from the paper’s conclusions “. . . individuals appear less risk averse when making decisions over OTHER people’s money, and some are actually risk loving. This general pattern is statistically significant. The difference does not appear to be driven by an attempt to pick risk attitudes that reflect the risk attitudes of others.”

Reynolds et al. [2009] run an experiment in which each participant first chooses between a risky and a safe option for themselves; and then makes a similar choice for the benefit of a group of other people. They find that when choosing for themselves, participants are more likely to choose the risky option than when choosing on behalf of others. Thus they find, using a no-skin design, that people are more risk averse when choosing on behalf of others.

Harrison et al. [2012] consider an individual multiple-price list choice among lotteries, and a social outcome through voting. They find no significant difference in risk preferences between individual and social choices, unless the participants are provided with information about the level of risk aversion of the other group members, in which case they are more risk averse when choosing socially. The design is quite different from ours, though, in that the social treatment involves majority voting in groups. Interestingly, Harrison et al. [2012] attribute the difference between their results and those of Chakravarty et al. [2011] to the lack of skin in the game in the latter design.

Exley [2016] designs an experiment to understand the effect of risk on charitable giving. While the motivation and goals of her study are distinct from ours, her experimental design is probably the closest to our design. The core of Exley’s experiment is a series of multiple price lists (MPL) that involve outcomes for the decision maker and for a charity. These outcomes can be certain monetary rewards or lotteries. The MPLs are arranged in four blocks. One block, which she calls “Self Lottery vs. Charity-Certain Amount,” is the closest to our design. The block involves two kinds of choices. First, binary choices between a lottery for the participant and a certain amount for the charity. This elicits a kind of certainty equivalent expressed in the participant’s preferences for the charity’s welfare. Second, a choice between a lottery for the charity and a certain amount for the charity.

In her design, Exley varies the probability of winning in the lottery over a fine-

graded scale (the set $\{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95\}$). The reason is that she is interested in the effect of uncertainty on selfishness. The main conclusion of the experiment is that risk acts as an excuse to be selfish. This is reflected in a larger aversion to donating to the charity when the lottery that the charity would obtain is subject to a larger degree of risk. While the MPL's in Exley's design are similar to ours, the use that she makes of these is orthogonal to our interest in comparing preferences over own and others' consumption.

Finally, [Chakraborty and Henkel \[2025\]](#) consider the allocation of resources among one participant who belongs to a similar social group than does the decision maker, and one who does not. The focus in their experiment is to understand the role of interpersonal uncertainty: the degree to which it is difficult for one person to assess the welfare of another. To this end, they construct alternative tasks that have no social component but that feature similar levels of uncertainty. We share with this paper the preoccupation for preferences on behalf of others, but the experimental designs, and overall goals of the two studies, are quite distinct.

2 Model

We first discuss our methodological proposal in an abstract setting. The methodology is broadly applicable to many different domains and questions in economics. The model primitives are $(\Theta, O, o, Q, \succeq)$ where:

- Θ is a finite set of options.
- O is a set of *outcomes*.
- \succeq is a preference relation over O .
- $o : \Theta \times \mathbf{R}_+ \rightarrow O$ is a function mapping each option θ and each *scale* $x \geq 0$ into an *outcome* $o(\theta, x)$.
- Q is a finite set of *exchange rates*.

Consider two agents: ME and OTHER. The outcomes in O will occur to (or be enjoyed by) OTHER. The preference \succeq describes ME's choices over outcomes for OTHER. Suppose that ME has a strictly increasing and continuous utility $v : \mathbf{R}_+ \rightarrow \mathbf{R}$ over ME's own consumption of money. Suppose also that \succeq is represented by a utility function $w : O \rightarrow \mathbf{R}$.

Example 1. Let O be the set of all simple lotteries on some given set of monetary prizes: $O = \Delta_s(\mathbf{R})$ is the set of probability distributions on \mathbf{R} with finite support. Let Θ be a set of options, θ_p with $p = 0.1, 0.2, 0.3$. The function o assigns, to each θ_p and scalar $x \geq 0$, the lottery in O that pays x with probability p and 0 with probability $1 - p$.

Suppose that the utility function w over lotteries in O takes the expected utility form. For example, $w(\ell_p) = pu(x) + (1 - p)u(0)$, when $\ell_p \in O$ is the lottery that pays x with probability p and 0 with probability $1 - p$. The function u may, in addition, be assumed to be of the constant relative risk aversion form. So we get that $w(\ell_p) = Ap x^\gamma$, for some parameters $A > 0$ and $\gamma > 0$.

How can we elicit, or estimate, the preferences \succeq and the utility w ? The standard procedure in economics is to ask ME to make choices in O that will be consumed by OTHER. Essentially, preferences are choices. In our paper, we propose a different approach. It is also rooted in the notion that preferences are choices, but we offer a different channel for experimentally incentivizing those choices.

We shall impose the following assumptions on the model primitives:

1. Separability: $o(\theta, x) \succeq o(\theta', x)$ if and only if $o(\theta, x') \succeq o(\theta', x')$ for any $\theta, \theta' \in \Theta$ and $x, x' > 0$.
2. $x \mapsto w(o(\theta, x))$ is a strictly monotonically increasing, and continuous, function.
3. $v(1 - x) \geq \max\{w(\theta, qx) : \theta \in \Theta, q \in Q\}$ for all x small enough, and $v(1 - x) \leq \min\{w(\theta, qx) : \theta \in \Theta, q \in Q\}$ for all x large enough.

Separability assumptions are extremely common in economics. We use them to define a preference on some relevant sub-domain: for example, the separation of tastes from beliefs requires separability assumptions, as does the separation of time preferences from consumption within a period. Here we impose separability in order to define a preference \succeq^* over Θ by $\theta \succeq^* \theta'$ if and only if $o(\theta, x) \succeq o(\theta', x)$ for some $x > 0$. We may interpret the preferences \succeq^* as representing the choices that ME would make for OTHER to consume when the choice involves alternatives in Θ .

We now discuss two approaches to recover ME's preferences over OTHER's consumption. The first approach, "no skin in the game," is the standard procedure in the literature. The second approach, "skin in the game," is our main proposal. It is important to understand the differences between the two, not only because it matters

for our methodological contribution, but also for our experimental results. Our experimental design includes tasks given by both methodologies, and we are interested in the empirical results obtained under each of the two approaches.

2.1 No Skin In The Game Method

The most common experimental design asks participants to choose between two alternatives, $o(\theta, x)$ and $o(\theta', x)$. We call this design, no-skin-in-the-game method.

In principle, such a design allows us to recover the preferences that ME has over OTHER’s consumption. These would be the preferences \succeq^* in our previous discussion. However, this method is only incentivized to the extent that ME truly cares about OTHER’s consumption. The choices made by ME are physically paid out, and enjoyed, by OTHER. There is no “physical” connection between OTHER’s consumption and ME’s material consumption, hence the name, no-skin-in-the-game.

2.2 Skin In The Game Method

Our main methodological proposal is to make ME’s choices be costly for ME. Our proposal is inspired by [Andreoni and Miller \[2002\]](#), who allow for a tradeoff between ME and OTHER’s consumption at varying terms of trade. In our model, these terms of trade are given by the exchange rates $q \in Q$. We ask ME to choose between ME receiving an amount of money or giving money that scales an outcome that OTHER gets. The exchange rate q multiplies how goes into the scale (x) of the outcome that OTHER receives.

The total amount of money is m , which we assume for now is $m = 1$. Given x , then, ME chooses between getting $m - x$ or giving OTHER the outcome $o(\theta, qx)$. Under the assumptions we have made, the *cutoff* $x^*(\theta, q)$ is well defined by

$$v(1 - x^*(\theta, q)) = w(o(\theta, qx^*(\theta, q)))$$

The cutoff $x^*(\theta, q)$ is the monetary amount that equates ME’s utility from ME consuming $1 - x^*(\theta, q)$ with the utility that ME obtains from OTHER consuming the object $o(\theta, qx^*(\theta, q))$. This object could be a lottery, as in [Example 1](#), a dated reward, or other consumption items.

The assumptions we have made are minimal, but they suffice for the method we have outlined to enable recovery of ME’s preference for consumption by OTHER.

Proposition 1. $\theta \succeq^* \theta'$ if and only if $x^*(\theta, q) \leq x^*(\theta', q)$.

Proof. Let $x = x^*(\theta, q)$. Then $\theta \succeq^* \theta'$ if and only if

$$0 = v(1 - x) - w(o(\theta, qx)) \leq v(1 - x) - w(o(\theta', qx)).$$

So $x \leq x^*(\theta', q)$ as $x \mapsto v(1 - x) - w(o(\theta', qx))$ is strictly decreasing. \square

The skin-in-the-game method provides a direct connection between ME’s choice of OTHER’s consumption and ME’s own consumption. Choosing an amount x means two things: it means that OTHER consumes the object $o(\theta, qx) \in O$ and that ME consumes the monetary amount $1 - x$.

