Ordered Reference Dependent Choice

(RC) Lim, Xi Zhi*

Revision 11/11/2019

Abstract

We study how violations of structural assumptions like expected utility and exponential discounting can be connected to reference dependent preferences with set-dependent reference points, even if behavior conforms with these assumptions when the reference is fixed. This is done with the introduction of a unified framework under which both general rationality (WARP) and domain-specific structural postulates (e.g., Independence for risk preference, Stationarity for time preference) are jointly relaxed using a systematic reference dependence approach. The framework allows us to study risk, time, and social preferences collectively, where behavioral departures from WARP and structural postulates are explained by a common source—changing preferences due to reference dependence. In our setting, reference points are given by a linear order that captures the relevance of each alternative in becoming the reference point and affecting preferences. In turn, they determine the domain-specific preference parameters for the underlying choice problem (e.g., utility functions for risk, discount factors for time).

1 Introduction

The standard model of choice in economics faces two separate strands of empirical challenges. First, structural assumptions such as the *expected utility form* (Independence) and *exponential discounting* (Stationary) are violated in simple choice experiments, most notably the Allais paradox and present bias. Second, and separately, studies have shown that choices are affected by reference points, resulting in "*non-rational*" behavior that violates the weak axiom of revealed preferences (WARP). With

*Department of Economics, Columbia University. Email: xi.zhi.lim@columbia.edu. I thank Pietro Ortoleva, Mark Dean, and Navin Kartik for excellent advising. I also thank participants at various seminars for useful feedback, and in particular Collin Raymond, Yusufcan Masatlioglu, Gil Riella, Kfir Eliaz, Elliot Lipnowski, Qingmin Liu, Yeon-Koo Che, and Weijie Zhong.

few exceptions, these two classes of prominent departures from standard models have been studied separately, and independently for each domain of choice, propelling models that seek to explain one phenomenon in isolation of the others.[1]

In this paper, we propose a unified framework in which failures of WARP and violations of structural assumptions, across the risk, time, and social domains, are jointly explained by reference dependency. This allows us to study different types of documented departure from standard models as related to one another, and in doing so suggests new empirical directions.

The intuition comes from a simple observation: If decision makers have preferences (e.g., utility functions, discount factors) that depend on a reference point, then even if they are otherwise standard and maximize exponentially discounted expected utility, they would still violate both WARP and structural assumptions like the Independence and Stationarity axioms from time to time—when reference points change.

Working with choice behavior, we provide the axiomatic foundation for a set of four models—generic choice, risk preference, time preference, and social preference—in which behavioral anomalies are explained by a common source: changing preferences due to reference dependence. In these models, reference points are endogenously determined by *reference orders*, which rank each alternative by their relevance in becoming the reference point and affecting preferences.

To illustrate, consider a decision maker who contradicts expected utility theory by exhibiting increased risk aversion in the presence of safer options. This narrative is consistent with a myriad of anomalous choice documented in Herne (1999); Wakker & Deneffe (1996); Andreoni & Sprenger (2011), and prominently Allais (1953)'s paradox.[2] This behavior can be explained without fully rejecting the expected utility form—that decision makers maximize the expectation of some utility function for each choice problem—but by allowing for reference dependent utility functions. We propose, in the risk domain, that a decision maker's utility function depends on the safest available alternative, which reflects changing risk aversion. When the safest alternative is fixed, standard expected utility holds. But when reference point changes, then the safer the reference, the more concave the utility function.

[1]For reference dependence, see for example Kahneman & Tversky (1979), Kőszegi & Rabin (2006), Masatlioglu & Ok (2005), Masatlioglu & Ok (2013), Ok et al. (2015), and Dean et al. (2017). For models weakening the expected utility form see Quiggin (1982), Bell (1982); Loomes & Sugden (1982), Chew (1983); Fishburn (1983); Dekel (1986), Gul (1991), and Cerreia-Vioglio et al. (2015). For models weakening the discounted utility form see Loewenstein & Prelec (1992), Laibson (1997) and Frederick et al. (2002). For models that use reference dependency to explain violations of structural assumptions, see for example Kőszegi & Rabin (2007), Ortoleva (2010). An exception where both WARP and structural assumptions are relaxed is Bordalo et al. (2012).

[2]In the Allias paradox, a decision maker is drawn to the safe option when it is available, contradicting the irrelevance of common consequence assumption in standard expected utility theory. We will (re)introduce the Allais paradox and discuss the application of our model in Section 3. Herne (1999); Wakker & Deneffe (1996); Andreoni & Sprenger (2011) document other behaviors consistent with our risk model, discussed in Section 3.

We then show that the same concept can be applied to time preference and social preference. Hence, we have a "unified framework".

The framework we propose has two persistent components: (i) a complete and transitive binary relation that determines which is the reference point and (ii) preference parameters (e.g., utility functions, discount rates) that depend on the reference point.

Our first step is a general representation theorem for choices in a generic domain: *Ordered Reference Dependent Utility (ORDU)* (Section 2). In this model, the decision maker uses a reference order to identify the reference alternative of a choice problem. In turn, this determines a utility function that she maximizes. Hence, it is as if that alternatives are ranked by their relevance in affecting preferences, and the underlying preference is determined by the alternative ranked highest in this order among those that are available.

The key behavioral postulate underlying the model is *Reference Dependence* (RD): it posits that if we fix the reference point, WARP holds. Since we do not know which is the reference point, we scarcely posit that there is one option in every choice problem such that if we keep said option when taking subsets, WARP holds. To illustrate, consider two choice sets $B \subset A$ such that WARP is violated; for instance, when an alternative is available in both A and B but is only chosen from A. Our axiom RD makes the behavioral assumption that the reference alternative of A is not present in B, causing a change in reference and WARP violation. Hence for any choice set A, RD demands that choices from subsets of A satisfy WARP as long as a certain (reference) alternative remains present. A formal definition is provided in Section 2. This axiom, along with a standard continuity assumption when X is infinite, characterizes the ORDU representation.

Next, we consider the special case of risk preference in Section 3. Here the postulate becomes: preserving one of the safest alternatives in a choice set preserves WARP *and* the Independence condition. This follows the same intuition we had above: maintaining the reference point, normative postulates hold; for risk preference, this includes Independence. We call this *Risk Reference Dependence*. A second axiom, *Monotone Risk Aversion*, requires that if we add alternatives to a choice problem, choices can only become more risk averse, since an even safer reference increases risk aversion. Together with standard continuity and first order stochastic dominance we obtain the *Avoidable Risk Expected Utility* (AREU) representation (formally presented in Section 3), in which a decision maker's utility function depends on the safest alternative available, and safer references lead to more concave utility functions.

We then turn to the time domain in Section 4. The standard model for time preferences is Exponentially Discounted Utility, yet it is routinely challenged in empirical studies in which consumers tend to exhibit less patience in short-term decisions, or *present bias*.[3] We propose a model in which the decision maker has a single utility function, maximizes exponentially discounted utility, but uses a discount factor indexed by the earliest available payment in a choice problem. The availability of a sooner alternative makes the decision maker impatient. The key axiom, *Reference Dependent Stationarity*, is the counterpart of RD, where now we require WARP *and* Stationarity to hold only when we preserve the earliest alternative. The reference effect is characterized by the axiom *Increasing Patience*, which simply posits that symmetrically advancing the options can only increase delay aversion. The resulting model gives rise to the well-known violation of dynamic consistency, in which the same delay between consumption is tolerable in the future but not in the present.

The application for social preference is studied in Section 5. Often viewed as a desire to be *fair*, subjects in economics and psychology experiments display behavior consistent with increased altruism when a more equitable split of payment is available.[4] In this setup, an alternative is an allocation for the decision maker and another individual. A natural measure of equity is the (normalized) ratio between the incomes, and attainable equity is therefore the maximum of such measure in a choice problem. As is standard for choices involving money, we use quasi-linear utility as foundation, but introduce the innovation that utility for money shared is increasing in attainable equity. This modification reflects our unified framework adapted to this setting—the presence of certain alternatives, as given by a reference order, affects the underlying preference for sharing. Like before, the main axiom *Reference Dependent Social Preference* posits the conformity with WARP and quasi-linear preferences when we preserve the most-equal option. A second axiom, *Increasing Altruism*, posits that decision makers are weakly more willing to share when more options are added to a choice problem, which can only increase attainable equity. In addition to capturing changing altruism, the model also explains increased sharing when splitting a fixed pie due to the availability of a more equitable division, as well as increased tendency to forgo a larger pie in favor of sharing a smaller one.

In our applications, failures of WARP and violations of structural assumptions are tightly linked. For example in the risk domain, adding full WARP to our model implies full compliance with Independence, and vice-versa. Therefore, our model departs from standard expected utility *only when* both WARP and Independence fail. Equivalent results obtain for the time and social domains. These findings separate our work from models that weaken a structural assumption but maintains WARP—the

[3]See for example Laibson (1997), Frederick et al. (2002), and Benhabib et al. (2010).
[4]See for example Ainslie (1992), Rabin (1993), Nelson (2002), Fehr & Schmidt (2006), and Sutter (2007).

equivalence of a stable preference, it also suggests that necessary violation of WARP in our models is the behavioral manifestation of changing preferences. It hence provides a new perspective to study classic paradoxes like Allais and present bias.

In relation to the existing literature, we first note that reference points are not exogeneously observed in our models. This strikes a fundamental difference in primitives/datasets to prospect theory by Kahneman & Tversky (1979), the endowment effect by Kahneman et al. (1991), and models of status quo bias led by Masatlioglu & Ok (2005).[5] Our models belong to a separate set of literature built on *endogenous reference*, where reference points are neither part of the primitive nor directly observable, such as in Kőszegi & Rabin (2006) and Ok et al. (2015). Unlike these models, our reference alternatives are given by a *reference order*. With this added structure, reference points—albeit unobservable—can be easily pinned down. Subsection 2.3 discusses in details the differences between ORDU and these two work, whereas Subsection 2.2 discusses the identification of our reference points and the consequent out-of-sample predictions via the reference order.

Our most general model, in which choices are over generic alternatives, is most similar in spirit and concurrent to Kıbrıs et al. (2018) albeit having different axiomatimization. Their paper focuses on choices over generic alternatives and contains no counterpart to our applications in the risk, time, and social domains. Their axiom depicts a conspicuity ranking between any two alternatives: if dropping x in the presence of y results in a WARP violation, then dropping y in the presence of x does not. Our approach is different and more involved, as it requires comparison between multiple choice problems differing by more than one alternative. However, this allows us to accommodate a wide range of behavioral postulates (in addition to WARP), such as the Independence and Stationarity conditions, with which we deliver reference-dependent expected utility and reference-dependent exponential discounting respectively. Moreover, their model is limited to a finite set of alternatives, whereas we allow the set to be any separable metric space. This is not (just) a technical contribution, as the added generality is indispensable for choices over lotteries.

We compare our applications in risk, time, and social preferences to existing models in their respective sections. However our main contribution is, instead of a single model that captures a specific departure from standard theory, a unified framework. The closest work that resembles a unified framework is *salience*, pioneered by Bordalo et al. (2012, 2013), in which options are evaluated differently depending on which attribute is salient. We are different in that our framework comprises

[5] For other models of status quo bias, see Masatlioglu & Ok (2013) and Dean et al. (2017). Ortoleva (2010) extends this idea to preferences under uncertainty.

of a systematic reference dependence approach of weakening normative postulates, with which we apply universally to the risk, time, and social domains. This approach allows us to study reference dependence in risk as related to reference dependence in time, and failure of WARP as related to failure of structural assumptions. Indispensable to this innovation is the use of choice correspondences as opposed to preference relations as primitive *and* foregoing WARP—the conventional "rationality" assumption increasingly scrutinized by empirical evidence. Otherwise, behavior is summarized by binary comparisons, leaving on the table useful information about how people make decisions in real-life situations where they choose from more than two alternatives. This richer scope allows us to utilize behavior from non-binary choice problems to study and explain anomalies traditionally found in binary choice.

The remainder of the paper is organized as follows. In Section 2, we provide the axioms and the representation theorem for a generic ordered-reference dependent utility representations. Later in that section, we introduce a companion result to incorporate the accommodation of properties other than WARP, and a template for additional structure in the reference order R. Section 3, Section 4, and Section 5 each provides a representation theorem under this unified framework for the risk, time, and social preference settings respectively, discusses the model's implications, as well as compares it to related models in the literature.

2 Ordered-Reference Dependence

We start with most general model, in which a decision maker chooses from generic alternatives.

2.1 Reference Dependent Choice

We introduce a reference-based approach of imposing a standard behavioral postulate. In this section, said postulate is WARP.

Let X be an arbitrary set of alternatives, $\mathcal{A}$ the set of all finite and nonempty subsets of X, and $c : \mathcal{A} \to \mathcal{A}$, $c(A) \subseteq A$, a choice correspondence. Recall that c satisfies WARP if for all choices problems A, B such that $B \subset A$, $c(A) \cap B \neq \emptyset$ implies $c(A) \cap B = c(B)$.[6]

Even though choices may violate WARP, it may still be the case that they comply with it among a subset of all choice problems $\mathcal{S} \subset \mathcal{A}$. We define this notion formally.

[6]For an arbitrary $\mathcal{A}$, this definition of WARP is weaker than another popular version: $x \in c(A)$, $y \in c(B)$, and $x, y \in A \cap B$ implies $x \in c(B)$. They are equivalent whenever $\mathcal{A}$ contains all doubletons and tripletons subsets of X.

Definition 1. Let $c : \mathcal{A} \to \mathcal{A}$ be a choice correspondence and $\mathcal{S} \subseteq \mathcal{A}$. We say c satisfies *WARP* over $\mathcal{S}$ if for all $A, B \in \mathcal{S}$,

$$B \subset A,\ c(A) \cap B \neq \emptyset \Rightarrow c(A) \cap B = c(B).$$

WARP is hence equivalent to the statement "c satisfies WARP over $\mathcal{A}$."

Our first axiom is a reference-based generalization of WARP.

Axiom 1 (Reference Dependence (RD))**.** *For every choice problem $A \in \mathcal{A}$, there exists an alternative $x \in A$ such that c satisfies WARP over $\mathcal{S} = \{B \subseteq A : x \in B\}$.*

Note that this axiom generalizes WARP, since "c satisfies WARP over $\mathcal{A}$" implies "c satisfies WARP over $\mathcal{S}$" for any $\mathcal{S} \subseteq \mathcal{A}$.

We explain the intuition of Axiom 1. Suppose choices between choice problems A and $B\,(\subset A)$ violate WARP; for example, $y \in c(A)$ but $y \in B \backslash c(B)$. We postulate that this is due to a change in reference point. Specifically, that the reference alternative of A must have been removed when take subset B of A, that is, it is in the set $A \backslash B$. Then, a natural limitation of WARP violations arise: have we *not* removed the reference alternative of A when taking an arbitrary subset B of A, choices would have complied with WARP. To put it differently, suppose that when taking subsets of A, if by preserving some alternative x in this process choices from these subsets comply with WARP. x is hence an endogenous candidate for "the reference alternative of A".[7] Axiom 1 demands that every choice problem contains (at least) one candidate alternative that achieves this.

Next we provide an example of compliance. Consider the following choice correspondence for $X = \{a, b, c, d\}$, where the notation $\{a, \underline{b}, c, d\}$ means b is chosen from the choice problem $\{a, b, c, d\}$.

$$\begin{array}{c} \{a, \underline{b}, c, d\} \\ \{a, \underline{b}, c\} \quad \{a, \underline{b}, d\} \quad \{a, c, \underline{d}\} \quad \{\underline{b}, \underline{c}, d\} \\ \{a, \underline{b}\} \quad \{\underline{a}, c\} \quad \{a, \underline{d}\} \quad \{\underline{b}, c\} \quad \{\underline{b}, d\} \quad \{\underline{c}, d\} \end{array}$$

This choice correspondence does not satisfy WARP globally (there are three instances of WARP violations: (i) between $\{a, \underline{b}, c, d\}$ and $\{\underline{b}, \underline{c}, d\}$, (ii) $\{\underline{b}, \underline{c}, d\}$ and $\{\underline{b}, c\}$, and (iii) $\{a, c, \underline{d}\}$ and $\{\underline{c}, d\}$). Yet WARP is satisfied from choice sets that contain a. To reconcile with Axiom 1, when $S = X$, a is a candidate reference alternative. This is also true for any choice set S that contains a. Likewise, for $S = \{b, c, d\}$, d is a candidate reference, and this is true for any choice set S that contains d but not a.