Note the role played by the cutoff $x^*(\theta, q)$ in Proposition 1. Operationally, in our experimental design, we shall use a multiple price list (MPL) to elicit a point of indifference between consumption for ME and for OTHER.

3 Experimental Design

We designed an experiment that elicits both ordinal and cardinal comparisons of participants’ preferences over outcomes consumed by themselves (ME) and by another individual (OTHER), allowing us to compare the SIG and no-SIG methods. As shown below, our methodology is broadly applicable. In this paper, we illustrate it in two domains that are central to the literature: risk and time preferences.

We begin by describing the types of questions our method uses and the non-parametric identification strategy for comparing preferences over outcomes for ME and OTHER (Section 3.1). We then present a detailed overview of the experimental protocol (Section 3.2) and discuss its key design features (Section 3.3).

3.1 Ordinal Comparison of Preferences: Identification

Our **skin-in-the-game** method (**SIG**) centers on three core, incentive-compatible questions that pin down how people value objects consumed by either ME or OTHER. We recover these valuations from an indifference condition: the point at which the utility ME gets from OTHER consuming an object exactly matches the utility ME gets from ME consuming a complementary object. In this way, ME’s utility from

OTHER’s consumption becomes directly comparable – and commensurate – to ME’s own consumption utility.

Let $v(o)$ and $w(o)$ denote the ME’s utility for outcome o when consumed by ME and OTHER, respectively. In the risk domain, outcomes $o \in O$ are monetary lotteries; in the time domain, they are consumption bundles specifying monetary payments delivered at different dates. For convenience, we slightly abuse notation and write $v(x, t)$ and $w(x, t)$ for the Bernoulli utility of ME for receiving a payment of $\$x$ in t days by ME and OTHER, respectively. For example, $v(5, 0)$ denotes ME’s utility of ME consuming $\$5$ today, while $w(5, 14)$ denotes ME’s utility of OTHER consuming $\$5$ in 14 days.

The three core questions are valuation tasks, implemented via a multiple-price list (see Section 3.2). These questions elicit three indifference points— (x^*, y^*, z^*) —in each domain, as specified below:

- Risk domain

$$(Q1) \quad v(K - x^*, 0) = p \cdot v(a(x^* + c), 0) + (1 - p) \cdot v(0, 0)$$

$$(Q2) \quad w(q(K - x^*), 0) = v(y^*, 0)$$

$$(Q3) \quad p \cdot w(qa(x^* + c), 0) + (1 - p) \cdot w(0, 0) = v(z^*, 0)$$

- Time domain

$$(Q1) \quad v(K - x^*, 0) = v(a(x^* + c), t)$$

$$(Q2) \quad w(q(K - x^*), 0) = v(y^*, 0)$$

$$(Q3) \quad w(qa(x^* + c), t) = v(z^*, 0)$$

The parameters p and t specify the main features of the studied objects: p is the probability of a high prize in the risk domain, while t is the delay in the time domain. The parameter q scales dollar payments to ME when the same payment is received by OTHER. Think of q as a price. It is the exchange rate between consumption by ME and by OTHER, as in [Andreoni and Miller \[2002\]](#).¹ Finally, the remaining parameters (K, a, c) are constants calibrated to target empirically relevant regions of preferences.

¹[Andreoni and Miller \[2002\]](#) vary q substantially across different treatments. We do not. This reflects an important difference between their objectives and ours. They wanted to test for rationality, which requires looking at the response to price variation. We instead choose q so as to make the tradeoff between ME and OTHER’s consumption be meaningful.

All parameters are fixed in advance and known to the participants. We report these parameters and discuss their role in Section 3.2.

In the risk domain, the first question, Q1, elicits the amount x^* for which ME is indifferent between receiving $K - x^*$ for sure and a lottery paying $a(x^* + c)$ with probability p and zero otherwise. The next two questions, Q2 and Q3, elicit y^* and z^* , which capture ME's valuation of OTHER receiving the same fixed amount $K - x^*$ and the same lottery, respectively, where other's payments are scaled by q .

The three values (x^*, y^*, z^*) suffice to compare ME's risk preferences for outcomes consumed by ME versus OTHER. If $y^* < z^*$ then ME is more risk-averse for lotteries consumed by ME than for OTHER. This follows directly from the monotonicity assumption: $y^* < z^*$ implies that $v(y^*, 0) < v(z^*, 0)$, which in turn means that ME's certainty equivalent for consuming the lottery paying $a(x^* + c)$ with probability p (and zero otherwise) is smaller than ME's certainty equivalent for OTHER consuming the same lottery. If, however, $y^* > z^*$ then ME is less risk-averse for lotteries consumed by ME than for OTHER.

Similarly, in the time domain, Q1 elicits the amount x^* for which ME is indifferent between receiving $K - x^*$ today and $a(x^* + c)$ in t days. Questions Q2 and Q3 then elicit y^* and z^* , which capture ME's valuation of OTHER receiving the same amounts today and in t days, respectively. As always, both payments are scaled by the exchange rate q . If $y^* < z^*$, then ME is more impatient for delays in own payments compared to other's payments, since $y^* < z^*$ implies that $v(y^*, 0) < v(z^*, 0)$, which in turn indicates that the certainty equivalent of ME receiving $a(x^* + c)$ in t days is smaller than certainty equivalent of OTHER receiving the same delayed amount scaled by q . Conversely, if $y^* > z^*$, ME is more patient for delays in own payments than in other's payments.

The **no-skin-in-the-game** method, the **no-SIG** hereafter, instead relies on the comparison of the values elicited using Q1 (specified above) and one additional question:

- Risk domain

$$(Q4) \quad p \cdot w(a(x^* + c), 0) + (1 - p) \cdot w(0, 0) = w(m^*, 0)$$

- Time domain

$$(Q4) \quad w(a(x^* + c), t) = w(m^*, 0)$$

In the risk domain, Q4 elicits the certainty equivalent of the same lottery as in Q1, but consumed by OTHER and expressed in terms of OTHER’s sure amount m^* . An important feature of Q4 is that the decision-maker (ME) focuses entirely on choosing outcomes for OTHER, without any reference to their own outcomes. This differs from Q2 and Q3, where ME evaluates objects for OTHER relative to objects consumed by ME. Using (x^*, m^*) , we can compare the ordinal risk preferences of ME for outcomes consumed by ME versus OTHER: if $K - x^* < m^*$, then ME is more risk-averse for lotteries consumed by ME than for those consumed by OTHER, since the certainty equivalent of the same lottery in ME’s sure amount is smaller than in OTHER’s sure amount. Similarly, in the time domain, $K - x^* < m^*$ indicates that ME is more impatient for delays in their own payments than in OTHER’s payments.

In addition to the questions Q1-Q4 we have described above, our experiments included one more question:

- Risk domain

$$(Q5) \quad v(K - s^*, 0) = p \cdot w(qa(s^* + c), 0) + (1 - p) \cdot w(0, 0)$$

- Time domain

$$(Q5) \quad v(K - s^*, 0) = w(qa(s^* + c), t)$$

In the risk domain, Q5 elicits the amount s^* that makes ME indifferent between receiving $K - s^*$ for sure and OTHER receiving the lottery that pays $a(s^* + c)$ with probability p . In the time domain, Q5 elicits the amount s^* that makes ME indifferent between receiving $K - s^*$ today and the OTHER person receiving $a(s^* + c)$ in t days. Q5 was primarily used for the structural estimation of preference parameters, as discussed in Section 4.4.

3.2 Experimental Protocol

The experiment consisted of 20 rounds and 20 MPLs, one per round. Each MPL consists of a sequence of “left” and “right” option. In each row of the MPL, the participant has to choose between the left and the right options. These are structured so as to find a point of indifference between the left and right options. At the end of the experiment we included several control tasks. We summarize the main features of the experiment. The Online Appendix features the instructions that were given to our participants, as well as screenshots from the experimental interface.

Valuation tasks. MPLs were used to elicit the indifference points $(x^*, y^*, z^*, m^*, s^*)$. Each round featured one MPL, which varied across two dimensions: (i) whether the left option remained fixed (a standard MPL) or varied with the right option (a modified MPL), and (ii) the identity of the recipient of the left and right outcomes (ME –ME, OTHER –ME, ME –OTHER, OTHER –OTHER).

A **standard MPL** consists of a sequence of rows, each displaying two options (left and right). The left option remains fixed across rows, while the right option changes, becoming progressively more attractive as one moves down the list. An illustration is provided in the left panel of Figure 1, which shows question Q2 in the risk domain. Subjects were asked to indicate the point at which they switched from preferring one option to the other, thereby enforcing a single switching point consistent with our theoretical framework, in which individuals reveal a unique indifference point.² To implement this, clicking the left option in a given row automatically selected the left option for that row and all rows above it, and the right option for all rows below; clicking the right option worked analogously. Participants could revise their selection at any time before submitting their final choice for each list.

A **modified MPL** also consists of a sequence of rows with two options in each; but in this case both the left and right options vary across rows: the left option becomes progressively less attractive, while the right option becomes progressively more attractive. An illustration is provided in the right panel of Figure 1, which shows question Q1 in the time domain. As with the standard MPL, subjects were asked to indicate the point at which they switched from preferring one option to the OTHER, thereby enforcing a single switching point consistent with our theoretical framework.

The modified MPL is meant to mimic the design in [Andreoni and Miller \[2002\]](#). In their design, as participants evaluate two choices on a budget line, they are simultaneously considering an increase in ME’s consumption and a decrease in OTHER’s. This tradeoff is obtained in our modified MPL by making an option for ME more attractive at the same time as the outcome for OTHER becomes worse.

Parameters. The five questions described in Section 3.1 were administered under two parameterizations in each domain. For risk preferences, we used $p = 0.50$ and $p = 0.80$. For time preferences, we used $t = 7$ and $t = 14$ days. In consequence, we obtain

²This is a standard experimental technique used most recently in [Nielsen and Rehbeck \[2022\]](#) and [McGranaghan et al. \[2024\]](#).

FIGURE 1: Two types of multiple-price lists

standard MPL						modified MPL					
OTHER gets	\$37.00 with prob 50%	or	YOU get	\$0.00	for sure	YOU get	\$20.00 today	or	YOU get	\$18.00 in 7 days	
	\$0.00 with prob 50%					YOU get	\$19.50 today	or	YOU get	\$18.25 in 7 days	
OTHER gets	\$37.00 with prob 50%	or	YOU get	\$1.00	for sure	YOU get	\$19.00 today	or	YOU get	\$18.50 in 7 days	
	\$0.00 with prob 50%					YOU get	\$18.50 today	or	YOU get	\$18.75 in 7 days	
OTHER gets	\$37.00 with prob 50%	or	YOU get	\$2.00	for sure	YOU get	\$18.00 today	or	YOU get	\$19.00 in 7 days	
	\$0.00 with prob 50%					YOU get	\$17.50 today	or	YOU get	\$19.25 in 7 days	
OTHER gets	\$37.00 with prob 50%	or	YOU get	\$3.00	for sure	YOU get	\$17.00 today	or	YOU get	\$19.50 in 7 days	
	\$0.00 with prob 50%					YOU get	\$16.50 today	or	YOU get	\$19.75 in 7 days	
...	
OTHER gets	\$37.00 with prob 50%	or	YOU get	\$24.00	for sure	YOU get	\$2.00 today	or	YOU get	\$27.00 in 7 days	
	\$4.00 with prob 50%					YOU get	\$1.50 today	or	YOU get	\$27.25 in 7 days	
OTHER gets	\$37.00 with prob 50%	or	YOU get	\$25.00	for sure	YOU get	\$1.00 today	or	YOU get	\$27.50 in 7 days	
	\$5.00 with prob 50%					YOU get	\$0.50 today	or	YOU get	\$27.75 in 7 days	
OTHER gets	\$37.00 with prob 50%	or	YOU get	\$26.00	for sure	YOU get	\$0.00 today	or	YOU get	\$28.00 in 7 days	
	\$5.00 with prob 50%										

Notes: The left panel presents Q2 question for risk preferences elicitation with $p = 0.50$. The right panel presents question Q1 for time preferences with $t = 7$ days.

$5 \times 2 + 5 \times 2 = 20$ rounds. The remaining parameters (q, K, a, c) were set to $(\frac{3}{2}, 18, \frac{1}{2}, 28)$ and $(2, 18, \frac{1}{2}, 31)$ for risk questions with $p = 0.50$ and $p = 0.80$, respectively and to $(\frac{3}{2}, 20, \frac{1}{2}, 36)$ and $(2, 20, \frac{1}{2}, 30)$ for time questions with $t = 7$ and $t = 14$ days, respectively. In the risk questions, the steps between MPL rows were \$1, while in the time questions, it was \$0.50. These parameters were calibrated to capture the relevant preferences regions as established in the vast previous experimental literature that studies time and risk preferences and to minimize the number of dominated choices in each MPL. Table A.1 in the appendix summarizes all the questions that appeared in the experiment.³

Structure of the Experiment. The rounds were grouped into four blocks: block A included 4 modified MPL ME-ME questions (Q1); block B included 8 standard MPL OTHER –ME questions (Q2 and Q3); block C included 4 standard MPL OTHER –OTHER questions (Q4); and block D included 4 modified MPL ME –OTHER questions (Q5). To control for order effects, we used a quasi-randomized block order, subject to the constraint that block A always preceded blocks C and D, since the latter blocks relied on values elicited in block A. After every four questions, we implemented brain breaks to reduce fatigue. During the brain breaks, participants were asked to find the distinction between two pictures depicting animals.

³Our experimental design is static in the sense that all decisions are made at a single date. It would be interesting to study behavior in a dynamic choice setting, as in Halevy [2015].

Control tasks. At the end of the experiment, participants completed a demographic questionnaire stating their age, gender, education, income, number of children and siblings, and political orientation on the liberal-conservative spectrum. Participants also answered two additional questions: one on their financial risk-taking behavior (“What percentage of your retirement savings is invested in risky assets such as stocks or mutual funds?”) and one on prosocial behavior (“How often do you give money to charities?”).

Incentives. Participants received a \$4 completion payment. In addition, 20% of participants were randomly selected for a bonus payment: one row from one of their MPL decisions was randomly implemented for payment. OTHER recipients were recruited in separate sessions, and bonus payments for choices affecting them were paid accordingly. The study lasted on average 25 minutes. The average total earnings per participant were approximately \$9. This average includes payment for both the ME and the OTHER participant, depending on which question was chosen for payment, and the choice made by the participant in this question.

Sample and Recruitment. The experiment was programmed and administered via Prolific in March 2025, with recruitment restricted to U.S.-based adults. A total of 300 participants completed the study. We excluded 33 participants who made two or more dominated choices—indicative of inattentiveness—leaving a final analysis sample of 267 participants.⁴

Ethics and Pre-registration. The study was approved by the Caltech (IRB protocol IR21-1167) and Berkeley Institutional Review Boards (CPHS protocol 2023-05-16357), and was pre-registered on AsPredicted (#130714).

3.3 Discussion of Experimental Design

Two features of our methodology are worth discussing. First, while we describe the approach commonly used in the literature as a no-SIG method, we acknowledge that if the participant (ME) cares about the OTHER’s consumption, then this approach does in fact provide incentives: the participant is choosing between objects that will be

⁴In addition, the completion time for inattentive subjects is significantly shorter than for attentive subjects ($p < 0.01$). Inattentive subjects also tend to click only once on the first row of the MPL, allowing them to move quickly through the questions.

consumed by the OTHER. However, if the participant is purely selfish and derives no utility from the OTHER’s consumption, then no such incentives exist. By contrast, our SIG approach avoids this problem entirely. It features built-in incentives regardless of whether the participant values the OTHER’s consumption because objects consumed by the OTHER are always evaluated in terms of the participant’s own consumption. We view this as a key advantage of our method.