[7] Using the language in Ok et al. (2015), this alternative can be called a *potential reference alternative* of A.

The only choice set left to be checked is $S = \{b, c\}$, but since the only non-singleton subset of $\{b, c\}$ is itself, WARP is trivial.

Although Axiom 1 allows for WARP violations, it is falsifiable as long as $|X| \geq 3$ (i.e., as soon as WARP is non-trivial). For example, the following choice correspondence violates Axiom 1.

$$\boxed{\{a, \underline{b}, c\} \quad \{\underline{a}, b\} \quad \{b, \underline{c}\} \quad \{\underline{a}, c\}}$$

In this example, instances of WARP violations are (i) between $\{a, \underline{b}, c\}$ and $\{\underline{a}, b\}$ and (ii) between $\{a, \underline{b}, c\}$ and $\{b, \underline{c}\}$. So when $A = \{a, b, c\}$, a does not preserve WARP since the first instance is not excluded, b does not preserve WARP since neither instance is excluded, and c does not preserve WARP since the second instance is not excluded. Hence the axiom does not hold.

Another way of "measuring" falsifiability is to count the number of observations (choice problems) required to falsify an axiom. For standard WARP that number is 2: for example, when WARP is violated between $\{a, \underline{b}, c\}$ and $\{\underline{a}, b\}$. Whereas for Axiom 1, a weakening of WARP, that number is 3: for example $\{a, \underline{b}, c\}$, $\{\underline{a}, b\}$, and $\{\underline{a}, c\}$, since the reference of $\{a, b, c\}$ is in $\{a, b\}$ and/or $\{a, c\}$, but WARP is violated both between $\{a, \underline{b}, c\}$, $\{\underline{a}, b\}$ and between $\{a, \underline{b}, c\}$, $\{\underline{c}, b\}$.[8] Thus reference dependence makes Axiom 1 harder to reject relative to WARP by one additional observation.

When X is infinite, we also assume Continuity. Say (X, d) is a metric space.

Axiom 2 (Continuity)**.** *We say $c : \mathcal{A} \to \mathcal{A}$ satisfies Continuity if it has a closed-graph (with respect to the Hausdorff distance): $x_n \to_d x$, $A_n \to_H A$, and $x_n \in c(A_n)$ for every $n = 1, 2, ...$ implies $x \in c(A)$.*[9]

2.2 Ordered-Reference Dependent Utility Functions

Let R be a complete and transitive binary relation, $\arg\max_{x \in A} R$ denotes the set $\{x \in A : xRy \, \forall y \in A\}$.

Definition 2 (Ordered-Reference Dependent Utility)**.** c admits an Ordered Reference Dependent Utility (ORDU) representation if there exist a complete, transitive, and antisymmetric reference order R on X and a set of reference-indexed utility functions $\{u_x : X \to \mathbb{R}\}_{x \in A}$ such that

$$c(A) = \arg\max_{y \in A} u_{r(A)}(y),$$

where $r(A) = \arg\max_{x \in A} R$.

[8] This can be generalized: Axiom 1 is falsified when there are WARP violations between A, B_1 and between A, B_2 such that $B_1 \cup B_2 = A$, where $A, B_1, B_2 \in \mathcal{A}$.

[9] By $\to_H$ we mean convergence in the Hausdorff distance, defined by $d_H(X, Y) = \max\left\{\sup_{x \in X} \inf_{y \in Y} d_2(x, y), \sup_{y \in Y} \inf_{x \in X} d_2(x, y)\right\}$.

Proposition 1.

1. *Let X be a finite set. c satisfies RD if and only if it admits an ORDU representation.*

2. *Let X be a separable metric space. c satisfies RD and Continuity if and only if it admits an ORDU representation where $c(A) = \arg\max_{y \in A} u_{r(A)}(y)$ has a closed-graph.*

ORDU represents a special type of context-dependent preferences. A decision maker's preference may change with the choice set, but depends only on its reference alternative, characterized by *reference-dependent utilities*. Reference-dependent utilities are more restrictive than *set-dependent utilities*, where each choice problem has its own utility function.[10] When $|X|$ is finite, there are at most $|X|$ distinct utility functions but around $2^{|X|}$ choice problems, and this difference increases exponentially in $|X|$. Furthermore, a linear order, called *reference order*, uniquely pins down the reference point for each choice problem.

The reference order has natural interpretations in richer settings, as we demonstrate in the risk, time, and social preference sections. When the setting is choices over generic alternatives, an interpretation of the reference order is a subjective salience ranking of alternatives. The most salient alternative determines the underlying preference used with the problem. In this setting, it is as if that the decision maker's attention is drawn to a certain salient alternative, and her preference ranking depends on that alternative. It is the fact that her attention is not always drawn to the same (reference) alternative that gives rise to WARP violations. But when she has the same reference alternative for a set of choice problems, her choices are consistent with a stable preference ranking.

The suggestion that certain salient component of a choice problem affect choices is not new, for example Bordalo et al. (2012, 2013). In their model, alternatives have attributes, and depending on which alternatives are being compared certain attributes are more salient than others, and weighted differently, from one choice problem to another. This is the source of WARP violations in their model. In ORDU, attributes are not part of the primitive/model, allowing for a different characterization of salience when the modeler either does not observe attributes or do not know the relevant attributes that play in role in decision making.

Combining *reference-dependent utilities* with a *reference order* yields out-of-sample predictions. For example, when the reference alternative in choice problem A is present in the choice problem $B \subset A$, that alternative is still the reference, and the preference ranking remains the same. This is a

[10] Set-dependent utilities, that each choice problem A has a utility function $u_A(x)$ that is maximized, puts no restriction on behavior, since we can simply set $u_A(c(A)) = 1$ and $u_A(x) = 0$ for all $x \neq c(A)$.

feature of the reference order. So once we have identified the reference alternative r of A, we know that subsets of A that contain r use the same utility function.

Furthermore, reference alternatives are identified whenever we observe WARP violation upon removing them, since it is *only* through changes in references points that WARP violations arise. For example if WARP is violated between A and B, and $B = A \setminus \{x\}$, then x is the reference of A. Hence we infer the presence of reference points, and pin them down uniquely, through inconsistency or incoherence in choice with respect to WARP. Instead, if a decision maker complies with WARP, the idea that preferences are reference dependent cannot be substantiated.

Together, the model allows us to make out-of-sample predictions between two worlds: On one end, the decision maker's reference alternatives are not identifiable with choice data precisely because she satisfies WARP and maximizes a single preference ranking; on the other end, the decision maker's choices are reference dependent and result in WARP violations, which allows us to identify reference points and subsequently make predictions using the reference order.

The idea behind ORDU has natural applications. The rest of the paper demonstrates it in the risk, time, and social preference domains. In each setting, a domain-specific interpretation is given to the reference order. In the risk setting, the minimum amount of risk the decision maker must take, as measured by the safest alternative in a choice problem, may influence risk aversion. In the time setting, the earliest available payment characterizes the imminence of consumption, which may affect the decision maker's patience. In the social preference setting, how much a decision maker is allowed to share may affect how generous she is, where greater attainable equity increases generosity. We formally these models in Section 3, Section 4, and Section 5 respectively.

2.3 Comparison with other non-WARP models

We conclude this section with a summary discussion of ORDU as compared with other axiomatized utility representations that model WARP violations. Technical details and complete characterizations of the examples used are deferred to Appendix B.

Under comparable setups, ORDU neither nests nor is nested by any of the following models: (i) Ok et al. (2015)'s *revealed (p)reference*, (ii) Kőszegi & Rabin (2006)'s *reference-dependent preferences* (*personal equilibrium*), (iii) Manzini & Mariotti (2007)'s *rational shortlist method*, and (iv) Masatlioglu et al. (2012)'s *choice with limited attention*.[11] Our focus will be the former two, as they also use

[11] That is, for each of the four external models we consider, there are choice correspondences that admit ORDU but not the eternal model and vice versa. Complete specifications of the choice correspondences used are provided in Appendix

reference formation to explain choices. The latter two are models in which reference formation is not used, but the addition (or removal) of alternatives directly contribute to WARP violations.

A key observation separates ORDU from other non-WARP models. In ORDU, reference points are given by a reference order and choices maximize reference-dependent utilities. Thus, either $c(\{a,b\})$ or $c(\{b,c\})$ must agree with $c(\{a,b,c\})$ in terms of being consistent with the maximization of a single utility function. This is because the reference point of $\{a,b,c\}$ is either in $\{a,b\}$ or in $\{b,c\}$ (or in both). In fact, since the reference point of A is in every subset of A that contains it, this condition generalizes into Remark 1.

Remark 1. Suppose c admits an ORDU representation. Take any finite collection of choice problems $A_1, ..., A_n$. For some $x \in A_1 \cup A_2 \cup ... \cup A_n$, choices from the collection of choice problems $\{A_i : x \in A_i\}$ must comply with standard utility maximization.

Recall that notation "$\{\underline{a}, b, c\}$" means choice problem $\{a,b,c\}$, from which a is chosen.

In Ok et al. (2015)'s (endogeneous) *reference dependent choice*, the decision maker maximizes a single utility function, but only chooses from alternatives that are better than the reference in all (endogenously determined) attributes, where references do not necessarily come from an order. In their model, a *decoy* d may block the choice of a in $\{a, \underline{b}, d\}$ and $\{a, \underline{c}, d\}$ due to the *attraction effect,* where b and c are *elevated* since they are better than d in all attributes, but a isn't. However, since reference formation is more flexible, as contrasted with the use of a reference order in ORDU, d need not be the decoy in $\{\underline{a}, b, c, d\}$, resulting in the choice of a. Remark 1 excludes this behavior from ORDU. Conversely, intransitive behavior in binary choice problems, such as $\{\underline{a}, b\}, \{\underline{b}, c\}, \{\underline{c}, a\}$, can be explained by ORDU but are ruled out by Ok et al. (2015), since the absence of a third alternative impedes their decoy-effect from taking place. Hence the two models are not nested.

Kőszegi & Rabin (2006)'s *reference-dependent preferences* is another related model. Gul et al. (2006) provided the axiomatic foundation for *personal equilibrium* (PE), in which a decision maker has a joint utility function $v : X \times X \to \mathbb{R}$ and chooses $PE(A) = \{x : v(x|x) \geq v(y|x)\ \forall y \in A\}$. That is, the choice maximizes a reference-dependent utility function, and the reference point is itself the eventually chosen alternative (hence "equilibrium"). This permits the following behavior: b is not chosen in $\{\underline{a}, b, c, d\}$, but is chosen in the subset $\{\underline{a}, \underline{b}, c\}$, in which d—an alternative better than b under $v(\cdot|b)$—was removed. Suppose further that d is chosen in $\{\underline{a}, \underline{d}\}$ for the same reason—c is better than d under $v(\cdot|d)$, but now c is removed. Remark 1 concludes that this behavior cannot be accommodated

B.

by ORDU, a consequence of the reference order, where the reference point of $\{\underline{a}, b, c, d\}$ and $\{\underline{a}, \underline{d}\}$ must both be d under ORDU. Hence a WARP violation between these two choice problems rules out ORDU. Conversely, an immediate implication PE is, if $x \in c(A)$ and $x \in B \subset A$, then $x \in c(B)$. A simple intransitive choice pattern $\{\underline{a}, b\}, \{\underline{b}, c\}, \{\underline{c}, a\}$ is hence admissible by ORDU but not PE.[12] We conclude that the two models are not nested.

Non-nestedness between ORDU and Manzini & Mariotti (2007)'s *rational shortlist method*, as well as between ORDU and Masatlioglu et al. (2012)'s *choice with limited attention*, are shown and explained in Appendix B.

2.4 A unified framework for structural anomalies

Reference Dependence (Axiom 1) weakened WARP by demanding that WARP is satisfied among choice problems that share a reference alternative (as opposed to all choice problems). This method of generalizing an axiom is not only applicable to WARP, but also a wide range of behavioral properties defined on choice behavior. For example, we can demand compliance with Independence in a similar way, where Independence is not necessarily satisfied between every two choices, but is complied with whenever the choices come from choice problems that have the same reference point.

This reference dependence approach of weakening an arbitrary postulate serves as the starting point of our models in the risk, time, and social domains. In their respective sections, we adapt Axiom 1 to postulates of the form "For every choice problem A, there exists an alternative $x \in A$ such that c satisfies $\mathcal{T}$ over $\mathcal{S} = \{B \subseteq A : x \in B\}$". $\mathcal{T}$ is "WARP and Independence" for the risk domain, "WARP and Stationarity" for the time domain, and "WARP and quasi-linearity" for the social domain.

The result is as anticipated—ordered-reference expected utility, ordered-reference exponentially discounted utility, and ordered-reference quasi-linear utility. In fact, the representation theorems for all four models in the present paper start with a quintessential result in Appendix A, Lemma 2, which demonstrates the wide applicability of our approach by accommodating a class of behavioral postulates we call *finite properties*, of which WARP, Independence, Stationarity, quasi-linearity, transitivity, convexity, monotonicity, stochastic dominance, etc. are examples. Then, complemented with additional structure on reference orders, we obtain the reference-dependent versions of the corresponding utility representations.

The next three sections are natural applications of this approach.

[12] Gul et al. (2006) shows that Kőszegi & Rabin (2006)'s *personal equilibrium* is equivalent to the maximization of a complete (but not necessarily transitive) preference relation.

3 Risk Preferences

We now turn to an application in the domain of risk, where we provide a utility representation, with axiomatic foundation, that explains increased risk aversion when safer options are present than when they are not. Consider a decision maker whose willingness to take risk depends on how much of it is avoidable, as measured by the safest alternative among those that are available. This depends on the underlying choice set: Sometimes, we have the option to fully avoid risk by keeping our asset in cash or by buying an insurance policy, and so the safest option is quite safe. In other situations, all options are risky and we are forced to take some risk, and so the safest option is quite risky. The premise of our model, in the risk setting, is that a decision maker's risk aversion may differ between these two types of choice problems in a particular way: she could be more risk averse when risk is avoidable than when it is not.

Suggestive evidence for this behavior is present in the literature. In the well-known paradox introduced by Allais (1953), when one choice problem contains a safe option and the other does not, subjects tend to chose the safer option in the former. This observation is consistent with increased risk aversion when safer options are present. We provide a quick recap of the Allais paradox and its relevance as pertain to our model when we discuss applications. We will also present a result that shows that Allais-type behavior is the consequence of changing utility functions by concave transformations, which characterizes greater risk aversion under the expected utility form.

In a separate setting meant to test for the *compromise effect*, Herne (1999) showed that the presence of a safer option results in WARP violations in the direction of more risk averse behavior. Wakker & Deneffe (1996) introduced the *tradeoff method* to elicit risk aversion without using a sure prize and showed that the estimated utility functions are in general less concave relative to the standard certainty equivalent / probability equivalent methods.[13] Andreoni & Sprenger (2011) reinforces this observation when the safest option is close to certainty.

3.1 Preliminaries

Consider a finite set of prizes $X \subset \mathbb{R}$. Let $\Delta(X)$ be the set of all lotteries over X endowed with the Euclidean metric d_2. Let $\mathcal{A}$ be the set of all finite and nonempty subsets of $\Delta(X)$. We call $A \in \mathcal{A}$ a choice problem. We take as primitive a choice correspondence $c : \mathcal{A} \to \mathcal{A}$ that gives, for each choice

[13] *Certainty equivalent method* finds the value of a sure prize such that a subject is indifferent to a fixed lottery. *Probability equivalent metho*d fixes the sure prize and alters the probability of a lottery until the subject is indifferent. *Tradeoff method* finds the indifferent point between two lotteries by varying one of the prizes.

problem A, a subset $c(A) \subseteq A$. We make the following notational simplifications: Per convention, δ_x denotes the lottery that gives prize $x \in X$ with probability 1. For $p, q \in \Delta(X)$ and $\alpha \in [0,1]$, we denote by $p^\alpha q$ the convex combination $\alpha p + (1-\alpha) q \in \Delta(X)$. Let $b := \max_{\geq} X$ and $w := \min_{\geq} X$ denote the highest and lowest prizes respectively. We denote by $q(x)$ the probability lottery q gives prize $x \in X$.