Second, a potential concern with our design is that participants’ answers in Q1 determine the payments used in Q2 and Q3. This raises the possibility that subjects might strategically manipulate their responses in Q1 to influence later choices, thereby threatening incentive compatibility. We believe that this concern is largely mitigated in our setting for several reasons. First, it is highly unlikely that participants could detect the link between Q1 and Q2/Q3; much less learn how to manipulate the interface to their advantage. Participants faced multiple parameterizations of the core questions in randomized order, and the indifference points x^* elicited in Q1 were transformed into amounts in Q2 and Q3 through algebraic manipulations with constants (q, a, c, K) , effectively obscuring the connection between the different questions. Second, as we show in the results section, the recovered distributions of risk and time preferences for ME are similar to what has been found in the literature. This suggests that the answer to Q1 were not manipulated. Third, under the standard “isolation” assumption—that participants treat each question as independent—incentive compatibility is preserved. (For empirical support of the isolation assumption, see [Kahneman and Tversky \[1979\]](#) and [Hey and Lee \[2005a,b\]](#).) For a detailed theoretical discussion of the assumptions underlying incentive compatibility in different payment mechanisms, see [Azrieli et al. \[2018\]](#). Most relevant to our design is their result on the random problem selection mechanism: paying subjects for one randomly selected question is incentive compatible provided preferences satisfy a (state-wise) monotonicity condition.⁵

Our paper is not the first to employ a dynamic design. For further discussion of gaming concerns, see two recent papers that implement alternative versions of dynamic designs: the dynamically optimized sequential experimentation (DOSE) method by [Chapman et al. \[2024\]](#) and the Bayesian rapid optimal adaptive design (BROAD) by [Ray et al. \[2012\]](#). Both papers provide evidence that gaming behavior

⁵If the choice objects are lotteries, then the relevant preferences are over compound lotteries defined by the payment mechanism.

is quite rare and unlikely to undermine preference elicitation.⁶

4 Results

This section is organized as follows. We begin with a broad overview of our main findings. The subsequent subsections present the analyses deriving these results. Section 4.1 describes the identification of selfish types. Section 4.2 presents a non-parametric analysis based on the SIG method, while Section 4.3 compares non-parametric results obtained using the SIG and no-SIG methods. Section 4.4 introduces a structural parametric model of preferences and estimates its parameters under both methods. Finally, Section 4.5 evaluates the relative performance of the two approaches through an out-of-sample prediction exercise using the estimates from Section 4.4.

Heterogeneity and non-parametric comparisons. The application of the SIG method reveals substantial individual heterogeneity in how decision-makers distinguish between their own preferences and those they hold for others. The population is divided into three groups: those who are more risk averse with respect to OTHER’s than over ME’s consumption constitute the largest group; those who are less risk averse are the smallest; while those who are equally risk averse comprise about 30% of the participants. Despite this heterogeneity, at the level of individual participants, the findings are consistent across tasks and lottery specifications. In the time domain, while the results are more symmetric than in the risk domain, there is a tendency toward greater impatience regarding the other’s delays compared to one’s own. Crucially, the SIG methodology successfully identifies selfish types, which constitute approximately 20% of participants in the risk domain and 12% in the time domain.

Methodological differences: SIG vs. no-SIG. Comparing the SIG results with the traditional no-SIG method exposes clear contradictions between the two approaches. The no-SIG method yields a distribution of types that differs significantly

⁶There exists an alternative method that can, in principle, eliminate gaming concerns. Applied to our setting, this would require eliciting Q2 and Q3 for all possible values of x^* from Q1. This would, however, require an infeasible number of questions. Johnson et al. [2021] suggest a workaround for this problem by defining the grand set of such questions, but administering only a subset in the experiment. For payment, one question is randomly drawn from the grand set. If the selected question was part of the administered subset, payment is based on the subject’s original response; if not, the question is posed as an additional one. We do not expect our results to change under this alternative procedure, and therefore leave this as a robustness check for future work.

from the SIG method, even when selfish types are excluded. Most notably, in the risk domain, the no-SIG method suggests that the majority of participants are more risk-averse for themselves than for others, directly contradicting the modal finding of the SIG method. Furthermore, the no-SIG approach drastically underestimates the fraction of participants who hold identical preferences for self and other ($< 13\%$ vs. $> 30\%$ in SIG). This discrepancy suggests that the no-SIG method’s reliance on valuations expressed entirely in the “other’s” domain fails to capture the actual underlying preference structure, likely due to the decoupling of the decision-maker’s own utility from the valuation task.

Interestingly, it is not the presence of selfish agents that explains the differences between the SIG and no-SIG findings. Even altruistic participants behave differently in the SIG than in the no-SIG treatments.

Structural Estimation and Predictive Power. Our structural estimation results, assuming Constant Relative Risk Aversion (CRRA) and exponential discounting, corroborates the ordinal findings of the SIG method. The estimated risk parameter for ME ($\theta_M = 0.62$) is significantly higher than for OTHER ($\theta_O = 0.55$), confirming that agents behave as if they are more risk-averse when evaluating the other’s consumption. Similarly, the estimated discount factor is higher for self ($\delta_M \approx 0.98$) than for the other ($\delta_O \approx 0.95$), indicating greater patience for one’s own delays. In contrast, the no-SIG structural estimates are noisy and fail to reject the null hypothesis that preferences are identical across domains.

Out of sample prediction. Finally, an out-of-sample prediction exercise using a holdout question (Q5) demonstrates that the SIG parameter estimates achieve a significantly lower root mean squared error (RMSE) than the no-SIG estimates, validating the superior predictive power and goodness of fit of the skin-in-the-game approach.

4.1 Identifying selfish types

An important aspect of our design is that it allows us to categorize some participants as selfish. In each domain, there are six questions in which participants evaluate objects received by others in terms of the objects received by themselves; these are the questions Q2, Q3, and Q4 with different parameterizations. We use responses

to these questions to identify *selfish types*: participants who consistently allocate all available resources to their own consumption and prefer even the smallest amount of their own consumption over any amount allocated to others. Such participants seem to derive utility exclusively from their own consumption and none from the consumption of others.

Specifically, we classify a participant as *selfish* in domain d if they choose an option with any positive potential payoff for themselves (even if probabilistic) over an option that provides a positive payoff to the OTHER in at least two out of six questions. In other words, such participants select options benefiting the OTHER only when their own payoff is strictly zero. Based on this criterion, 54 out of 267 participants (20%) are selfish in the risk domain, 31 out of 267 (12%) are selfish in the time domain, and 57 out of 267 (21%) are selfish in at least one domain. In Appendix B, we show that although the proportion of selfish types naturally depends on the chosen cutoff, our qualitative results remain unchanged under both more stringent and more lenient definitions of selfishness.⁷

4.2 SIG, non-parametric analysis

Table 1 compares own and OTHER’s risk and time preferences using the ordering of y^* and z^* (Section 3.1). The data exhibit substantial heterogeneity: some participants are more risk-averse or more impatient for ME than for OTHER, some are less so, and many display similar attitudes for ME and OTHER. The distribution of comparative attitudes is stable across parameterizations in each domain. In the risk domain, type classifications for ME vs. OTHER are positively and significantly correlated across parameterizations ($p < 0.01$ for all three comparisons). Similarly, in the time domain, classifications are positively and significantly correlated ($p < 0.05$ for all three comparisons).

Our conclusions hold for the strict definition used in Table 1 and the relaxed definition reported in Table B.3 (Appendix B).⁸

In the risk domain, more participants are *less* risk-averse for ME than for OTHER

⁷Table B.1 presents the distribution of selfish choices in different number of questions. Table B.2 replicates the main ordinal classification presented in Table 1 in Section 4.2.

⁸The relaxed definition allows small perturbations. In the risk domain, we classify ME as more risk-averse than OTHER if $y^* < z^* - \$1$ (the step size in the MPL for risk questions). In the time domain, we classify ME as more impatient than OTHER if $y^* < z^* - \$0.50$ (the step size in the MPL for time questions).

TABLE 1: Preferences for ME vs OTHER, ordinal skin-in-the-game approach

ME -- than OTHER	Risk domain		ME -- than OTHER	Time domain	
	$p = 0.50$	$p = 0.80$		$t = 7$ days	$t = 14$ days
more risk-averse	0.24	0.28	more impatient	0.33	0.28
equally risk-averse	0.31	0.33	equally impatient	0.42	0.43
less risk-averse	0.45	0.39	less impatient	0.24	0.29
nb subjects	$i = 211$	$i = 208$		$i = 227$	$i = 225$

Notes: We focus on non-selfish participants in each domain with non-dominant choice in Q1 in each parameterization. The comparison of attitudes is done using questions Q2 and Q3: if $y^* < z^*$ then ME is more risk-averse (more impatient) than OTHER; see Section 3.1 for details.

than have the opposite preferences. So the share of those who are *less* risk-averse for ME than for OTHER exceeds the share of those who are *more* risk-averse for ME than for OTHER (Test of Proportions: $p < 0.05$ in both parameterizations). By contrast, in the time domain, we find no consistent asymmetry across parameterizations: most participants display similar time preferences for ME and OTHER.⁹

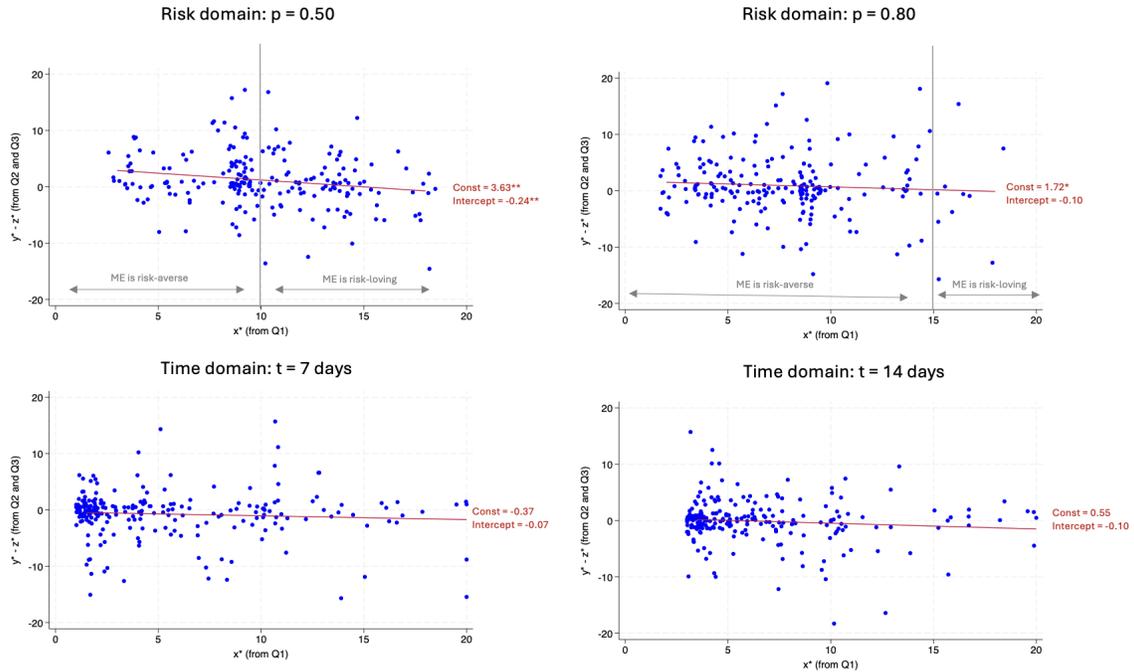
Figure 2 shows the relationship between ME’s preferences for objects consumed by OTHER and those consumed by ME, reported separately for each domain. First, we note that for the objects consumed by ME, we reproduce common findings regarding risk and time preferences. In the risk domain, most subjects are risk-averse (75% for $p = 0.50$ and 77% for $p = 0.80$), with a median certainty equivalent of \$9 for a lottery paying \$18.50 with probability 0.50 and \$10 for a lottery paying \$19.50 with probability 0.80. In the time domain, the median subject is indifferent between \$19.75 in 7 days and \$16.50 today (implying a daily interest rate of 2.81%), and between \$17.50 in 14 days and \$15 today (implying a daily interest rate of 2.38%). Figure B.1 in Appendix B presents the distribution of x^* in each parameterization.

The novelty in our design is the comparison between risk and time attitudes for ME and OTHER. Our data exhibit substantial heterogeneity at the participant level. There is no systematic relationship across participants between own risk preferences and risk preferences for the consumption of OTHER. The data points in Figure 2 are nearly symmetric around zero, indicating that the certainty equivalent of ME consuming a given lottery does not predict ME’s certainty equivalent of OTHER consuming the same lottery (expressed in ME-payments).¹⁰ An analogous pattern holds

⁹When $t = 7$ days we observe more participants who are more impatient for ME than for OTHER ($p < 0.05$); when $t = 14$ days the difference is not significant ($p = 0.92$).

¹⁰For $p = 0.50$, both the intercept and slope of the linear fit are statistically different from zero but their magnitudes are negligible.

FIGURE 2: Connection between preferences, ordinal skin-in-the-game approach



Notes: For each parameterization and each domain, we plot the difference between indifference points y^* and z^* versus indifference point x^* . For the risk domain, we split the region of x^* into the risk-loving and risk-averse/neutral regions. The red line is the linear fit and the estimated intercepts and slopes are recorded next to the line. ** indicates significance at 5% level.

for time preferences: an individual's patience for consumption bundles consumed by ME is unrelated to their patience for bundles consumed by OTHER.

Finally, Table B.4 in the Appendix B presents the correlation between the additional control tasks elicited at the end of the experiment and the risk- and time-attitudes we recover for participants for ME vs OTHER. We find no robust evidence that any of these covariates are related to the differential risk- or time-preferences that people exhibit for their own consumption versus the consumption of others.

4.3 Comparing SIG and no-SIG, non-parametric analysis

We next compare the SIG and the no-SIG methods. Table 2 reports the classification of types under each method, separately for each domain. Recall that the no-SIG method cannot identify purely selfish types, since the objects consumed by the OTHER are evaluated only in terms of OTHER-payments. This explains the difference in the

number of observations reported in the first two columns of each parameterization. The third column shows the classification based on the no-SIG method after removing the selfish types identified by the SIG method.

The distributions of types recovered by the two methods differ substantially. Consider first the risk domain. Under the no-SIG approach, similar risk attitudes of ME and OTHER are quite rare—less than 13% for $p = 0.50$ and less than 8% for $p = 0.80$. In contrast, the SIG approach finds that more than 30% of participants exhibit similar risk attitudes for objects consumed by ME and by OTHER.¹¹ Moreover, while the no-SIG approach suggests that the majority of participants are more risk-averse for ME than for OTHER, the SIG approach identifies the opposite pattern as modal behavior.¹² Interestingly, removing the selfish types and reapplying the no-SIG approach leaves the distribution of types virtually unchanged. This occurs because in the no-SIG case, the majority of purely selfish types are misclassified as being more risk-averse for ME than OTHER when asked to state the certainty equivalent of the lottery consumed by OTHER in OTHER-terms. Indeed, the no-SIG method recovers a similar distribution of risk attitudes for ME and OTHER when applied only to selfish types (identified by the SIG approach), all non-selfish types, or all subjects together.¹³

In the time domain, the patterns are similar. The fraction of participants who display similar time preferences for ME and OTHER is significantly smaller under the no-SIG than under the SIG approach.¹⁴ Moreover, while the SIG approach recovers roughly equal fractions of participants who are more rather than less impatient for ME than for OTHER, the no-SIG approach produces a skewed distribution, with substantially more participants classified as more impatient for ME than for OTHER.¹⁵ Furthermore, as in the risk domain, removing the selfish types and reapplying the no-

¹¹The differences in the fractions of participants who have similar risk-attitudes in the SIG and no-SIG approaches are statistically significant, $p < 0.01$.

¹²Under the no-SIG approach, the fraction of subjects classified as more risk-averse for ME than for OTHER is significantly larger than the fraction classified as less risk-averse: $p < 0.01$ in both parameterizations.

¹³The no-SIG approach applied to selfish participants yields the following type distributions: for $p = 0.50$, 50% are more risk-averse, 17% equally risk-averse, and 33% are less risk-averse for ME than OTHER; for $p = 0.80$, 57% are more risk-averse, 6% equally risk-averse, and 37% are less risk-averse for ME than OTHER.

¹⁴We obtain $p < 0.01$ in both parameterizations.

¹⁵In the SIG approach, the fraction of participants classified as more impatient for ME than for OTHER exceeds the fraction classified as less impatient at $t = 7$ days ($p = 0.03$), but not at $t = 14$ days ($p = 0.92$). In contrast, under the no-SIG approach, the fraction classified as more impatient is significantly higher than the fraction classified as less impatient in both parameterizations ($p = 0.02$ for $t = 7$ days; $p < 0.01$ for $t = 14$ days).

TABLE 2: Comparing types recovered by SIG and no-SIG methods

	Risk domain: $p = 0.50$			Risk domain: $p = 0.80$		
	SIG	no-SIG	no-SIG non-selfish	SIG	no-SIG	no-SIG non-selfish
ME more risk-averse than OTHER	0.24	0.60	0.63	0.28	0.56	0.56
ME equally risk-averse as OTHER	0.31	0.13	0.12	0.33	0.08	0.08
ME less risk-averse than OTHER	0.45	0.27	0.25	0.39	0.36	0.36
nb subjects	$i = 211$	$i = 265$	$i = 211$	$i = 208$	$i = 262$	$i = 208$
	Time domain: $t = 7$ days			Time domain: $t = 14$ days		
	SIG	no-SIG	no-SIG non-selfish	SIG	no-SIG	no-SIG non-selfish
ME more impatient than OTHER	0.33	0.44	0.45	0.28	0.48	0.48
ME equally impatient as OTHER	0.42	0.22	0.20	0.43	0.26	0.24
ME less impatient than OTHER	0.24	0.34	0.34	0.29	0.26	0.28
nb subjects	$i = 227$	$i = 257$	$i = 227$	$i = 225$	$i = 254$	$i = 225$

Notes: SIG stands for skin-in-the game approach, while no-SIG stands for no-skin-in-the-game approach.