3.2 Risk Reference Dependent Choice

Recall that in Section 2 we defined what it means for WARP to hold on an arbitrary set of choice problems. We now do the same for *Independence.*

Definition 3. Let $c : \mathcal{A} \to \mathcal{A}$ be a choice correspondence and $\mathcal{S} \subseteq \mathcal{A}$. We say c satisfies *Independence* over $\mathcal{S}$ if for all $A, B \in \mathcal{S}$ and $\alpha \in (0,1)$,

1. $p \in c(A)$, $q \in A$, $q^\alpha s \in c(B)$ and $p^\alpha s \in B \Rightarrow p^\alpha s \in c(B)$, and

2. $p^\alpha s \in c(A)$, $q^\alpha s \in A$, $q \in c(B)$ and $p \in B \Rightarrow p \in c(B)$.

In standard expected utility, c satisfies both WARP and Independence over $\mathcal{A}$.

Now, we are interested in the behavior where changes in the safest available alternatives affect risk aversion, but WARP and Independence are complied with whenever the safest alternatives of a collection of choice problems are the same.

First we define what "the safest available alternative" means through the use of two partial orders. A *mean-preserving spread* (MPS) is clearly not safest, this is our first order. However, mean-preserving spread is a (very) partial order, and many lotteries are left unranked, making it hard to predict when should WARP and Independence hold.

To account for this limitations we also deem riskier any lottery that is an *extreme spread*, our second risk order, which we now define. We call p is an extreme spread of q ($p\mathrm{ES}q$) if $p = \beta q + (1-\beta)(\alpha(\delta_b) + (1-\alpha)(\delta_w))$ for some $\beta \in [0,1]$ and $\alpha \in (q(b), 1-q(w))$. This captures lotteries that assigned more probability to extreme prizes while being proportionally identical for intermediate prizes. Extreme spread shares the core intuition of Aumann & Serrano (2008)'s risk index (which in their paper only applies to gain-loss prospects), where lotteries are deemed safer than another in the "economics sense"—if more-risk-averse decision makers prefer them whenever less-risk-averse decision makers do.[14]

[14]For every q, the set of extreme spreads of q is small and lives entirely within the probability triangle containing q, δ_b,

The two risk orders are compatible with each other but are not nested. Now we characterize the set of alternatives that are not risky by these two measures. Let $MPS(A) = \{p \in A : \exists q \in A \text{ s.t. } p\text{MPS}q\}$ and $ES(A) = \{p \in A : \exists q \in A \text{ s.t. } p\text{ES}q\}$ denote the mean-preserving spreads and extreme spreads in A respectively.

Definition. Define the least risky set of A by $\Psi(A) := A \backslash (MPS(A) \cup ES(A))$.

We now replace Reference Dependence from Section 2 with a stronger axiom that demands (i) reference-dependent compliance with both WARP *and* Independence and that (ii) the reference is a least risky lottery.

Axiom 3 (Risk Reference Dependence (RRD))**.** *For every choice problem $A \in \mathcal{A}$, there exists $p \in \Psi(A)$ such that c satisfies WARP and Independence over $\mathcal{S} = \{B \subseteq A : p \in B\}$.*

Axiom 3 identifies a candidate reference for choice problem A. If $\Psi(A) = \{p\}$, and B_1 and B_2 are subsets of A containing p, then neither a violation of WARP nor a violation of Independence is produced between $c(B_1)$ and $c(B_2)$.

Like Reference Dependence (Axiom 1 in Section 2), Risk Reference Dependence postulates that there is a reference point, in the sense that WARP holds in its presence. But it additionally postulates that Independence also holds, and that this reference point is in $\Psi(A)$—a least risky lottery.[15]

Finally, note that Axiom 3 weakens the axioms of standard expected utility, which demands compliance of WARP and Independence over the entire $\mathcal{A}$.

3.3 Monotone Risk Attitude

Our motivation is that the decision maker's choices vary only in terms of magnitude of risk aversion. Moreover, increases in risk aversion are due to the presence of a safer alternative. Consider the following axiom.

Axiom 4 (Monotone Risk Attitude)**.** *For any choice problems $A, B \in \mathcal{A}$ such that $B \subset A$,*

$$\delta^\alpha r \in c(B),\ p^\alpha r \in B, p^\beta q \in c(A),\ \delta^\beta q \in A \ \Rightarrow\ \delta^\beta q \in c(A),$$

and δ_w. In this probability triangle, it is exactly the set of lotteries such that a more risk loving decision maker would prefer (over q) whenever a more risk averse one does, under the framework of standard expected utility. In particular, it is a superset of mean-preserving spreads in this triangle. The intuition behind this notion is that, when probabilities are allocated to the most extreme prizes, even if mean is not preserved, we should still deem the resulting lottery riskier. Note that an extreme spread need not be a mean-preserving spread, and vice versa.

[15] This is where the decision maker's subjectivity enters the model: For two lotteries not ranked by objective notions of risk, one individual may deem one lottery riskier, whereas another individual disagrees. The axiom demands that a reference point exists and is a least risky alternative, but in instances where $|\Psi(A)| > 1$, the decision maker's choices determine which lottery in $\Psi(A)$ is the reference.

where $p, q, r \in \Delta(X)$, δ is a degenerate lottery, and $\alpha, \beta \in [0,1]$.

It is standard that $\succsim_1$ is deemed *more risk averse* than $\succsim_2$ if for any degenerate alternative δ and lottery p, $\delta \succsim_2 p \Rightarrow \delta \succsim_1 p$. Here we extend this definition to lotteries that are not entirely riskless, but differ by a degenerate and (possibly) non-degenerate components: $\delta^\alpha q$ and $p^\alpha q$ (where δ is a degenerate alternative).[16] Under standard expected utility this extension is without loss, i.e., the two notions coincide.[17] It is precisely because we depart from the standard expected utility model that we require this extended definition—the choice between δ and p does not pin down the choice between $\delta^\alpha q$ and $p^\alpha q$ due to changing risk aversion.

Axiom 4 postulates that as a choice problem expands, the decision maker is *not more risk loving.* Our intuition is that the introduction of new alternatives can only reduce minimum risk / increase avoidable risk, and consequently the decision maker views risk less favorably and becomes more risk averse in the choices she makes. This is the source of Independence violation in our model, but only one type of violation is allowed: that choices become *more risk averse*, where other channels remain shut.

Note that standard expected utility satisfies this axiom trivially—an expected utility maximizer can neither be more risk loving nor more risk averse between any two choice sets, a consequence of the Independence axiom. Therefore, our departure from expected utility is to permit increased risk aversion when new alternatives are added to a choice set.

Final two axioms are standard: that choice is continuous (defined in Section 2) and abides by first order stochastic dominance.

Axiom 5 (FOSD)**.** *For any $p, q \in \Delta(X)$ such that $p \neq q$ and p first order stochastically dominates q, $p \in A$ implies $q \notin c(A)$.*

3.4 Representation Theorem

We now introduce the utility representation.

Definition. We say an order R is risk-consistent if, whenever (i) p is a mean-preserving spread of q or (ii) p is an extreme spread of q (or both), we have qRp.

[16]It is straightforward to show that $p^\alpha q$ is obtained from $\delta^\alpha q$ by moving probabilities from one prize to one or more prizes. We hence deem $\delta^\alpha q$ safer than $p^\alpha q$, and say that a *more risk averse* decision maker prefers $\delta^\beta q$ to $p^\beta q$ whenever a less risk averse decision maker prefers $\delta^\alpha r$ to $p^\alpha r$.

[17]This is the consequence of the Independence axiom of standard expected utility, in which $\delta^\alpha q$ is chosen over $p^\alpha q$ if and only if $\delta^\beta r$ is chosen over $p^\beta r$ if and only if δ is chosen over p.

Definition 4. c admits an Avoidable Risk Expected Utility (AREU) representation if there exist (i) a complete, transitive, and antisymmetric reference order R on $\Delta(X)$ and (ii) a set of strictly increasing utility functions $\{u_p : X \to [0,1]\}_{p\in\Delta(X)}$, such that

$$c(A) = \arg\max_{p\in A} \mathbb{E}_p u_{r(A)}(x),$$

where

- $r(A) = \arg\max_{q\in A} R$,
- R is risk-consistent,
- qRp implies $u_q = f \circ u_p$ for some concave $f{:}[0,1] \to [0,1]$, and
- $\arg\max_{p\in A} \mathbb{E}_p \left[u_{r(A)}(x)\right]$ has a closed-graph.

Proposition 2. *Let $c : \mathcal{A} \to \mathcal{A}$ be a choice correspondence. The following are equivalent:*

1. *c satisfies Risk Reference Dependence, Monotone Risk Attitude, FOSD and Continuity.*
2. *c admits an AREU representation.*

Furthermore in every AREU representation, given R, u_p is unique for all $p \neq (\delta_b)^{\alpha}(\delta_w)$.

When choices admit an AREU representation, it is as if the decision maker goes through the following decision making process: Facing a choice problem, she first looks for the safest alternative using R, which is risk consistent—it ranks safer alternatives higher. This determines the (Bernoulli) utility function for the choice problem and she proceeds to choose the option that maximizes expected utility. Moreover, the safer the reference, a more concave utility function is used, resulting in weakly more risk averse choices. This generalizes the standard model where a decision maker chooses the option that maximizes expected utility using a single utility function. It departs from standard expected utility by allowing greater risk aversion when alternatives are added to a choice set, but prohibits any other types of preference changes.

Note that utility functions in AREU are generically unique (up to an affine transformation). This property guarantees that their relationships by concave transformations are not arbitrary, and choices manifest changing risk aversions. Here, each utility function u_p is used to evaluate options for a set of choice problems that deem p as the safest alternative. When $p \neq (\delta_b)^{\alpha}(\delta_w)$, there are many of these choice problems in which p is not the chosen alternative, making u_p non-arbitrary.

3.5 Applications of AREU

We now show that AREU is compatible with the Allais paradox.

In experimental settings, subjects tend to choose the degenerate lottery $p_1 = \delta_{3000}$ over the lottery $p_2 = 0.8\delta_{4000} + 0.2\delta_0$, but choose $q_2 = 0.2\delta_{4000} + 0.8\delta_0$ over $q_1 = 0.25\delta_{3000} + 0.75\delta_0$. Note that the second pair of options are derived from the first pair using a common mixture, $q_1 = 0.2p_1 + 0.8\delta_0$ and $q_2 = 0.2p_2 + 0.8\delta_0$. Under expected utility theory, those who prefer p_1 to p_2 should prefer q_1 to q_2, and vice versa. Hence choices of p_1 and q_2 is a direct contradiction of expected utility theory. This is called the Allais paradox (and in particular the *common ratio effect*), a prominent "anomaly" in the study of choices under uncertainty, began with Allais (1953).[18]

AREU is compatible with this phenomenon. Given a reference order R that deems the safest alternative in the first choice problem—in which a sure prize is available—as safer, a decision maker is more risk averse and uses a more concave utility function. That is, where $A = \{p_1, p_2\}$ and $B = \{q_1, q_2\}$, we have $u_{r(A)} = f \circ u_{r(B)}$ for some concave transform f.[19] It is because of this change in utility function characterizing increased risk aversion that makes p_1, the safe option, appealing in the first choice problem. More generally, AREU captures the favoring of risk-free options whenever they are available, similar to Kahneman & Tversky (1979)'s *certainty effect.*

However, AREU is incompatible with choices of p_2 and q_1—violation of expected utility theory in the opposite direction. Since the decision maker is more risk averse in the first choice problem, if instead p_2 is chosen over p_1, then q_2 must be chosen over q_1 in the second choice problem since a less concave utility function will make q_2 more appealing. Analogously, a choice of q_1 implies a choice of p_1. To summarize, behaviors compatible with AREU are: (p_1, q_1), (p_2, q_2), and (p_1, q_2). The opposite behavior in which the decision maker in more risk loving in the first choice problem, (p_2, q_1), is ruled out.[20] This stipulates a specific type of departure from standard expected utility—more risk aversion in the presence of safer alternatives.

AREU's compatibility with the Allais paradox is not limited to the above specification. Moreover, it captures both the *common ratio effect* and the *common consequence effect.* We generalize the previous arguments when $|X| := |\text{supp}(\{\delta, p, q\})| = 3$ in the following statement:

[18]This example is taken from Starmer (2000). Camerer (1995) and Starmer (2000) provide an in-depth survey.

[19]The choice in the first problem can be explained by a (Bernoulli) utility function u_A if and only if, after normalization ($u_A(0) = 0$ and $u_A(4000) = 1$), $u_A(3000) > 0.8$. Similarly, the choice in the second problem can be explained by a normalized utility function u_B if and only if $u_b(3000) < 0.8$. This is from solving $u_A(3000) > 0.8u_A(4000) + 0.2u_A(\$0)$ and $0.25u_B(3000) + 0.75u_B(0) > 0.2u_B(4000) + 0.8u_B(0)$ for $u_A(3000)$ and $u_B(3000)$. Furthermore, as long as $u_A(3000) > u_B(3000)$, we have $u_A = f \circ u_B$ for some concave $f : [0,1] \to [0,1]$.

[20]This is consistent with the behavioral postulate referred to as *Negative Certainty Independence* in Dillenberger (2010); Cerreia-Vioglio et al. (2015).

Fact. *Consider a degenerate lottery* δ *and a lottery* p *such that neither of them first order stochastically dominates another. Consider lotteries* $\delta' = \delta^{\alpha} q$ *and* $p' = p^{\alpha} q$ *for any* $\alpha \in (0,1)$ *and lottery (degenerate or otherwise)* q*. Suppose* $|X| = 3$*, then*

1. *If* $\delta \in c(\{\delta, p\})$ *and* $p' \in c(\{\delta', p'\})$*, then*

   (a) *For all* $u_1, u_2 : X \to \mathbb{R}$ *such that* u_1 *explains the first choice and* u_2 *explains the second,* $u_1 = f \circ u_2$ *for some concave function* $f : \mathbb{R} \to \mathbb{R}$*.*

   (b) *Moreover, the choices admit an AREU representation such that* $r(\{\delta, p\}) R r(\{\delta', p'\})$*.*

2. *If the choices* $c(\{\delta, p\})$*,* $c(\{\delta', p'\})$ *admit an AREU representation, then*

   (a) *If* $p \in c(\{\delta, p\})$*, then* $p \in c(\{\delta', p'\})$*.*

   (b) *If* $\delta' \in c(\{\delta', p'\})$*, then* $\delta \in c(\{\delta, p\})$*.*

Last we consider one other application of AREU. A known phenomenon in behavioral finance is *reaching for yield*, in which investors invest less when the risk-free rate is higher, which is at odds with the standard expected utility model with commonly used specifications such as those that exhibit constant relative risk aversion. Lian et al. (2017) shows that this behavior is at odds with utility functions exhibiting constant or decreasing absolute risk aversion, capturing a large class of utility functions typically used in behavioral finance. The authors also provided evidence of this behavior.

AREU is consistent with this observation, and more specifically that the addition of a *better* sure prize increases risk aversion. Consider a choice set A which contains a sure prize of \$5 and choice set B which is A with an added option: a sure prize of \$7. Although risk is fully avoidable in both choice problems due to the availability of sure prizes, it is intuitive that a decision maker may display greater risk aversion in B. AREU captures this behavior using the specification $\delta_x R \delta_y$ whenever $x > y$, in which the decision maker maximizes expected utility with a more concave utility function when a *better* sure prize is present. This extends the intuition of increased risk aversion from the addition of safer alternative to the addition of safer *and* better alternatives.