SIG approach does not change the distribution of recovered types.¹⁶

The similarity between the no-SIG distributions with and without selfish types, combined with the stark differences from the distributions recovered by the SIG approach, suggests that the discrepancies between the two approaches are not merely due to the no-SIG approach’s inability to identify selfish types. Rather, they reflect a more fundamental divergence in how individuals evaluate objects in terms of their own versus others’ consumption.

4.4 SIG, structural parametric estimation

We now turn to a structural estimation of a parametric model of preferences. In particular, we assume that participants’ utilities over their own consumption, as well as the utility over OTHER’s consumption, takes the form of exponential discounting with constant relative risk aversion (CRRA). Under these parametric assumptions, we structurally estimate the parameters governing risk and time preferences: the coefficient of relative risk aversion and the discount factor.¹⁷ We use non-linear least squares to jointly estimate four parameters: risk for ME (θ_M), risk for OTHER (θ_O),

¹⁶The no-SIG approach applied to selfish participants yields the following type distributions: for $t = 7$ days, 33% are more impatient, 37% equally impatient, and 30% less impatient for ME than OTHER; for $t = 14$ days, 41% are more impatient, 41% equally impatient, and 17% less impatient.

¹⁷MARINA: do the robustness test same exercise with CARA and put in Appendix.

time for ME (δ_M), and time for OTHER (δ_O). This estimation is performed both at the aggregate level (Table 3) and at the individual level (Figure 3).

The results of the aggregate estimation are reported in Table 3. Regression (1) reports the estimated parameters for non-selfish participants in both domains using the three core questions of SIG method. For comparison, regression (2) reports estimates based on the two questions required to recover risk and time parameters under the no-SIG method.

TABLE 3: Estimating risk and time preferences parameters for ME and OTHER

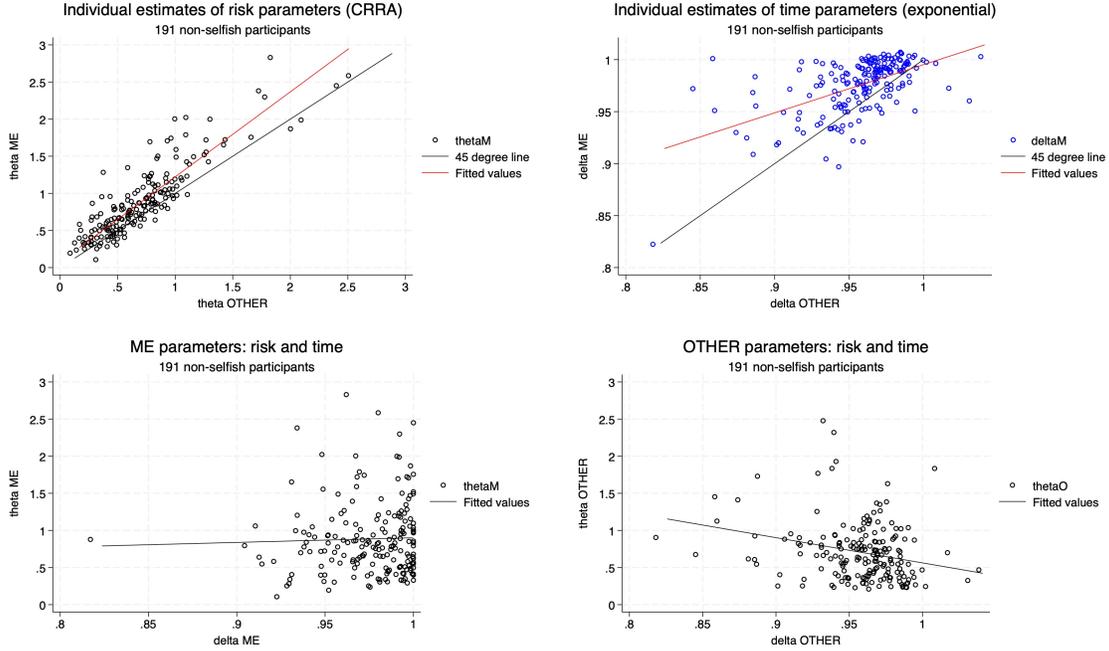
	SIG method Reg (1)	no-SIG method Reg (2)
risk parameter for ME θ_M	0.62 (0.02)	0.63 (0.02)
risk parameter for OTHER θ_O	0.55 (0.02)	0.67 (0.04)
time parameter for ME δ_M	0.98 (0.002)	0.98 (0.002)
time parameter for OTHER δ_O	0.95 (0.002)	0.98 (0.002)
H0: $\theta_M = \theta_O$	$p < 0.01$	$p = 0.270$
H0: $\delta_M = \delta_O$	$p < 0.01$	$p = 0.127$
sample	Q1, Q2, Q3	Q1, Q4
nb obs	$n = 2412$	$n = 2043$
nb subjects	$i = 210$	$i = 266$

Notes: Nonlinear least-squares estimation is performed assuming CRRA utility and exponential discounting, i.e., $v(x, t) = \delta_M^t \cdot x^{\theta_M}$ and $w(x, t) = \delta_O^t \cdot x^{\theta_O}$. Robust standard errors are reported in the parenthesis with clustering at the participant level.

Table 3 conveys several interesting findings. First, both the SIG and no-SIG method recover standard preferences for both ME and OTHER: on average, ME is risk-averse and impatient for both ME and OTHER, and the parameter estimates are consistent with the range of empirical estimates in the literature. Second, according to the SIG method, ME is more risk-averse when evaluating lotteries consumed by OTHER than by ME, and more impatient when delays are incurred by OTHER than by ME. The estimated risk and time parameters for ME and OTHER differ significantly at the 1% level. These aggregate estimations align closely with the non-parametric individual-level results presented in the previous section. The only difference between the two is the following: while the non-structural approach shows roughly equal proportions of participants who are more or less impatient for ME than for OTHER, the structural estimation detects a slight asymmetry in time attitudes. This difference, however, is quantitatively very small.

Third, the no-SIG estimates show that ME is slightly less risk-averse for OTHER

FIGURE 3: Individual Structural Estimates of Risk and Time Parameters



Notes: Individual estimates of risk and time parameters are plotted for non-selfish participants. We exclude 19 participants for whom we cannot reliably recover their individual estimates. The estimation is based on Q1, Q2, and Q3 in each parameterization in both domains for a maximum of 12 questions per person. The top left (right) panel depicts the relationship between risk (time) preferences for ME (OTHER). The bottom left (right) panel depicts the relationship between own (other's) parameters in the risk and time domain.

than for ME and is equally impatient for both. However, the estimates are quite noisy: in fact, in both domains the estimated parameters for ME and OTHER are not statistically different ($p > 0.10$). These results are inconsistent with the non-structural no-SIG results reported in the previous section. Specifically, the non-structural no-SIG approach shows that more participants are more risk-averse and more impatient for ME than for OTHER (Table 2) contradicting the structural estimates.

The results of the individual-level estimation are reported in Figure 3, which graphs our individual-level parameter estimates for the risk and time-preference parameters. The findings in Figure 3 largely corroborate the message of our estimation with aggregate data. We focus on the 191 non-selfish participants who chose at least five interior certainty equivalents out of the 12 questions used for this estimation. The top two panels compare risk (time) preferences for ME and OTHER, while the bottom two panels plot both domain parameters separately for ME and OTHER.

The individual-level SIG estimates closely mirror the aggregate results discussed above. A majority of participants (69%) are risk-averse for both ME and OTHER, while 14% are risk-loving for both. Importantly, majority of participants (81%) are more risk-averse for OTHER than for ME, and roughly 10% display identical risk attitudes across the two. Turning to time preferences, nearly all participants (97%) are impatient for both ME and OTHER. Among them, a majority (53%) exhibit the same degree of impatience, while 41% are more impatient for OTHER than for ME.¹⁸

4.5 Predictive ability of SIG and no-SIG methods

In our final exercise, we assess the relative performance of the two methods by conducting an out-of-sample prediction exercise based on the estimated parameters reported in Table 3. Standard goodness-of-fit measures such as R^2 , root MSE, Akaike Information Criterion, or Bayesian Information Criterion are not appropriate in this case, since the SIG and no-SIG methods are estimated on different samples with different numbers of observations. So we use an out of sample prediction exercise to assess both goodness of fit and predictive ability. Specifically, we evaluate how well the parameter estimates obtained under each method predict behavior in Q5, which was not used in the structural estimations. The results indicate that the SIG method achieves a significantly lower root MSE than the no-SIG method (Wilcoxon sign-rank test: $p < 0.01$).