3.6 Linkage between violations of WARP and violations of Independence

Risk Reference Dependence (Axiom 3) demands WARP and Independence over certain subsets of all choice problems. Given an AREU choice correspondence, we now state the consequences of imposing each of Transitivity, WARP, and Independence over the (entire) set of all choice problems $\mathcal{A}$. It turns out adding any one of these assumptions bring us back to standard expected utility, giving us a formal separation of AREU from the wide range of non-expected utility models in which WARP / complete preference ranking is maintained.

First, we adapt Transitivity, typically defined on a preference relation, to the framework of choice.

Definition 5. Let $c : \mathcal{A} \to \mathcal{A}$ be a choice correspondence. We say c satisfies *Transitivity* over $\mathcal{S} \subseteq \mathcal{A}$ if for any $\{p, q\}, \{q, s\}, \{s, p\} \in \mathcal{S}$,

$$p \in c(\{p, q\}) \text{ and } q \in c(\{q, s\}) \Rightarrow p \in c(\{q, s\}).$$

Transitivity on c is analogous to the standard definition of Transitivity on a preference relation (where $p \succsim q$ and $q \succsim s \Rightarrow p \succsim s$). In the framework of choice, c may satisfy Transitivity over the entire collection of choice problems $2^Y \setminus \{\emptyset\}$ but violate WARP. For example: $x \in c(\{x, y\})$, $y \in c(\{y, z\})$, $x \in c(\{x, z\})$, but $z \in c(\{x, y, z\})$. However, the reversed implication is true: If c satisfies WARP over $2^Y \setminus \{\emptyset\}$, then c satisfies Transitivity over $2^Y \setminus \{\emptyset\}$.[21]

We say c admits a utility representation if there exists $U : \Delta(X) \to \mathbb{R}$ such that $c(A) = \arg\max_{p \in A} U(p)$.

Proposition 3. *Suppose c admits an AREU representation. The following are equivalent:*

1. *c satisfies Transitivity (over $\mathcal{A}$).*
2. *c satisfies WARP (over $\mathcal{A}$).*
3. *c satisfies Independence (over $\mathcal{A}$).*
4. *c admits an expected utility representation.*
5. *c admits a utility representation.*

This result states that AREU cannot independently accommodate Transitivity, WARP, or Independence. Put it differently, although AREU weakens multiple postulates (WARP and Independence), it is in fact a "tight" deviation from standard expected utility in that the reinstiution of either postulate recovers the expected utility model.[22] In particular, the intersection of AREU and models maintaining

[21] When $S \neq 2^Y \setminus \{\emptyset\}$, Transitivity and WARP are not nested. For instance take $\mathcal{S} = \{\{x, y\}, \{y, z\}, \{x, z\}\}$, the choice correspondence $x \in c(\{x, y\})$, $y \in c(\{y, z\})$, $z \in c(\{x, z\})$ satisfies WARP (over $\mathcal{S}$) but not Transitivity.

[22] Transitivity and WARP do not imply one another. While obtaining Transitivity from WARP requires little additional assumptions, the other direction is typically difficult to achieve without explicitly assuming WARP. A classic example of a choice correspondence that satisfies Transitive but not WARP is $c(\{a, b\}) = \{a\}$, $c(\{b, c\}) = \{b\}$, $c(\{a, c\}) = a$, $c(\{a, b, c\}) = b$. In models of context-dependent choice, violations of WARP are sometimes accommodated alongside violations of Transitivity, and in other cases they are accommodated whilst keeping Transitivity imposed.

WARP, or equivalently a complete and transitive preference relation, is expected utility.

AREU is a model that explains non-expected utility behavior as a consequence of changing (risk) preferences. To this end, Proposition 3 shows that AREU leaves no extra explanatory power in explaining the violations separately. Instead, the violations are inextricably linked to one another, and resolving either one will bring us back to standard expected utility. When a choice correspondence admits a utility representation, choices are interpreted as the consequence of a stable preference ranking. Proposition 3 formally states that AREU is in line with this interpretation—standard expected utility ensues when preferences over lotteries are stable, and failure of expected utility is due to changing preferences.

This sets us apart from the majority of non-EU models where Independence is weakened under the assumption of WARP. In those cases, a single utility function is maximized, but it doesn't take the expected utility form. In our case, choices come from utility functions that conform with the expected utility form, but there are many of them, which characterize changing preferences. This is the result of a joint weakening of both WARP and Independence, the core idea of the unified framework we propose. In the time and social domain, we show that the same results and arguments hold.

3.7 Comparison to other non-expected utility models

In this section, we investigate the consequence of adding various restrictions to AREU, and use them to explore the relationship between AREU and other non-expected utility models of risk preferences.

3.7.1 AREU with Transitivity in a probability triangle

While Proposition 3 provides a strong separation between AREU and many non-EU models, we can more meaningfully recover the extent to which AREU is related to other models by imposing Transitivity "partially". To this end, we turn our attention to Marschak-Machina triangles (also "probability triangle") for the next part of our analysis.

We will show that AREU is very close to Betweenness, a well-known property first introduced (on preference relations) by Chew (1983); Fishburn (1983); Dekel (1986). Like expected utility, models of betweenness preferences have the characteristic of linear indifference curves (or in higher dimensions, hyperplanes). Their departure from standard expected utility is that the indifference curves need not have the same slopes (resp. gradients). Since Betweenness is typically defined on a preference relation, we first proceed to define Betweenness on a choice correspondence.

Definition 6. Let $c : \mathcal{A} \to \mathcal{A}$ be a choice correspondence. We say c satisfies *Betweenness* over $\mathcal{S} \subseteq \mathcal{A}$ if for any $\{p, q\}, \{p, p^\alpha q\}, \{p^\alpha q, q\} \in \mathcal{S}$ and $\alpha \in (0, 1)$,

1. $c(\{p, q\}) = \{p\} \Rightarrow c(\{p, p^\alpha q\}) = \{p\}$ and $c(\{p^\alpha q, q\}) = \{p^\alpha q\}$,

2. $c(\{p, q\}) = \{p, q\} \Rightarrow c(\{p, p^\alpha q\}) = \{p, p^\alpha q\}$ and $c(\{p^\alpha q, q\}) = \{p^\alpha q, q\}$.

We are ready for the first result when restricting attention to a Marschak-Machina triangle. For any three prizes $\{a, b, c\} \subseteq X$, consider the set of all lotteries induced by them, $\Delta(\{a, b, c\})$. Let $\mathcal{B}_{a,b,c}$ denote the set of all finite and nonempty subsets of $\Delta(\{a, b, c\})$. Note that $\mathcal{B}_{a,b,c} \subseteq \mathcal{A}$. Going forward, we omit subscripts and use the notation $\mathcal{B}$. We begin with a few standard definitions. Let p be a mean-preserving spread of q. We say c is *weakly risk averse* (resp. *risk loving*) *over* $\mathcal{B}$ if $\{p, q\} \in \mathcal{B}$ implies $q \in c(A)$ (resp. $p \in c(A)$). If $c(\{p, q\}) = \{p, q\}$ whenever $\{p, q\} \in \mathcal{B}$, we say c is *risk neutral*. We say that indifference curves fan out (resp. fan in) if they become weakly steeper (resp. flatter) in the first order stochastic dominance direction.

Proposition 4. *Suppose c admits an AREU representation. If c satisfies Transitivity over $\mathcal{B}$, then:*

1. *c satisfies Betweenness over $\mathcal{B}$.*

2. *c is either weakly risk averse over $\mathcal{B}$, weakly risk loving over $\mathcal{B}$, or risk neutral over $\mathcal{B}$.*

3. *Indifference curves fan out if c is weakly risk averse.*

4. *Indifference curves fan in if c is weakly risk loving.*

Proposition 4 connects AREU to linear indifferent curves—a property of standard expected utility. Moreover, Proposition 4 also pins down the set of admissible indifferent curves. Even though AREU allows a decision maker to have varying magnitudes of risk aversion, compliance with Transitivity would bound the increase in risk aversion such that choices are either exclusively risk averse or exclusively risk loving (in this probability triangle). In each case, a particular direction of fanning is also prescribed (Figure 3.1). These results provide testable predictions for AREU, and separates it from other models, which we discuss next.

3.7.2 Comparison to other non-expected utility models

Various alternatives to expected utility were introduced by Quiggin (1982), Chew (1983); Fishburn (1983); Dekel (1986), Bell (1985); Loomes & Sugden (1986), Gul (1991), Kőszegi & Rabin (2007), and

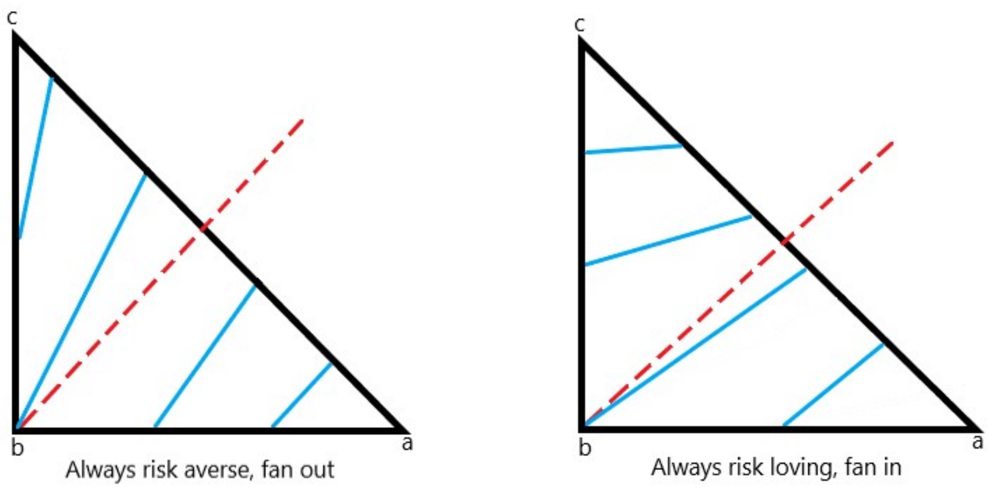

Figure 3.1: Let $a < b < c$. Dotted (red) lines are the mean-preserving spread lines. Solid (blue) lines are indifferent curves. Referring to Proposition 4, the picture on the left corresponds to point 3 and the picture on the right corresponds to point 4.

Cerreia-Vioglio et al. (2015). We now use Proposition 3 and Proposition 4 to study their relationship to AREU.

The AREU model has a close relationship with *betweenness preferences* introduced by Chew (1983); Fishburn (1983); Dekel (1986). Although the two intersect only at expected utility, a direct application of Proposition 3, the two make similar predictions for binary choices when Transitivity is added to AREU in a probability triangle.

Among models of betweenness preferences, Gul (1991)'s *disappointment aversion* is closest in spirit to AREU, but the two predict different behavior. In *disappointment aversion*, the set of possible outcomes of each lottery is decomposed into elevation prizes and disappointment prizes, and the utilities from disappointment prizes are discounted using a function of the probability of disappointment. An implication of disappointment aversion is the property of mixed fanning, in which indifference curves first fan in and then fan out, for example. AREU cannot accommodate mixed fanning, a direct application of Proposition 4, and so the two models differ in their coverage of non-expected utility behavior.

For the same reason, AREU and Cerreia-Vioglio et al. (2015)'s *cautious expected utility* put fourth different behavioral predictions. In their model, a decision maker evaluate each lottery as its worst certainty equivalence under a set of (Bernoulli) utility functions. The result is a behavior that resembles cautiousness. A property resembling mixed fanning is an implication of their model, where indifferent curves are steepest in the middle, a consequence of the axiom *Negative Certainty Independence*: $p \succsim \delta$ implies $p^{\alpha}q \succsim \delta^{\alpha}q$.

Like AREU, Kőszegi & Rabin (2007)'s *reference-dependent risk preferences* uses reference points to explain non-expected utility behavior. However, both the identification of reference points and the consequence of changing reference points differ. In AREU, reference alternatives are given by the safest alternatives in choice problems, and they serve as a proxy for changing risk preferences. In Kőszegi & Rabin (2007), a decision maker is subjected to gain-loss utility relative to a reference point, where the reference point is the lottery she expects to receive. We focus on *choice-acclimating personal equilibrium* (CPE), in which reference points are endogenously set as the eventually-chosen alternatives. Masatlioglu & Raymond (2016) shows that when a CPE specification satisfies first order stochastic dominance, the implied behavior can be explained by the *quadratic utility* functionals of Machina (1982); Chew et al. (1991). Yet, Chew et al. (1991) demonstrates that quadratic functionals intersect with betweenness preferences only at expected utility, and hence the CPE model of Kőszegi & Rabin (2007) intersects with AREU only at expected utility.[23]

The model closest to AREU, to my knowledge, is the *context-dependent gambling effect* by Bleichrodt & Schmidt (2002). In their model, a decision maker's preferences are explained by two (Bernoulli) utility functions, one for comparisons that involve a riskless option and another for the rest. Unlike AREU, their model only applies to binary decisions, which results in different axioms and applicability. Furthermore, when a degenerate lottery is slightly perturbed into a non-degenerate one, it produces a choice reversal, which seems implausible. Their model also does not accommodate violations of expected utility in choice problems without a riskless option, such as variations of the Allais paradox. Finally, while their axioms are separately imposed on binary decisions involving and not involving riskless options, our axioms are imposed on the choice correspondence without such discrimination.

4 Time Preferences

In this section, we provide an application of our unified framework for choices over delayed consumption. The canonical model for this setting is Discounted Utility, axiomatized by Fishburn & Rubinstein (1982), in which a decision maker evaluates each payment-time pair (x, t) by $\delta^t u(x)$. However, Dis-

[23]Similar conclusions of non-intersection with AREU (other than expected utility) can be made for Quiggin (1982)'s *rank dependent utility* (see Chew & Epstein (1989)) and Bell (1985); Loomes & Sugden (1986)'s *disappointment theory.* Some of these results, and a comprehensive summary, are provided by Masatlioglu & Raymond (2016).

counted Utility has routinely failed experimental tests as subjects violate Stationarity: the choice between two payments changes when the decision is made in advance, typically favoring the later option for the long-term decision.[24] To accommodate this violation, we weaken the axioms of Fishburn & Rubinstein (1982) using an approach analogous to Section 2's Reference Dependence. The outcome is a utility representation in which choices maximize exponentially discounted utilities using a discount factor that depends on the timing of the earliest payment.

4.1 Preliminaries

Let $X = [a, b] \subset \mathbb{R}_+$ be an interval of non-negative payments and let $T = [1, \bar{t}] \subset \mathbb{R}_+$ be an interval of non-negative time points. $X \times T$ is the set of alternatives, where $(x, t) \in X \times T$ denotes a payment of x at time t. We endow $X \times T$ with the standard Euclidean metric. Let $\mathcal{A}$ be the set of all finite and nonempty subsets of $X \times T$. Finally, let $c : \mathcal{A} \to \mathcal{A}$, $c(A) \subseteq A$, be a choice correspondence.

We maintain the following standard axioms for time preference, that higher payments and sooner payments are better.

Axiom 6.

1. *Outcome Monotonicity: if* $x > y$, *then* $c(\{(x, t), (y, t)\}) = \{(x, t)\}$.

2. *Impatience: if* $t < s$, *then* $c(\{(x, t), (x, s)\}) = \{(x, t)\}$.

4.2 Reference Dependent Patience

Time consistency in choice is captured by a well-known behavioral property called Stationarity. Under Stationarity, a decision maker's preference between two future payments is consistent regardless of when the decision is made. For this reason, Stationarity is often deemed a normative postulate in economic analysis.

Similar to what we did to WARP and Independence in previous sections, we first define what it means for a choice correspondence c to satisfy Stationarity over a subset of all choice problems $\mathcal{S} \subseteq \mathcal{A}$.

Definition 7. Let $c : \mathcal{A} \to \mathcal{A}$ be a choice correspondence and $\mathcal{S} \subseteq \mathcal{A}$. We say c satisfies *Stationarity* over $\mathcal{S}$ if for all $A, B \in \mathcal{S}$, $a > 0$,

[24] See for example Laibson (1997), Frederick et al. (2002), and Benhabib et al. (2010).

$$(x,t) \in c(A),\ (y,q) \in A,\ (y,q+a) \in c(B), \text{ and } (x,t+a) \in B \Rightarrow (x,t+a) \in c(B).$$

Supplied with Axiom 6, a direct adaption of Fishburn & Rubinstein (1982) into the framework of choice gives that c satisfies WARP and Stationarity over $\mathcal{A}$ if and only if it admits a (exponential) Discounted Utility representation.