We conclude that the structural SIG approach outperforms the no-SIG approach in its out-of-sample predictive ability. Moreover, it produces estimates that are consistent across both the aggregate and individual levels and, importantly, align closely with the non-structural SIG results discussed in the previous section. In contrast, the

¹⁸To make these individual comparisons we allowed for some slack in the estimates. In particular, we say that ME is more risk-averse for OTHER than for ME if $\theta_M > \theta_O + 0.02$; similarly, we say that ME is more impatient for OTHER than for ME if $\delta_M > \delta_O + 0.02$.

no-SIG approach yields noisier estimates that lack internal consistency. We view this coherence as an additional strength of the SIG method.

A Parameters

Table A.1 summarizes the parameters used in each question of the experiment. The numbers shown in purple vary depending on the switching row in Question 1 (Q1) for each domain and parameterization:

- For risk questions with $p = 0.50$, we report the Q2–Q4 values corresponding to a decision-maker who is indifferent between receiving \$9 for sure and a lottery paying \$18.50 with probability 0.50 and \$0 with probability 0.50.
- For risk questions with $p = 0.80$, we report the Q2–Q4 values corresponding to a decision-maker who is indifferent between receiving \$13 for sure and a lottery paying \$18 with probability 0.80 and \$0 with probability 0.20.
- For time questions with a 7-day delay ($t = 7$), we present the Q2–Q4 values corresponding to indifference between receiving \$18.50 today and \$18.75 in 7 days.
- For time questions with a 14-day delay ($t = 14$), we present the Q2–Q4 values corresponding to indifference between receiving \$16.50 today and \$16.75 in 14 days.

TABLE A.1: Parameters of experiment

	Left option		Right option		Changes across rows			nb rows		
	recipient	object	recipient	object	x_{\min}	x_{\max}	x_{step}			
RISK										
Q1	$p = 0.50$	ME	\$(18-x)\$	ME	\$(14+0.5x)\$	prob p	0	18	1	19 rows
Q2	$p = 0.50$	OTHER	\$18	ME	\$x		0	26	1	27 rows
Q3	$p = 0.50$	OTHER	\$37 prob p	ME	\$x		0	26	1	27 rows
Q4	$p = 0.50$	OTHER	\$18.50 prob p	OTHER	\$x		0	26	1	27 rows
Q5	$p = 0.50$	ME	\$(18-x)\$	OTHER	\$(28+x)\$	prob p	0	18	1	19 rows
Q1	$p = 0.80$	ME	\$(18-x)\$	ME	\$(15.50+0.5x)\$	prob p	0	18	1	19 rows
Q2	$p = 0.80$	OTHER	\$19.50	ME	\$x		0	26	1	27 rows
Q3	$p = 0.80$	OTHER	\$27 prob p	ME	\$x		0	26	1	27 rows
Q4	$p = 0.80$	OTHER	\$18 prob p	OTHER	\$x		0	26	1	27 rows
Q5	$p = 0.80$	ME	\$(18-x)\$	OTHER	\$(23.25+0.75x)\$	prob p	0	18	1	19 rows
Q1	$t = 7$ days	ME	\$(20-x)\$ today	ME	\$(18+0.5x)\$	in t	0	20	0.50	41 rows
Q2	$t = 7$ days	OTHER	\$27.75 today	ME	\$x	today	0	20	0.50	41 rows
Q3	$t = 7$ days	OTHER	\$28.13 in t	ME	\$x	today	0	20	0.50	41 rows
Q4	$t = 7$ days	OTHER	\$18.75 in t	OTHER	\$x	today	0	20	0.50	41 rows
Q5	$t = 7$ days	ME	\$(20-x)\$ today	OTHER	\$(27+0.75x)\$	in t	0	20	0.50	41 rows
Q1	$t = 14$ days	ME	\$(20-x)\$ today	ME	\$(15+0.5x)\$	in t	0	20	0.50	41 rows
Q2	$t = 14$ days	OTHER	\$33 today	ME	\$x	today	0	20	0.50	41 rows
Q3	$t = 14$ days	OTHER	\$33.50 in t	ME	\$x	today	0	20	0.50	41 rows
Q4	$t = 14$ days	OTHER	\$16.75 in t	OTHER	\$x	today	0	20	0.50	41 rows
Q5	$t = 14$ days	ME	\$(20-x)\$ today	OTHER	\$(30+x)\$	in t	0	20	0.50	41 rows

B Additional Analysis

TABLE B.1: Proportions of subjects behaving selfishly in different number of questions

	Risk domain	Time domain
never	0.73	0.84
1 question	0.07	0.05
2 questions	0.03	0.02
3 questions	0.02	0.005
4 questions	0.04	0.04
5 questions	0.02	0.01
6 questions	0.09	0.05
nb subjects	267	267

Notes: Selfish behavior means choosing options that allocate all available resources to own rather than other’s consumption. We consider questions Q2, Q3, and Q4 in each parameterization and in each domain for a total of six questions per domain.

TABLE B.2: Risk and time preferences for ME vs OTHER under different definitions of selfishness, ordinal skin-in-the-game approach

ME -- than OTHER	Main definition		Definition 1		Definition 2	
	$p = 0.50$	$p = 0.80$	$p = 0.50$	$p = 0.80$	$p = 0.50$	$p = 0.80$
more risk-averse	0.24	0.28	0.24	0.28	0.26	0.29
equally risk-averse	0.31	0.33	0.33	0.35	0.31	0.32
less risk-averse	0.45	0.39	0.42	0.37	0.44	0.39
nb subjects	$i = 211$	$i = 208$	$i = 192$	$i = 191$	$i = 226$	$i = 223$

ME -- than OTHER	Main definition		Definition 1		Definition 2	
	$t = 7$ days	$t = 14$ days	$t = 7$ days	$t = 14$ days	$t = 7$ days	$t = 14$ days
more impatient	0.33	0.28	0.33	0.27	0.33	0.28
equally impatient	0.42	0.43	0.43	0.44	0.42	0.43
less impatient	0.24	0.29	0.24	0.29	0.25	0.29
nb subjects	$i = 227$	$i = 225$	$i = 215$	$i = 214$	$i = 232$	$i = 229$

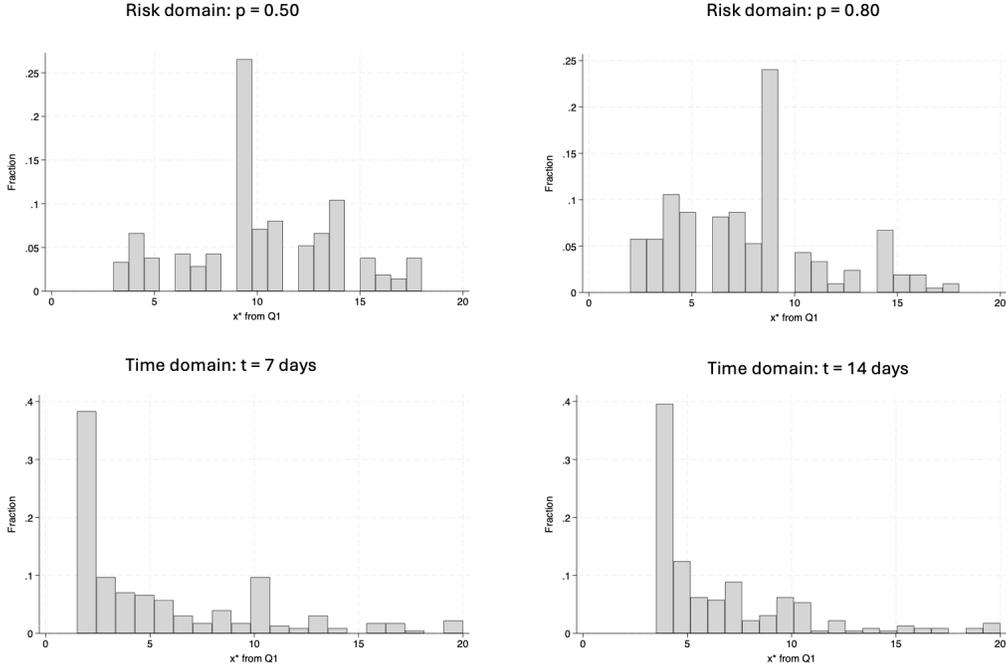
Notes: We focus on non-selfish participants in each domain with non-dominant choice in Q1 in each parameterization. The main definition uses the definition of selfishness defined in Section XX in the paper, according to which participant is classified as non-selfish if they choose options that allocate all available resources to own rather than other’s consumption in less than 2 out of 6 questions. Definition 1 is more stringent: non-selfish participants never allocate all available resources to only own consumption. Definition 2 is more lenient: non-selfish participant allocate all resources to own consumption in at most 3 out of 6 questions. The comparison of attitudes is done using questions Q2 and Q3: if $y^* < z^*$ then ME is more risk-averse (impatient) than OTHER; see Section 3.1 for details.