A choice correspondence that exhibits time inconsistency fails to satisfy Stationarity over $\mathcal{A}$. However, the choice correspondence may still satisfy Stationary over some subsets of $\mathcal{A}$. Consider the following axiom, which states that Stationarity is satisfied between any two choice problems that share an earliest payment.

Axiom 7 (Reference Dependent Patience (RDP))**.** *For any $A, B \in \mathcal{A}$, if A and B share an earliest payment, then c satisfies WARP and Stationary over $\{A, B\}$.*

The axiom posits that a violation of WARP and Stationarity between two choice problems can only occur if they do not share an earliest payment. If we interpret compliance with Stationarity as having a stable level of patience, the axiom proposes that patience may depend on how soon any payment can be attained. This allows us to capture behavior in which compliance with WARP and Stationarity is not necessarily upheld between long-term and short-term choice problems, such as those exhibited in time consistency experiments.

Note that this postulate can be rewritten in the style of Reference Dependence (Axiom 1) and Risk Reference Dependence (Axiom 3) from previous sections, stated formally in the following lemma.

Lemma 1. *Fix a choice correspondence c, the following are equivalent.*

1. *c satisfies Axiom 7.*
2. *For every choice problem $A \in \mathcal{A}$ and every earliest payment (x,t) in it, c satisfies WARP and Stationarity over $\{B \subseteq A : (x,t) \in B\}$.*

Albeit straightforward, the lemma reassures us that the unified method of weakening standard postulates proposed in this paper is not dissimilar to demanding compliance between pairs of choice problems.

In fact, their equivalence in this setting is due to two details. First, unlike our general model (Section 2) and application in the risk domain (Section 3), in which the reference order is either fully or partly subjective, the reference points in the present setting is completely objective—the earliest payments in the choice sets. Because of this objectivity, the reference order is pinned down axiomatically, and the axiom does not involve an existential statement that allows for subjectivity in determining reference points. Second, WARP and Stationarity are properties between pairs of choices (and not more). This is not the case for all postulates. For example, Transitivity is an axiom that is trivially satisfied between any pair of choices, but a violation can be found when more choices are considered. Identifying this equivalence, and the reasons thereof, allows us to design more efficient tests of the axioms in our unified framework.

4.3 Increasing Patience

We postulate that patience (may) increase when options are postponed.

Consider prizes $x_1 < x_3$ arriving at time $t_1 < t_3$ respectively. We posit that by *postponing* the options by $a > 0$, the decision maker is (weakly) more patient and will choose $(x_3, t_3 + a)$ over $(x_1, t_1 + a)$ if she chose (x_3, t_3) over (x_1, t_1). The postulate differs from Stationarity as it allows for the choice of (x_1, t_1) over (x_3, t_3) but $(x_3, t_3 + a)$ over $(x_1, t_1 + a)$, or *present bias.* To summarize, it allows for violation of Stationarity in one direction but not the other.

However, this falls short of capturing changes in patience. Difference in delay aversion between individuals cannot be directly categorized into difference in discounting and difference in consumption utility, an issue discussed in Ok & Benoît (2007). Just because a decision maker chooses a sooner option, and another a later one, it is not conclusive that the first decision maker discounts more. It could be that there is difference in consumption utility, where the first decision maker's marginal utility for money is a lot lower than that of the second decision maker inducing the choice of a sooner but smaller prize.

We introduce a technique that allows us to "fix" consumption utilities and only allow discounting to change. This can be used to characterize a set of individuals whose consumption utility is (as-if) the same, and differ only in their patience level. For this paper, we use it to restrict the choice behavior of a single individual to those that can be explained by a single consumption utility function while allowing patience to vary.

Consider $c(\{(x, t), (y, q), (z, s)\}) = \{(x, t), (y, q), (z, s)\}$, where (x, t) gives the smallest payment but arrives earliest, (z, s) gives the largest payment but arrives latest, and (y, q) is intermediate in both. Now consider the new choice problem $\{(x, \lambda t), (y, \lambda q), (z, \lambda s)\}$, where $\lambda < 1$; that is, all payments will now arrive at a (common) fraction of time. Under Stationarity, a decision maker only cares about the delay between the alternatives and would now strictly prefer the latest option since the time-difference between any two options is smaller. Yet it is ambiguous how a decision maker of the present model would behave. On one hand, an earlier choice problem causes the decision maker to choose more impatiently; on the other hand, delays between alternatives have decreased, which favor later options. The competing forces render the choice ambiguous. The same competing forces occur when $\lambda > 1$: the decision maker is more patient, but delays between options are larger. In these situations, we restrict the decision maker's behavior in the following way: if the decision maker chooses both the earliest and the latest alternatives after such a transformation (and recall that he was indifferent between all three before), then he also chooses the intermediate option in the new choice problem.

The same restriction is imposed when the transformation is of the form $\{(x, \lambda t - a), (y, \lambda q - a), (z, \lambda s - a)\}$. Note that only λ changes the delay between the options, and a symmetrically shifts arrival time. Also, the only instance this restriction is non-trivial is when the aforementioned competing forces are present: when "$\lambda < 1$ and $\lambda t - a < t$", where the decision maker becomes less patient but she no longer has to wait as long for a better payment, and when "$\lambda > 1$ and $\lambda t - a > t$", where she is more patient but also has to wait longer for a better payment.

This gives rise to the following axiom.

Axiom 8 (Increasing Patience)**.** *For all $t_1 < t_2 < t_3$, $A = \{(x_1, t_1), (x_2, t_2), (x_3, t_3)\}$, and $A' = \{(x_1, \lambda t_1 + a), (x_2, \lambda t_2 + a), (x_3, \lambda t_3 + a)\}$,*

1. $c(\{(x_1, t_1), (x_3, t_3)\}) = \{(x_3, t_3)\} \Rightarrow c(\{(x_1, t_1 + a), (x_3, t_3 + a)\}) = \{(x_3, t_3 + a)\}$ *for all* $a > 0$,
2. $c(A) = A$ *and* $(x_1, \lambda t_1 + a), (x_3, \lambda t_3 + a) \in c(A') \Rightarrow (x_2, \lambda t_2 + a) \in c(A')$ *for all* $\lambda, a \in \mathbb{R}$.

This postulate is trivially satisfied by a decision maker whose behavior fully complies with Stationarity, since she can neither be more patient nor less patient when options are symmetrically postponed.

4.4 Representation Theorem

We are ready for the utility representation and representation theorem.

Definition 8. c admits a Present-Biased Discounted Utility representation (PBDU) if there exist a strictly increasing and continuous utility function $u : X \to \mathbb{R}$ and a set of time-indexed discount factors $\{\delta_t\}_{t \in T}$ such that

$$c(A) = \arg\max_{(x,t) \in A} \delta_{r(A)}^{t} u(x),$$

where

- $r(A) = \min\{t : (x,t) \in A\}$,
- $t < t' \Rightarrow \delta_t \leq \delta_{t'}$,
- δ_t is continuous on $[1, \bar{t}]$.

Proposition 5. *Let $c : \mathcal{A} \to \mathcal{A}$ be a choice correspondence. The following are equivalent:*

1. *c satisfies Reference Dependent Patience, Increasing Patience, Outcome Monotonicity, Impatience, and Continuity.*
2. *c admits a PBDU representation.*

Furthermore, in every PBDU representation, discount factors δ_t are unique given u.

In this model, it is *as if* the decision maker maximizes exponentially discounted utility, but with discount factors that depend on the timing of the earliest available payment. When the earliest available payment arrives sooner in one choice problem than another, then the decision maker uses a lower discount factor in the former. Since discount factors are often interpreted as a measure of *patience*, our model can be viewed as one in which the decision maker's patience changes systematically across choice problems, where she is less patient when an earlier payment is available.

This model conforms with present bias, the empirically prevalent failure of dynamic consistency in which decision makers exhibit less delay aversion for long-term decisions. Take for example the classic observation of present bias, where \$$x$ today is preferred to \$$y$ tomorrow (choice problem A) but the opposite decision is made when both payments are postponed by a year (choice problem B). The model we propose explains the behavior with the simple interpretation that, since the earliest alternative for A arrives sooner than that for B (i.e., $r(A) < r(B)$), the decision maker is less patient in the former (i.e., $\delta_{r(A)} < \delta_{r(B)}$).

Moreover, even though the choice between \$$x$ at time t and \$$y$ a day later is not consistent across the time horizon t, the model predicts that as we gradually postpone both options with s, the choice can only switch from $(x, t+s)$ to $(y, t+1+s)$. That is, if there is a point in time at which the decision maker becomes sufficiently patient to choose \$$y$ over \$$x$, she must continue to do so as we further postpone both options. This "single switching" property is closely connected to the ordered-reference nature of PBDU and provides testable predictions.

Finally, the underpinning of PBDU is the simultaneous weakening of WARP and Stationarity in a reference-dependent approach. Reminiscent of the observation made in Proposition 3, WARP and Stationarity are interconnected in our model: neither of which can be independently weakened, formally stated in the following result.

Proposition 6. *Suppose c admits a PBDU representation. Then the following are equivalent:*

1. *c satisfies WARP (over $\mathcal{A}$).*
2. *c satisfies Stationarity (over $\mathcal{A}$).*
3. *c admits an exponential discounting utility representation.*
4. *c admits a utility representation.*

4.5 Related models of time preferences

The biggest difference between Present-Biased Discounted Utility (PBDU) and *hyperbolic discounting,* a class of models in which future options are discounted disproportionately less, is that PBDU (when non-trivial) necessitates WARP violations and hyperbolic discounting models satisfy WARP. Furthermore, unlike models of hyperbolic discounting, PBDU evaluates all alternatives in a choice problem using a single discount factor.[25] However, the empirically informed intuition that discount factors vary across time is shared between models of hyperbolic discounting and PBDU, albeit implemented differently. For our model, PBDU, discount rate changes at the choice problem level, whereas for hyperbolic discounting it changes at the alternative level. The difference is stark when we consider choice problems that contain more than two alternatives. In hyperbolic discounting, the preference between any two options stays the same regardless of what choice problems they appear in, hence WARP is never violated. This is not the case for PBDU, where a sooner option may become superior to a later one from the introduction of a third (but not necessarily chosen) alternative, and results in WARP violations in PBDU.

Exponential discounting has advantageous properties in economic applications, propelling Laibson (1997)'s well-known *quasi-hyperbolic discounting.* In their model, behavior complies with Stationarity as long as the choice is between two future payments, and present bias only arises when an immediate payment is involved. This is not the case in PBDU, as the switch from choosing the earlier payment

[25] See for instance Loewenstein & Prelec (1992) and Laibson (1997).

to choosing the later one can occur at any time as we gradually shift both payments into the future. Another implication of quasi-hyperbolic discounting in our setting is the failure of continuity, where an instantaneous change in choice occurs when the earlier payment arrives at time 1 ("today"). Our model complies with continuity of choice, and instead forgoes WARP to explain dynamic inconsistency.

We now turn to two other models that both explain dynamic inconsistency and can explain WARP violations.

Lipman et al. (2013) provides an explanation of dynamic inconsistency that builds on Gul & Pesendorfer (2001)'s introduction of *temptation*. In Gul & Pesendorfer (2001), a decision maker has commitment utilities and temptation utilities, and chooses a menu (a choice problem) taking into account both. The result is that a larger menu may be inferior, a departure from the conventional understanding that more options should never be worse. Lipman et al. (2013) extends this to the setting of time preference and proposes that a decision maker assess current consumption using temptation utility and future consumption using commitment utility. When making decisions in advance, it is as if the decision maker is choosing between singleton menus for her future self, and the absence of temptation utility allows her to make a more patient decision relative to choices over immediate consumption. Like *quasi-hyperbolic discounting*, present-bias is restricted to immediate consumption, whereas PBDU allows present-bias to kick in at time frame and as long as its effect is persistent when both options are further postponed—the "single switching" property discussed earlier.

More recently, Freeman (2016) introduced a framework in which WARP is weakened, and reversals are explained by time-inconsistent preferences. In their model, a decision maker chooses when to complete a task, and may exhibit choice reversal when additional opportunities for completions are introduced (a expansion of the choice set). In particular, in response to the addition of an opportunity for completion, a sophisticated decision maker may choose to complete the task earlier (and never later) in anticipation that allowing her future self to make that decision would result in an eventual completion time that is worse than completing the task now. A naive decision maker, however, could only end up completing later.

While our model, PBDU, allows for choice reversal in the direction of choosing an earlier option when the choice set expands, it is in fact incompatible with Freeman (2016)'s decision makers' behavior (other than WARP-conforming behavior). In PBDU, a reversal can only occur when the discount rate changes, which only happens when an alternative earlier than any other already available is added. However, in Freeman (2016)'s model, either that this added alternative is chosen (which is not a reversal) or the decision problem becomes identical to before and so WARP is complied with. Indeed, the necessary conditions of their model, Irrelevant Alternatives Delay (for a naive agent) and Irrelevant Alternatives Expedite (for a sophisticated agent) only hold in PBDU if WARP holds.

5 Social Preference

We now turn to our last application.

Consider a decision maker who has a particular type of set-dependent social preference—she shares more when greater equity is attainable. Experiments in economics and psychology have shown that, instead of being fully selfish and maximize monetary payment to oneself, people are often willing to share their wealth. This leads to models of *other-regarding preferences* and *inequality aversion,* first introduced by Fehr & Schmidt (1999); Bolton & Ockenfels (2000); Charness & Rabin (2002). Furthermore, one's desire to share, or inequality aversion, may be affected by the options they have, often in the direction where the availability of more equitable options results in greater sharing. One explanation for this behavior is outcome-based, where a decision maker becomes more inequality averse in the presence of more equitable distributions. Another explanation is intention-based, where the decision maker seeks to be perceived as fair.[26] Our model does not distinguish between these two causes for increased altruism, we refer interested readers to surveys by Fehr & Schmidt (2006); Kagel & Roth (2016) for the vast evidence and suggested explanations.

To illustrate, suppose a decision maker is endowed with \$10 and is given a number options to share it with another individual. When she is asked to choose between giving \$2 and giving \$3, giving \$2 may seem reasonable. However, when the choice is between giving \$2, \$3 or \$5, the decision maker may opt for \$3 (and keeping \$7) instead. The pair of choices (over income distributions) $c\left(\left\{\left(\$8,\$2\right),\left(\$7,\$3\right)\right\}\right) = \left\{\left(\$8,\$2\right)\right\}$ and $c\left(\left\{\left(\$8,\$2\right),\left(\$7,\$3\right),\left(\$5,\$5\right)\right\}\right) = \left\{\left(\$7,\$3\right)\right\}$ violates WARP. Hence the assumption of utility maximization, even if the utility function captures other-regarding preferences and inequality aversion, is incapable of explaining this behavior.

Using our unified framework, where both WARP and a standard postulate are weakened using reference-dependence, we provide a model in which a decision maker's degree of inequality aversion increases when equitable options are added to a choice problem.

[26] See for example Ainslie (1992), Rabin (1993), Nelson (2002), and Sutter (2007).

5.1 Preliminaries

Let $X = \mathbb{R}_+ \times \mathbb{R}_+$ be the set of all pairs of non-negative monetary payments. We call a pair $(x, y) \in X$ an income distribution, where x is the dollar amount for the decision maker and y for a second individual. We endow X with the standard Euclidean metric. Let $\mathcal{A}$ be the set of all finite and nonempty subsets of X and $c : \mathcal{A} \to \mathcal{A}$, $c(A) \subset A$ a choice correspondence.

The first axiom is standard, an income distribution that gives everyone weakly more, and at least one person strictly more, is strictly preferred.

Axiom 9 (Monotonicity). *$c(\{(x, y), (x', y')\}) = \{(x, y)\}$ whenever $x \geq x'$, $y \geq y'$, and $(x, y) \neq (x', y')$.*

5.2 Fairness Dependence

Our first axiom for this section is a specialization of Axiom 1 from Section 2. It characterizes behavior in which choices from choice problems containing the same amount of attainable equity conform with quasi-linear preferences. The use of quasi-linear preferences for choices involving money is common in the economics. Since our model introduces reference-dependent utility functions, using quasi-linear utilities when preferences are stable provides meaningful restrictions to choices.

Definition 9. Let $c : \mathcal{A} \to \mathcal{A}$ be a choice correspondence and $\mathcal{S} \subseteq \mathcal{A}$. We say c satisfies *quasi-linearity* over $\mathcal{S}$ if for all $A, B \in \mathcal{S}$ and $a \in \mathbb{R} \setminus \{0\}$,

$$(x, y) \in c(A), \ (x', y') \in A, \ (x' + a, y') \in c(B), \text{ and } (x + a, y) \in B \Rightarrow (x + a, y) \in c(B).$$

In order to characterize attainable equity, we first need a measure of equity. A nature candidate is the ratio between x, y for any income distribution (x, y). We define the equity index of (x, y) as $e_{(x,y)} := \min \left\{ \frac{x}{y}, \frac{y}{x} \right\}$. The use of min is to treat income distributions (a, b) and (b, a) indiscriminately in terms of measuring equity. Note that $e_{(x,y)}$ is always weakly less than 1, and values closer to 1 correspond to greater equity. The index captures how "close" are the two payments within an income distribution (a, b), adjusting for scale, and is ordinally equivalent to the form taken by the Gini coefficient $|a - b|/a$. The cardinality of this index plays no role in our analysis.

Analogous to our approach in previous sections, we demand that choices comply with WARP and quasi-linearity when the fairest income distribution of a choice problem is unchanged. Departing from complete compliance with WARP and quasi-linearity, we allow the decision maker to choose different income distributions when the fairest income distribution is dropped, potentially violating WARP. Formally, we impose a weakening of WARP and quasi-linearity in the following way:

Definition 10. For any set of income distributions $A \in \mathcal{A}$, we call $\Psi(A) := \left\{(x,y) \in A : e_{(x,y)} \geq e_{(x',y')} \, \forall (x',y') \in A\right\}$ the set of *fairest* income distributions in A.

Axiom 10 (Fairness Dependence (FD))**.** *For every choice problem $A \in \mathcal{A}$ and any fairest distribution $(x,y) \in \Psi(A)$, c satisfies WARP and quasi-linearity over $\{B \subseteq A : (x,y) \in B\}$.*

5.3 Increasing Altruism

We study choices that exhibit increased sharing when greater equity is attainable. Consider the following postulate. Suppose in a choice problem an income distribution (x,y) is chosen over $(x',0)$, a distribution where the decision maker does not give at all. We postulate that by adding any other income distributions into the choice set, since this can only (weakly) increase attainable equity, she does not switch to not sharing, $(x',0)$. Additionally, we extend this postulate to cases where the comparison is between (x,y) and (x',y') such that $y' < y$. Effectively, this restriction imposes a direction on which willingness to share changes—the decision maker is weakly more altruistic when more options are available, which weakly increases attainable equity. Formally:

Axiom 11 (Increasing Altruism)**.** *For any $A, B \in \mathcal{A}$ such that $A \subset B$ and $(x,y), (x',y') \in A$ such that $y > y'$. If $(x,y) \in c(A)$ and $(x',y') \notin c(A)$, then $(x',y') \notin c(B)$.*

5.4 Representation Theorem

Consider the following utility representation in which utility from receiving the amount \$$x$ is always evaluated consistently but utility from giving amount \$$y$ depends on how much equity is attainable from the choice problem.

Definition 11. c admits a Fairness-based Social Preference Utility representation (FSPU) if there exists a set of strictly increasing utility functions $\{v_r : \mathbb{R}_+ \to \mathbb{R}\}_{r \leq 1}$ such that

$$c(A) = \underset{(x,y) \in A}{\arg\max}\, x + v_{r(A)}(y),$$

where

- $r(A) = \max\limits_{(x,y) \in A} e_{(x,y)}$,

- $r > r' \Rightarrow v_r(y) - v_r(y') \geq v_{r'}(y) - v_{r'}(y')$ for all $y > y'$,
- $\arg\max_{(x,y)\in A}\ x + v_{r(A)}(y)$ has a closed-graph.

Proposition 7. *Let $c : \mathcal{A} \to \mathcal{A}$ be a choice correspondence. The following are equivalent:*

1. *c satisfies Fairness Dependence, Increasing Altruism, Monotonicity, and Continuity.*
2. *c admits a FSPU representation.*

Furthermore, in every FSPU representation, v_r is unique for all r.

In this model, the decision maker's utility from giving dollar amount y, $v_{r(A)}(y)$, depends on how much equity is attainable in the underlying choice problem, as measured by $r(A)$. Recall that $e_{(x,y)}$ is weakly less than 1 and a number closer to 1 represents greater equity. Hence, attainable equity from choice set A is simply the highest value $e_{(x,y)}$ among available income distributions $(x, y) \in A$, or $r(A) = \max\left\{e_{(x,y)} : (x, y) \in A\right\}$. When $r(A)$ is greater, the decision maker values any given shared amount y more. Consequently, even if income distribution (x, y) is chosen over (x', y') in some choice problem, where $y' > y$, adding a very fair option could cause the decision maker to switch to (x', y').

This model accommodates increased willingness to give when distributing a fixed pie with different splitting options. To illustrate, suppose a decision maker must allocate a fixed amount of money, say \$100, between her and another individual, but she is not allowed to split the amount however she likes. Instead, there is a set of feasible distributions characterized by $D \subset [0, 1]$; she can choose to allocate $\alpha \cdot \$100$ to herself if and only if $\alpha \in D$. By specifying two different sets of feasible distributions, D and D', we have effectively specified two choice problems in our setup. Say $D = \{0.5, 0.6, 0.7\}$ and $D' = \{0.6, 0.7\}$. If $\alpha = 0.7$ is chosen in D' (the decision maker keeps \$70 for herself and \$30 is given to the other individual), she might choose to keep less in D due to increased altruism from greater attainable equity. However, if she chose $\alpha = 0.6$ in D', then she must not choose $\alpha = 0.7$ in D; this is a testable prediction.

In FSPU, altruism is maximal when a perfectly balanced income distribution is available. In particular, note that increased altruism is not the result of opportunity to *give* more; instead, it is attainable equity that drives altruism. To illustrate the difference, consider the same example but with $D = \{0.5, 0.3, 0.2\}$ and $D' = \{0.3, 0.2\}$. Even though D contains alternatives that achieve greater equity, the decision maker's ability to give is the same across the two choice problems. Yet, since the feasible allocations are always unfavorable to her (she can never keep more than half), higher attainable equity results from her ability to *take* more. In this setting, our decision maker can be interpreted as being less altruistic when the world is unfair to her, and she becomes more altruistic when fairer options are added.

Lastly, FSPU allows for willingness to forgo a greater surplus in favor of giving more. Suppose the decision maker must choose between $(\$30, \$20)$ and $(\$60, \$0)$. The second option is appealing in that the total amount of money extracted is greater, whereas the first option sacrifices both surplus and payment to oneself in favor of providing a share to the other individual. Suppose $(\$60, \$0)$ is chosen. The model allows for the behavior in which the addition of $(\$25, \$25)$ to the choice set causes the decision maker to switch from $(\$60, \$0)$ to $(\$30, \$20)$ due to increased generosity. While this behavior is reasonable, it cannot be accommodated by any model that complies with WARP.

The familiar linkage between WARP violation and violation of standard postulate, in this case quasi-linearity, is summarized in the following statement.

Proposition 8. *Suppose c admits a PBDU representation. Then the following are equivalent:*

1. *c satisfies WARP (over $\mathcal{A}$).*
2. *c satisfies quasi-linearity (over $\mathcal{A}$).*
3. *c admits a quasi-linear utility representation.*
4. *c admits a utility representation.*

5.5 Related Literature

Other-regarding preferences have been extensively studied, and well-known models are introduced by Fehr & Schmidt (1999); Bolton & Ockenfels (2000). However, the primary focus of these models is to capture the characterization of *inequality aversion* using functional forms. In particular, a single and persistent preference ranking of income distributions is assumed throughout these models. Charness & Rabin (2002) introduced a departure that allows for *reciprocity* using a term that lowers utility from giving when the other player is deemed to have “misbehaved”.

FSPU, departs from these models by introducing preferences over income distributions that may change from one choice problem to another. In particular, utility from giving depends on how much equity is attainable in the underlying choice set. The vast literature on distributional preferences provides suggestive evidence of this behavior. List (2007); Bardsley (2008); Korenok et al. (2014)

showed that in a dictator game, adding (or increasing) the option to take from the receiver significantly reduces a dictator's willingness to give, and in some cases result in choice reversals (WARP violations). However, although the narratives are related, the design of their experiments does not provide a complete test for the predictions of FSPU, as additions of less equitable distributions do not affect preferences in FSPU.

The study of *audience effect* also provides empirical evidence that decision makers care about how others perceive their choices. In Dana et al. (2006), dictators were given the option exit (avoid) a $10 dictator game and receive $9, a option that leaves the receiver with nothing. Since a payoff of $9 (and $10) can be achieved by going through with the dictator game, exiting is interpreted as a costly effort to avoid the dictator game. 28% of the subjects chose to exit. When the game is conducted such that the decision to exit or not is completely veiled from the receivers, only 4% chose to exit.

In a separate study, Dana et al. (2007) provides dictators a costless opportunity to find out how much the receivers will receive from each of their two options, $(6, 1)$ and $(5, 5)$, before making a choice (payoffs to themselves, the first number in each pair, are always displayed). 44% of dictators chose not to find out, and among them 86% chose "$(6, ?)$" over "$(5, ?)$". Only 47% of dictators chose to reveal the payoffs and subsequently chose $(5, 5)$ over $(6, 6)$. In the baseline, in which all payoffs are displayed by default, 74% of subject chose $(5, 5)$ over $(6, 6)$. Based on subjects' apparent exploitation of this "moral wiggle room", the authors conclude that fair behavior is primarily motivated by the desire to appear fair, either to themselves or to others.

In game theoretic settings, Rabin (1993)'s pioneering work introduced intention based reciprocity through a notion of *kindness*. In their model, kindness is measured using the set of payoffs an opponent *could* induce. A player's kindness depends on how kind the opponent is, due to the desire to be *fair*, and vice versa, leading to the solution concept term *fairness equilibrium*. Since kindness is measured using the set of available actions, the Rabin (1993)'s model and FSPU share some conceptual similarity. However, since FSPU is built on a decision theoretic framework, it is unable to capture the type of reciprocity concerns depicted in Rabin (1993). The same argument separates FSPU from related models in game theory.

To my knowledge, Cox et al. (2016) is the only other paper, with a decision theoretic setup, that introduces a model to explain WARP violations of this kind. Unlike FSPU, they take *endowment* into account, which allows for the study of *giving* versus *taking*. This is different to the approach in FSPU, where only income distributions are relevant and endowments are not part of the primitive. Based on an intuition related to FSPU, Cox et al. (2016) uses *moral reference points* to explain changes in dictator's willingness to allocate, where a moral reference point more favorable to the dictator (and/or less favorable to the receiver) results in allocating more to the dictator herself. However, unlike FSPU, their reference points are not alternatives, but instead a vector of reference payoffs that depend on multiple allocations within the feasible set as well as the endowment. Consequently, there are many choice problems in which the addition of a more equitable alternative cannot result in choice reversal in their model, since it does not affect the moral reference point, yet preference reversals as a result of adding more equitable alternatives is precisely the behavioral tenet in FSPU.[27] Although the two models are different in many ways, they both seek to capture the increasingly evident intuition that social preference depends on the set of feasible allocations, which results in WARP-violating behavior.

6 Conclusion

This paper presents a unified framework for ordered-reference dependence choice. The framework, a reference-oriented weakening of standard postulates, is adaptable to suitably accommodate a wide range of reference orders and reference effects. We demonstrate this universality by providing (axiomatized) utility representations in the context of risk, time, and social preference, where we use reference dependent preferences to account for well-known behavioral anomalies. The resulting models are akin to their standard counterparts, inheriting many of the standard models' properties while explaining non-conforming behavior through intuitive changes in specifications. This is possible primarily due to the use of choice correspondences as primitive (instead of preference relations) along with the weakening of WARP. Together, they allow us to prescribe a new way of weakening standard postulates using behavior in non-binary choice problems.

A natural question is the generality of this exercise—does every choice theoretic model have an ordered-reference dependence version by simply having their axioms weakened using an adapted version of Axiom 1? We are able to provide some answers to this question.

In Appendix A, we provide a sufficient condition for an arbitrary standard postulate to be accommodated by our method. We call these standard postulates *finite properties*, they are axioms that are satisfied whenever a violation fails to be substantiated with just finitely many observations. For example, WARP is a finite property, since it is inherently a property between a pair (hence "finite")

[27]For example, if a choice problem contains income distributions $(0, 1)$ and $(1, 0)$, then adding (x, x) for any $x \in (0, 1)$ will not change Cox et al. (2016)'s moral reference point, and their model demands compliance with WARP.

of choices. To put it differently, if a choice correspondence fails WARP, a violation can be substantiated with just two observations. A non-example is continuity, since a choice correspondence can fail continuity whilst a violation can never be substantiated with finitely many observations. However, at this level of generality, we are only able to achieve a result for ordered-reference dependent choice, and not for ordered-reference dependent utility representations. We formalize and discuss this limitation in Appendix A.

Appendix A: Unified Framework and Finite Properties

In this technical section, we state a companion result to Proposition 1 that allows for (i) either fully or partially prescribing the reference order R and (ii) expanding the accommodated property from WARP to a much larger class. For the latter, we call them *finite properties*, which we will now define.

Let X be an arbitrary set of alternatives, $\mathcal{A}$ the set of all finite and nonempty elements of 2^X. For any $\mathcal{B} \subseteq \mathcal{A}$, we call $c : \mathcal{B} \to \mathcal{A}$, where $c(A) \subseteq A$ for all $A \in \mathcal{B}$, a choice correspondence. Let $\mathcal{C}$ be the set of all choice correspondences one can possibly observe from X and $\mathcal{A}$. Formally,

$$\mathcal{C} := \{c : \mathcal{B} \to \mathcal{A} \text{ s.t. } \mathcal{B} \subseteq \mathcal{A}\}.$$

A property imposed on a choice correspondence can be viewed as a subset of $\mathcal{C}$ that is itself closed under subset operations (where each choice correspondence, a member of $\mathcal{C}$, is viewed as a set of pairs). For instance, the set of all choice correspondences satisfying WARP form a collection of choice correspondences defined by the WARP property. We use this notation to characterize an arbitrary property, formally:

Definition 12. We call $\mathcal{T} \subseteq \mathcal{C}$ a *property* if for all $c, \hat{c} \in \mathcal{C}$ such that $\hat{c} \subset c$, $c \in \mathcal{T}$ implies $\hat{c} \in \mathcal{T}$.

We use "c satisfies $\mathcal{T}$" and "$c \in \mathcal{T}$" interchangeably.

In decision theoretic terms, what we call properties here are features of a choice correspondence that are more likely satisfied when we have less observations (i.e. instead of observing c, we only observe $\hat{c}$). For example, WARP ($A \subset B$ and $c(A) \cap B \neq \emptyset \Rightarrow c(A) \cap B = c(B)$) is a property defined on a pair of a choice sets and their correspondencing choices. If the statement of WARP is satisfied for some $c : \mathcal{B} \to \mathcal{A}$, that is, all pairs of choice sets and their corresponding choices satisfy WARP, and $\hat{c} : \mathcal{B}' \to \mathcal{A}$ is where $\mathcal{B}' \subset \mathcal{B}$ and $\hat{c}(B) = c(B)$, then the statement of WARP is also satisfied for $\hat{c}$.