TABLE B.3: Preferences for ME vs OTHER, the ‘relaxed’ definition, ordinal skin-in-the-game approach

ME -- than OTHER	Risk domain		ME -- than OTHER	Time domain	
	$p = 0.50$	$p = 0.80$		$t = 7$ days	$t = 14$ days
more risk-averse	0.20	0.22	more impatient	0.29	0.22
equally risk-averse	0.43	0.44	equally impatient	0.52	0.53
less risk-averse	0.40	0.34	less impatient	0.19	0.25
nb subjects	$i = 211$	$i = 208$		$i = 227$	$i = 225$

Notes: We focus on non-selfish participants in each domain with non-dominant choice in Q1 in each parameterization. The comparison of attitudes is done using questions Q2 and Q3 allowing for small perturbations. If $y^* < z^* - \$1$ then ME is more risk-averse than OTHER, while if $y^* > z^* + \$1$ then ME is less risk-averse than OTHER. Similarly, if $y^* < z^* - \$0.50$ then ME is more impatient than OTHER, while if $y^* > z^* + \$0.50$ then ME is less impatient than OTHER.

FIGURE B.1: Risk and time preferences for ME



Notes: We plot the histogram of x^* elicited in Q1 of the corresponding domain/parameterization focusing on non-dominated choices of non-selfish participants.

TABLE B.4: Effect of Covariates on Risk and Time Preferences for ME vs OTHER

	Risk domain					
	more risk-averse	$p = 0.50$ same	less risk-averse	more risk-averse	$p = 0.80$ same	less risk-averse
Male	-0.04 (0.06)	0.25** (0.06)	-0.20** (0.06)	0.01 (0.06)	0.04 (0.07)	-0.04 (0.07)
Age (over 35)	-0.002 (0.06)	0.08 (0.06)	-0.08 (0.07)	-0.01 (0.06)	0.05 (0.07)	-0.04 (0.07)
College Degree (completed)	-0.14** (0.06)	0.004 (0.07)	0.14** (0.07)	-0.02 (0.07)	-0.09 (0.07)	0.12* (0.07)
No children	-0.07 (0.06)	-0.13** (0.07)	0.20** (0.07)	-0.003 (0.07)	-0.05 (0.07)	0.05 (0.07)
No siblings	0.16 (0.12)	-0.14 (0.14)	-0.02 (0.15)	0.04 (0.15)	-0.05 (0.07)	0.14 (0.15)
Liberal	0.10* (0.06)	-0.09 (0.06)	-0.007 (0.07)	0.06 (0.06)	0.06 (0.06)	-0.12* (0.07)
Charity (almost never)	-0.04 (0.10)	0.16* (0.09)	-0.12 (0.12)	-0.08 (0.11)	0.04 (0.11)	0.05 (0.11)
Retirement (not risky)	0.01 (0.06)	0.08 (0.06)	-0.09 (0.07)	0.05 (0.06)	0.03 (0.07)	-0.08 (0.07)
nb of participants	$i = 213$	$i = 213$	$i = 213$	$i = 213$	$i = 213$	$i = 213$

	Time domain					
	more impatient	$t = 7$ days same	less impatient	more impatient	$t = 14$ days same	less impatient
Male	0.001 (0.06)	0.09 (0.06)	-0.09* (0.06)	0.03 (0.06)	-0.001 (0.06)	-0.02 (0.06)
Age (over 35)	-0.14** (0.06)	0.21** (0.06)	-0.07 (0.05)	-0.12** (0.06)	0.15** (0.06)	-0.03 (0.06)
College Degree (completed)	-0.02 (0.07)	-0.10 (0.07)	0.12* (0.06)	0.004 (0.06)	-0.06 (0.07)	0.06 (0.07)
No children	-0.02 (0.07)	0.01 (0.07)	0.01 (0.06)	-0.01 (0.06)	0.14** (0.07)	-0.12* (0.06)
No siblings	0.05 (0.12)	-0.04 (0.13)	-0.008 (0.11)	0.19* (0.10)	-0.20 (0.14)	0.003 (0.13)
Liberal	0.11* (0.06)	-0.13** (0.06)	0.03 (0.05)	-0.007 (0.06)	-0.09 (0.06)	0.10* (0.06)
Charity (almost never)	0.06 (0.10)	-0.21* (0.10)	0.14* (0.09)	-0.006 (0.09)	-0.08 (0.11)	0.09 (0.10)
Retirement (not risky)	0.09 (0.06)	-0.03 (0.06)	-0.06 (0.06)	0.09 (0.06)	-0.10 (0.06)	0.01 (0.06)
nb of participants	$i = 236$	$i = 236$	$i = 236$	$i = 236$	$i = 236$	$i = 236$

Notes: We estimate Multinomial Logit regression with dependent variable recording whether the subjects is more, same, or less risk-averse for ME than OTHER in the top panel and more, same, less impatient impatient for ME than for OTHER in the bottom panel for each parameterization separately. We then report the average marginal effect (AME) of each covariate on the probability that the outcome takes each of the three values. In both domains we focus on non-selfish participants as defined in our results section.

References

- Andreoni, James and John Miller (2002), “Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism.” *Econometrica*, 70, 737–753.
- Azrieli, Yaron, Christopher Chambers, and Paul J. Healy (2018), “Incentives in Experiments: A Theoretical Analysis.” *Journal of Political Economy*, 126.
- Chakraborty, Anujit and Luca Henkel (2025), “The Role of Interpersonal Uncertainty in Prosocial Behavior.” Mimeo, UC Davis.
- Chakravarty, Sujoy, Glenn W Harrison, Ernan E Haruvy, and E Elisabet Rutström (2011), “Are you risk averse over other people’s money?” *Southern Economic Journal*, 77, 901–913.
- Chapman, Jonathan, Erik Snowberg, Stephanie Wang, and Colin Camerer (2024), “Dynamically Optimized Sequential Experimentation (DOSE) For Estimating Economic Preference Parameters.” *working paper*.
- Exley, Christine L (2016), “Excusing selfishness in charitable giving: The role of risk.” *The Review of Economic Studies*, 83, 587–628.
- Fisman, Ray, Shachar Kariv, and Daniel Markovits (2007), “Individual Preferences for Giving.” *American Economic Review*, 97, 1858–1876.
- Halevy, Yoram (2015), “Time Consistency: Stationarity and Time Invariance.” *Econometrica*, 83, 335–352. [10.3982/ECTA10872](https://doi.org/10.3982/ECTA10872).
- Harrison, Glenn W., Morten I. Lau, E. Elisabet Rutström, and Marcela Tarazona-Gómez (2012), “Preferences over social risk.” *Oxford Economic Papers*, 65, 25–46. [10.1093/oep/gps021](https://doi.org/10.1093/oep/gps021).
- Hey, J.D. and J. Lee (2005a), “Do Subjects Remember the Past?” *Applied Economics*, 37, 9–18.
- Hey, J.D. and J. Lee (2005b), “Do Subjects Separate (or Are They Sophisticated)?” *Experimental Economics*, 8, 233–265.

- Johnson, Cathleen, Aurelien Baillon, Han Bleichrodt, Zihua Li, Dennie VanDolder, and Peter Wakker (2021), “PRINCE: an improved method for measuring incentivized preferences.” *Journal of Risk and Uncertainty*, 62, 1—28.
- Kahneman, D and A. Tversky (1979), “Prospect Theory: An Analysis of Decision-sunder Risk.” *Econometrica*, 47, 263–291.
- McGranaghan, Christina, Kirby Nielsen, Ted O’Donoghue, Jason Somerville, and Charles Sprenger (2024), “Distinguishing Common Ratio Preferences from Common Ratio Effects Using Paired Valuation Tasks.” *American Economic Review*, 114, 307–347.
- Nielsen, Kirby and John Rehbeck (2022), “When Choices Are Mistakes.” *American Economic Review*, 112, 2237—2268.
- Ray, Debajyoti, Daniel Golovin, Andreas Krause, and Colin Camerer (2012), “Bayesian Rapid Optimal Adaptive Design (BROAD): Method and application distinguishing models of risky choice.” *working paper*.
- Reynolds, Douglas B, Jacob Joseph, Reuben Sherwood et al. (2009), “Risky shift versus cautious shift: determining differences in risk taking between private and public management decision-making.” *Journal of business & economics research (JBER)*, 7.