Fact. *The intersection of properties is a property.*[28]

Now, we consider a subset of all properties:

Definition 13. Let $\mathcal{T}$ be a property. We call $\mathcal{T}$ a *finite property* if for all $c \in \mathcal{C}$, $c \notin \mathcal{T}$ if and only if there exists a finite set of choice sets $A_1, ..., A_n \in \operatorname{dom}(c)$ such that $\hat{c} : \{A_1, ..., A_n\} \rightarrow \mathcal{A}$, where $\hat{c}(B) = c(B)$, is not in $\mathcal{T}$.

In words, a finite property is (defined as) a property in which non-compliance can be concluded with finitely many observations (i.e. choices from finitely many choice sets). The majority of decision theoretic axioms are finite properties.

Fact. *When X is finite, any property is a finite property.*[29]

When X is infinite, examples of finite properties include Convexity (either $a^\alpha b \in c(\{a^\alpha b, a\})$ or $a^\alpha b \in c(\{a^\alpha b, b\})$), Monotonicity ($c(\{a, b\}) = \{a\}$ if $a > b$), Transitivity ($a \in c(\{a, b\})$ and $b \in c(\{b, d\})$ implies $a \in c(\{a, d\})$), von Neumann-Morgenstern (vNM) Independence ($p \in c(\{p, q\})$ if and only if $\alpha p + (1 - \alpha) r \in c(\{\alpha p + (1 - \alpha) r, \alpha q + (1 - \alpha) r\})$), Betweenness, Stationarity, and Separability, to name a few.

Non-examples of finite properties (that are nonetheless properties) include various versions of continuity (e.g., $x_n \in c(A_n)$, $x_n \rightarrow x$, $A_n \rightarrow A$ implies $x \in c(A)$) and infinite acyclicity ($a_i \in c(\{a_i, a_{i+1}\})$ for $i = 1, 2, ..., \sigma$, where σ is an ordinal number, implies $a_1 \in c(\{a_1, a_\sigma\})$). Usually, the determination of whether a property is a finite property is immediate when a property is defined algorithmically (as in the axioms in this paragraph) as opposed to defined as an arbitrary subset of $\mathcal{C}$.[30]

Fact. *The intersection of finite properties is a finite property.*[31]

[28] To see this: Consider any $c \in \mathcal{C}$ such that $c \in \mathcal{T}_1 \cap \mathcal{T}_2$. So $c \in \mathcal{T}_1, \mathcal{T}_2$. And since $\mathcal{T}_1, \mathcal{T}_2$ are properties, we have $\hat{c} \in \mathcal{T}_1, \mathcal{T}_2$, and hence $\hat{c} \in \mathcal{T}_1 \cap \mathcal{T}_2$, for all $\hat{c} \in \mathcal{C}$ and $\hat{c} \subset c$.

[29] To see this: Fix any c. Sufficiency is a direct result of the definition of a property. Necessity is also straightforward: Let $A_1, ..., A_n = \operatorname{dom}(c)$, then $\hat{c} = c$, so $c \notin \mathcal{T}$ completes the proof.

[30] The empirical falsifiability of a property (that with finitely many observations the property can be falsified) is not sufficient to establish that it is a finite property. Consider the combination of WARP and continuity, there is no reason why this cannot be defined as a single property. It is empirically falsifiable, since WARP needs only two observations to falsify. Yet in the absence of a violation of WARP, a choice correspondence can very well violate the continuity portion, rendering the property unsatisfied but not falsified with finitely many observations. Conversely, if a property is empirically non-falsifiable, then it is a finite property if and only if it is always trivially satisfied.

[31] To see this: Suppose $\mathcal{T}_1$ and $\mathcal{T}_2$ are both finite properties, $\mathcal{T}_1 \cap \mathcal{T}_2$ is a property. We check Definition 13 that $\mathcal{T}_1 \cap \mathcal{T}_2$ is a finite property. Fix any $c \in \mathcal{C}$. Suppose $c \notin \mathcal{T}_1 \cap \mathcal{T}_2$. Then without loss of generality say $c \notin \mathcal{T}_1$, take the choice sets $A_1, ..., A_n \in \operatorname{dom}(c)$ such that $\hat{c} : \{A_1, ..., A_n\} \rightarrow \mathcal{A}$, where $\hat{c}(B) = c(B)$. Since $\hat{c} \notin \mathcal{T}_1$, so $\hat{c} \notin \mathcal{T}_1 \cap \mathcal{T}_2$, and the rest is straightfoward. Now suppose there exist $A_1, ..., A_n \in \operatorname{dom}(c)$ such that $\hat{c} : \{A_1, ..., A_n\} \rightarrow \mathcal{A}$, where $\hat{c}(B) = c(B)$, is not in $\mathcal{T}_1 \cap \mathcal{T}_2$. Without loss of generality say $\hat{c} \notin \mathcal{T}_1$, so $c \notin \mathcal{T}_1$, so $c \notin \mathcal{T}_1 \cap \mathcal{T}_2$.

For instance, let $\mathcal{T}_1$ be the subset of all choice correspondences that satisfy WARP and $\mathcal{T}_2$ the subset of all choice correspondences that satisfy vNM Independence. These are both finite properties. We can define "WARP and vNM Independence" as a single finite property $\mathcal{T}_1 \cap \mathcal{T}_2$. It characterizes the set of all choice correspondences that satisfies *both* WARP and Independence.

Fact. *The intersection of finite properties and properties that are not finite properties may or may not be finite properties.*[32]

Let $\Psi : \mathcal{A} \to \mathcal{A}$ be a correspondence with $\Psi(A) \subseteq A$ such that $a \in B \subset A$ and $a \in \Psi(A)$ implies $a \in \Psi(B)$.

Definition. We say that a linear order (R, X) is Ψ-*consistent* if $y \in A \backslash \Psi(A)$ implies xRy for some $x \in \Psi(A)$.

Lemma 2. *Consider a choice correspondence $c : \mathcal{A} \to \mathcal{A}$, a finite property $\mathcal{T}$, and a correspondence Ψ. The following are equivalent:*

1. *For every finite $A \in \mathcal{A}$, there exists $x \in \Psi(A)$ such that the choice correspondence $\tilde{c} : \{B : B \subset A \text{ and } x \in B\} \to \mathcal{A}$, where $\tilde{c}(B) = c(B)$, is in $\mathcal{T}$.*

2. *There exists a complete, transitive, antisymmetric, and Ψ-consistent binary relation (R, X) such that for all $x \in X$, the choice correspondence $\tilde{c} : \left\{B : \arg\max_{y \in B} R = x\right\} \to \mathcal{A}$, where $\tilde{c}(B) = c(B)$, is in $\mathcal{T}$.*

First, consider the case in which $\Psi(A) := \mathrm{id}(A) = A$. The first condition in Lemma 2 is satisfied when, for each choice problem, an alternative serves as an anchor that guarantees compliance with finite property $\mathcal{T}$ among choices from subsets of the choice problem containing this anchor. Like before, this anchor is a potential reference alternative with which desirable properties of c hold. When Ψ is not the identity function, we are demanding that at least one alternative in a restricted set of each choice problem (restricted according to Ψ) is a potential reference alternative.

This lemma is the backbone of the models in Section 3, Section 4, and Section 5. For now, we present a simple demonstration. Consider again the wine example, but now the set of all alternatives X contains multiple entries of the same wine at different prices. Each alternative is hence a wine-price

[32] We provide examples. Take $X = [0, 1]$. The intersection of WARP and Continuity is clearly not a finite property, since WARP can hold whereas Continuity will (trivially) hold for any set of choices from finitely many choice sets, but fails to hold in general. The intersection of Monotonicity (that $x > y \Leftrightarrow y \notin c(A)$ for all $A \ni x, y$) and Continuity, on the other hand, is a finite property; essentially, Monotonicity is so strong that Continuity holds whenever Monotonicity does, and since Monotonicity is a finite property (in fact, it is one where a violation can be detected with just the choice from one choice set), their intersection is a finite property.

pair (x, p). Like before, a decision maker was seen choosing a more expensive wine over a cheaper one, but sometimes the reverse (at the exact same prices). The economist postulates that for each choice problem, it is either the cheapest or the most expensive wine that the consumer's underlying preference depends on. Given this postulate, let $\Psi(A)$ be the set of cheapest and most expensive wine-price pairs in A. Furthermore, in addition to WARP, the economist would like to postulate that for a fixed reference, if the decision maker chooses wine x at price p over wine y at price q, then he would also choose wine x at price p over wine y at price $q' > q$; we will call this property "Money is Good". This is an example of a finite property on c.

Lemma 2 establishes that, for a choice correspondence that satisfies these postulates, a reference order (R, X) can be built such that WARP and Money is Good are satisfied among choice sets that share the same $R-$maximal element. Furthermore, for any three wine-price pairs, the intermediate-in-price option is either reference dominated by the more expensive option, the cheaper option, or both. A prediction follows: If a wine-price pair (x, p) reference dominates another wine-price pair (y, q), then all wine-price pairs (z, s) such that $s \in [\min\{p, q\}, \max\{p, q\}]$ are reference dominated by (x, p). That is, even if the economist hasn't fully pinned down this partially subjective R, she can conclude that among choice sets that contain (x, p) and (z, s), where s is between p and q, choices satisfy WARP and Money is Good.

If instead the economist makes the weaker postulate that some reference alternative exists (i.e., $\Psi = id$), then no structure on R can be guaranteed (other than it is a linear order). Conversely, if the economist makes the stronger postulate that the cheapest wine is exactly the reference alternative, then for any two wine-price pairs, the cheaper option reference dominates the other. This demonstrates the flexibility Ψ provides in the trade-off between explanation and prediction. If $\Psi(A)$ is a very restrictive set, such as a singleton, then the model is easy to test and provides strong predictions. If $\Psi(A)$ is very nonrestrictive, such as $\Psi(A) = A$, then the model is harder to test but accommodates more behavior.

To summarize, we expanded the result of 1 to include (i) how properties of R can be axiomatically introduced and (ii) what kind of properties, beyond WARP, of a choice correspondence, can be accommodated in this unified framework of ordered-reference dependent choice. Both are crucial for our specialized models.

Lemma 2 falls short of achieving a utility representation. Indeed, the underlying difficulty is related to the literature on limited datasets, in which one observes choices from a strict subset of all choice problems. de Clippel & Rozen (2014) points out that, in this case, even if observed choices are consistent with behavioral postulates, it need not be sufficient for a corresponding utility representation. In our case, even though we started with an exhaustive dataset ($c : \mathcal{A} \to \mathcal{A}$), we have effectively created a partition such that each part contains only a subset of all choice problems. Nevertheless, as demonstrated in Section 3, ordered-reference dependent expected utility can be achieved with normative restrictions on Ψ.

Appendix B: Additional Materials

ORDU vs other non-WARP models

To simplify notation, we use "$\{\underline{a}, b, c\}$" for $c(\{a, b, c\}) = \{a\}$.

In Ok et al. (2015)'s (endogeneous) *reference dependent choice*, the decision maker maximizes a single utility function, but only chooses from alternatives that are better than the reference in all (endogenously given) attributes. Consider the following choices accommodated by their model but not ORDU.[33]

$$\boxed{\boldsymbol{\{\underline{a}, b, c, d\} \quad \{a, \underline{b}, d\} \quad \{a, \underline{c}, d\}}}$$

An interpretation by Ok et al. (2015) is that a gives the highest utility but a *decoy* d blocks the choice of a from $\{a, b, d\}$ and $\{a, c, d\}$, the *attraction effect.* Yet their model does not require the decoy to function for $\{a, b, c, d\}$. On the contrary, ORDU requires that one of the reference points of $\{a, b, d\}$ and $\{a, c, d\}$ is the reference of $\{a, b, c, d\}$, a consequence of the reference order that excludes this behavior from ORDU. Finally, an intransitive choice correspondence $\{\underline{a}, b\}, \{\underline{b}, c\}, \{\underline{c}, a\}$ may be explained by ORDU, but is never accommodated by Ok et al. (2015). Hence the two models are not nested.

Kőszegi & Rabin (2006)'s *reference-dependent preferences* is another related model. In *personal equilibrium* (PE), decision makers has a joint utility function $v : X \times X \to \mathbb{R}$ and chooses $PE(A) = \{x : v(x|x) \geq v(y|x)\ \forall y \in A\}$; that is, the choice maximizes a reference-dependent utility function, and the reference point is itself the eventually chosen alternative. This permits the type of behavior

[33] The complete choice correspondence is

$\boldsymbol{\{\underline{a}, b, c, d\}}$	$\{\underline{a}, b, c\}$	$\boldsymbol{\{a, \underline{b}, d\}}$	$\boldsymbol{\{a, \underline{c}, d\}}$	$\{\underline{b}, c, d\}$	$\{\underline{a}, b\}$
$\{\underline{a}, c\}$	$\{\underline{a}, d\}$	$\{\underline{b}, c\}$	$\{\underline{b}, d\}$	$\{\underline{c}, d\}$	

Using an Ok et al. (2015) specification where $u(a) > u(b) > u(c) > u(d)$. $r(\{a, b, d\}) = r(\{a, c, d\}) = r(\{b, d\}) = r(\{c, d\}) = d$, $r(A) = \diamond$ otherwise, and $\mathcal{U} = \{U\}$ where $U(b) > U(c) > U(d) > U(a)$.

where x is not chosen in a set but is chosen in the subset—when the alternatives x fails to beat are removed. While ORDU also allows for $x \in c(B) \backslash c(A)$ where $B \subset A$, it does so with two implications: (i) an alternative $y \in A \backslash B$ must had been the reference point of A and so (ii) for some $y \in A \backslash B$, choices are consistent between $c(A)$ and $c(T)$ for all $T \subset A$ that contains y. Consider the following example.

$$\boxed{\boldsymbol{\{\underline{a}, b, c, d\}} \quad \boldsymbol{\{\underline{a}, \underline{b}, c\}} \quad \boldsymbol{\{\underline{a}, \underline{d}\}}}$$

Since $c(\{a, b, c, d\})$ and $c(\{a, b, c\})$ does not come from standard utility maximization, the reference of $\{a, b, c\}$ is d, and so $c(\{a, b, c, d\})$ and $c(\{a, d\})$ must maximize the same utility function. But this is not the case either, so this choice pattern is incompatible with ORDU. It is, however, compatible with PE.[34]An immediate implication PE is: $x \in c(A)$ and $x \in B \subset A$, then $x \in c(B)$. A simple intransitive choice pattern $\{\underline{a}, b\}, \{\underline{b}, c\}, \{\underline{c}, a\}$, admissible by ORDU, concludes that the two models are not nested.

Manzini & Mariotti (2007) proposes a non-WARP model without a reference point interpretation. In *rational shortlist method* (RSM), decision makers are endowed with two asymmetric relations P_1 and P_2. Facing a choice problem A, she first creates a shortlist by eliminating inferior alternatives according to P_1 (eliminate x if yP_1x for some $y \in A$), and then choose from this shortlist according to P_2. WARP violation appears when an alternative x is eliminated in a set S, but not in the subset $T \subset S$, where it subsequently chosen over the best alternative of S. An example of this behavior is displayed by the first of the following choices.

$$\boxed{\boldsymbol{\{\underline{a}, b, c, d\}} \quad \boldsymbol{\{a, b, \underline{d}\}} \quad \boldsymbol{\{b, \underline{c}, d\}} \quad \boldsymbol{\{\underline{b}, c\}}}$$

For ORDU to reconcile, c must be deemed the reference of $\{a, b, c, d\}$, but then choices from $\{b, c, d\}$ and $\{b, c\}$ must comply with standard utility maximization. Since this is not the case, ORDU does not nest RSM.[35] While ORDU is constrained by fixed reference points, the model is more flexible than

[34]The complete choice correspondence is

$\boldsymbol{\{\underline{a}, b, c, d\}}$	$\boldsymbol{\{\underline{a}, \underline{b}, c\}}$	$\{\underline{a}, b, \underline{d}\}$	$\{\underline{a}, c, d\}$	$\{b, \underline{c}, d\}$	$\{\underline{a}, \underline{b}\}$
$\{\underline{a}, c\}$	$\boldsymbol{\{\underline{a}, \underline{d}\}}$	$\{\underline{b}, \underline{c}\}$	$\{b, \underline{d}\}$	$\{\underline{c}, d\}$	

Gul et al. (2006) shows that PE is equivalent to choices maximizing a complete (but not necessarily transitive) preference relation. This choice correspondence is explained by $a \sim b$, $a \succ c$ $a \sim d$, $b \sim c$, $d \succ b$, $c \succ d$.

[35]The complete choice correspondence is

$\boldsymbol{\{\underline{a}, b, c, d\}}$	$\{\underline{a}, b, c\}$	$\boldsymbol{\{b, \underline{c}, d\}}$	$\{\underline{a}, c, d\}$	$\boldsymbol{\{a, b, \underline{d}\}}$	$\{\underline{a}, b\}$
$\{\underline{a}, c\}$	$\{a, \underline{d}\}$	$\boldsymbol{\{\underline{b}, c\}}$	$\{b, \underline{d}\}$	$\{\underline{c}, d\}$	

,

induced by (aP_1b, aP_1c, cP_1d, dP_1b) and (aP_2b, aP_2c, dP_2a, bP_2c, dP_2b, cP_2d).

RSM when reference points do change, since no restriction is put on the new utility function. RSM, however, cannot accommodate a choice that makes the shortlists in a small and large set but not an intermediate one. The result is the following behavior accommodated by ORDU but not RSM.[36]

$$\boxed{\boldsymbol{\{\underline{a}, b, c, d\}} \quad \boldsymbol{\{a, \underline{b}, d\}} \quad \boldsymbol{\{\underline{a}, b\}}}$$

We conclude that ORDU and RSM are not nested.

Last, we compare ORDU to Masatlioglu et al. (2012)'s *choice with limited attention* (CLA). A decision maker has a complete and transitive ranking $\succ_{CLA}$ of alternatives and an attention filter that limits choices to a subset of each choice problems, the "consideration set". When another choice problem is derived by removing choices not in the consideration set, the consideration set remains the same. Although a single ranking is used (as opposed to ORDU's many utility functions), flexibility in constructing consideration sets easily allows for behavior not accommodated by ORDU.

$$\boxed{\boldsymbol{\{a, \underline{b}, c\}} \quad \boldsymbol{\{\underline{a}, b\}} \quad \boldsymbol{\{b, \underline{c}\}} \quad \{\underline{a}, c\}}$$

CLA is provided under the framework of choice functions (no indifference), and with that restriction ORDU is nested by CLA.[37]However, the two models make different predictions under a comparable setup. For the analysis, we modify CLA by allowing for indifferences in the ranking of alternatives (replacing $\succ_{CLA}$ with $\succsim_{CLA}$), but preserve in entirety the attention filter / consideration set component. The following behavior is accommodated by ORDU but not CLA.[38]

$$\boxed{\boldsymbol{\{\underline{a}, \underline{b}, c, d\}} \quad \boldsymbol{\{a, \underline{b}, \underline{c}\}} \quad \boldsymbol{\{\underline{b}, c\}}}$$

When indifferences are allowed, the single ranking limitation of CLA becomes the bottleneck in explaining behavior. The two models are hence not nested under comparable setups.

[36]The complete choice correspondence is

$$\boxed{\begin{array}{cccccc} \boldsymbol{\{\underline{a}, b, c, d\}} & \{a, \underline{b}, c\} & \boldsymbol{\{a, \underline{b}, d\}} & \{\underline{a}, c, d\} & \{\underline{b}, c, d\} & \boldsymbol{\{\underline{a}, b\}} \\ \{\underline{a}, c\} & \{\underline{a}, d\} & \{\underline{b}, c\} & \{\underline{b}, d\} & \{\underline{c}, d\} & \end{array}}.$$

This is explained by the ORDU specification: $bRaRcRd$, $u_i(a) > u_i(b) > u_i(c) > u_i(d)$ when $i \in \{a, b, d\}$, and $u_c(b) > u_c(a) > u_c(c) > u_c(d)$.

[37]Consider any choice function c that admits an ORDU representation, define CLA's parameters as follows: attention filter $\Gamma(A) := \left\{\min(A, R), \arg\max_{x \in A} u_{\min(A,R)}(x)\right\}$ (singleton if $\min(A, R) = \arg\max_{x \in A} u_{\min(A,R)}(x)$) and CLA's preference $x \succ y$ if xRy.

[38]The complete choice correspondence is

$$\boxed{\begin{array}{cccccc} \boldsymbol{\{\underline{a}, \underline{b}, c, d\}} & \boldsymbol{\{a, \underline{b}, \underline{c}\}} & \{\underline{a}, \underline{b}, d\} & \{\underline{a}, c, d\} & \{\underline{b}, c, d\} & \{a, \underline{b}\} \\ \{a, \underline{c}\} & \{\underline{a}, d\} & \boldsymbol{\{\underline{b}, c\}} & \{\underline{b}, d\} & \{\underline{c}, d\} & \end{array}}.$$

This is explained by the ORDU specification: $cRbRaRd$, $u_d(a) = u_d(b) > u_d(c) > u_d(d)$, $u_a(b) = u_a(c) > u_a(a)$, $\underline{u_b(b) > u_b(c)}$. Now we show non-compliance with CLA (with the indifference extension): Since $a \in c(\{a, b, c, d\}) \setminus c(\{a, b, c\})$, CLA reconciles this by setting the consideration sets $\Gamma(\{a, b, c, d\}) = \{a, b, d\}$ and $\Gamma(\{a, b, c\}) = \{b, c\}$, so a is not considered in the smaller set. However, the property of consideration sets then requires $\Gamma(\{b, c\}) = \{b, c\}$, and $\{a, c\}$

Appendix C: Proofs

Proof of Lemma 2

Lemma 3. *Let Z be a set, and $\mathbb{Z}$ be the set of all finite and nonempty subsets of Z. Let$\mathcal{R}$ be a self map on $\mathbb{Z}$, $\mathcal{R}(S) \subseteq S$, such that*

(i) For all $S \in \mathbb{Z}$, $\mathcal{R}(S) \neq \{\emptyset\}$, and

(ii) α - for all $T, S \in \mathbb{Z}$, $x \in T \subseteq S$, if $x \in \mathcal{R}(S)$, then $x \in \mathcal{R}(T)$.

Then, there exist $\mathcal{R}^* \subseteq \mathcal{R}$ such that

(i) For all $S \in \mathbb{Z}$, $\mathcal{R}^*(S) \neq \{\emptyset\}$,

(ii) α - for all $T, S \in \mathbb{Z}$, $x \in Z$ such that $x \in T \subseteq S$, if $x \in \mathcal{R}^*(S)$, then $x \in \mathcal{R}^*(T)$, and

(iii) β - for all $T, S \in \mathbb{Z}$, $x, y \in Z$ such that $x, y \in T \subseteq S$, if $x \in \mathcal{R}^*(T)$ and $y \in \mathcal{R}^*(S)$, then $x \in \mathcal{R}^*(S)$.

Proof. We prove by construction.

1. We say $\mathcal{R}' \subseteq \mathcal{R}$ if $\mathcal{R}'(S) \subseteq \mathcal{R}(S)\ \forall S \in \mathbb{Z}$. Assume and invoke Zermelo's theorem to well-order the set of all doubletons in the domain of $\mathcal{R}$ (there may be uncountable many of them, hence Zermelo's theorem). Now we start the transfinite recursion using this order.

2. In the zero case, we have $\mathcal{R}_0 = \mathcal{R}$. This correspondence satisfies α and is nonempty-valued. Suppose $\mathcal{R}_\sigma$ satisfies α and is nonempty-valued.

3. For the successor ordinal $\sigma+1$, we take the corresponding doubleton $B_{\sigma+1}$ and take $x \in B_{\sigma+1}$ such that $\forall S \supset B_{\sigma+1}$, $\mathcal{R}(S) \setminus \{x\} \neq \emptyset$. Suppose such an x does not exist, then for both $x, y \in B_{\sigma+1}$, there are $S_x \supset B_{\sigma+1}$ and $S_y \supset B_{\sigma+1}$ such that $\mathcal{R}_\sigma(S_x) = \{x\}$ and $\mathcal{R}_\sigma(S_y) = \{y\}$ since $\mathcal{R}_\sigma$ is nonempty-valued. Consider $S_x \cup S_y \in \mathbb{Z}$. Since $\mathcal{R}_\sigma$ is nonempty-valued, $\mathcal{R}_\sigma(S_x \cup S_y) \neq \emptyset$. But since $\mathcal{R}_\sigma$ satisfies α, it must be that $\mathcal{R}_\sigma(S_x \cup S_y) \subseteq \mathcal{R}_\sigma(S_x) \cup \mathcal{R}_\sigma(S_y)$, hence $\mathcal{R}_\sigma(S_x \cup S_y) \subseteq \{x, y\}$. Suppose without loss $x \in \mathcal{R}_\sigma(S_x \cup S_y)$, then due to α again and that $x \in B_{\sigma+1} \subset S_y$, it must be that $x \in \mathcal{R}_\sigma(S_y)$, which contradicts $\mathcal{R}_\sigma(S_y) = \{y\}$. (That is, we showed that with nonempty-valuedness and α, no two elements can each have a unique appearance in the $\mathcal{R}_{(\cdot)}$-image of a set containing those two elements.) Hence, $\exists x \in B_{\sigma+1}$ such that $\forall S \supset B_{\sigma+1}$, $\mathcal{R}(S) \setminus \{x\} \neq \emptyset$. Define $\mathcal{R}_{\sigma+1}$ from $\mathcal{R}_\sigma$ in the following way: $\forall S \supset B_{\sigma+1}$, $\mathcal{R}_{\sigma+1}(S) :=$

$\mathcal{R}_\sigma(S) \setminus \{x\}$. Note: (i) Since x is deleted from $\mathcal{R}_\sigma(T)$ only if it is also deleted (if it is in it at all) from $\mathcal{R}_\sigma(S)$ $\forall S \supset T$, we are preserving α, and (ii) since x is never the unique element in $\mathcal{R}_\sigma(S)$ $\forall S \supset B_{\sigma+1}$, we preserve nonempty-valuedness.

4. For a limit ordinal λ, define $\mathcal{R}_\lambda = \cap_{\sigma<\lambda}\mathcal{R}_\sigma$. Note that since $\mathcal{R}_{\sigma'} \subset \mathcal{R}_{\sigma''}$ $\forall \sigma' > \sigma''$, $\cap_{\sigma\leq\bar{\sigma}} = \mathcal{R}_{\bar{\sigma}}$. Furthermore, for any $\sigma < \lambda$, $\mathcal{R}_\sigma$ is constructed such that α and nonempty-valuedness are preserved. Hence $\mathcal{R}_\lambda$ satisfies α and is nonempty-valued.

5. Note that this process terminates when all the doubletons have been visited, for we would otherwise have constructed an injection from the class of all ordinals to the set of all doubletons in $\mathbb{Z}$, which is impossible.

6. Finally, we check that $|\mathcal{R}_\lambda(S)| = 1$ for all $S \in \mathbb{Z}$, hence β is satisfied trivially. Suppose it is not a function, hence $\exists S \in \mathbb{Z}$ such that $x, y \in \mathcal{R}_\lambda(S)$. Then by α we have that $x, y \in \mathcal{R}_\lambda(\{x, y\})$, which is not possible as the recursion process has visited it and deleted something from $\mathcal{R}(\{x, y\})$.

7. Set $\mathcal{R}_\lambda = \mathcal{R}^*$.

$\square$

We state the following observation. Let $c : \mathcal{A} \to \mathcal{A}$ be a choice correspondence. Recall that $\mathcal{A}$ is the set of all finite and nonempty subsets of X. For $S \subseteq Y$ and $x \in S$, define $\mathbb{A}^x_S := \{A \subseteq S : A \in \mathcal{A} \text{ and } x \in A\}$. We use the notation $(c, \mathbb{A}^x_A)$ to denote the choice correspondence $\tilde{c} : \mathbb{A}^x_A \to \mathcal{A}$ where $\tilde{c}(B) = c(B)$. In other words, $(c, \mathbb{A}^x_A)$ is a subset of c where the domain is restricted to $\mathbb{A}^x_A$– the set of all subsets of A containing x.

Remark 2. Let $c : \mathcal{A} \to \mathcal{A}$ be a choice correspondence and $\mathcal{T}$ a finite property as defined in Definition 13. Define $\Gamma(S) := \{x \in S : (c, \mathbb{A}^x_A) \in \mathcal{T}\}$.

1. If $y \in \Gamma(A)$, then $y \in \Gamma(B)$ whenever $B \subset A$.

2. If $y \in \Gamma(A)$ for all finite $A \subseteq D$, then $y \in \Gamma(D)$.

We call x a reference alternative for S if $x \in \Gamma(S)$. Remark 2 states that if x is a reference alternative for some choice problem A, i.e., $(c, \mathbb{A}^x_A) \in \mathcal{T}$, then x is also a reference alternative for $B \subseteq A$. This is an immediate consequence of the definition of a property (and the fact that $\mathbb{A}^x_B \subseteq \mathbb{A}^x_A$ whenever $B \subseteq A$). In words, if a violation is undetected with more observations, then it cannot be detected with less. Furthermore, if x is a reference alternative for all finite subsets of an arbitrary set of alternatives D, then x is also a reference alternative for D; this, is due to $\mathcal{T}$ being a finite property. Otherwise, take a finite set of choice problems $\mathcal{S} = A_1, ..., A_n$, each of which a subset of D containing x, such that a finite property is violated, i.e., $\tilde{c} : \mathcal{S} \to \mathcal{A}$, where $\tilde{c}(B) = c(B)$, is not in $\mathcal{T}$. Since this is a finite tuple of finite choice problems, consider the finite set $A := \cup_i A_i$. Clearly, $x \notin \Gamma(A)$, but A is a finite subset of D, hence a contradiction. Intuitively, if x is not a reference alternative for some arbitrary set of alternatives D, then violation of a finite property would have been detected in a finite subset of D, rendering x not a reference alternative for some choice problem $A \subseteq D$.

Now, let $\mathcal{R}' : \mathcal{A} \to \mathcal{A} \cup \{\emptyset\}$ be a set valued function that picks out, for each choice problem $A \in \mathcal{A}$, the set of all reference alternatives $\mathcal{R}'(A) \subseteq A$; formally, $\mathcal{R}'(A) := \{x \in S : (c, \mathbb{A}_A^x) \in \mathcal{T}\}$. Since $\mathcal{T}$ is a finite property, by Remark 2, $\mathcal{R}'$ satisfies property α (as defined in Lemma 3). Furthermore, the hypothesis in Lemma 2 gives that $\mathcal{R}'(A) \cap \Psi(A)$ is nonempty for all $A \in \mathcal{A}$. Finally, define $\mathcal{R} : \mathcal{A} \to \mathcal{A}$ by $\mathcal{R}(A) := \mathcal{R}'(A) \cap \Psi(A)$. Since both $\mathcal{R}'(A)$ and $\Psi(A)$ satisfy property α, $\mathcal{R}(A)$ satisfies property α.

Putting our $\mathcal{R}$ through Lemma 3, we get a set-valued function $\mathcal{R}^* : \mathcal{A} \to \mathcal{A}$ that is now a singleton everywhere (i.e., $|\mathcal{R}^*(A)| = 1$ for all $A \in \mathcal{A}$). Furthermore, this function satisfies property α, and satisfies property β trivially. With this, we build the order (R, Y) by setting xRy if $\{x\} = \mathcal{R}^*(\{x, y\})$, and xRx. The result is a complete, transitive, and antisymmetric binary relation.

Lemma 4. *For an (R, Y) constructed according to the the aforementioned procedure, $y \in A \setminus \Psi(A) \Rightarrow xRy$ for some $x \in \Psi(A)$ (i.e. R is Ψ-consistent).*

Proof. Suppose not, say $y \in A \setminus \Psi(A)$ but yRx for all $x \in \Psi(A)$. Consider $\{\{x, y\} : x \in \Psi(A)\}$. Since this is a finite set of doubletons, suppose without loss of generality $\{x^*, y\}$ is the last one (in $\{\{x, y\} : x \in \Psi(A)\}$) visited by the procedure in Lemma 3, and denote the step corresponding to $\{x^*, y\}$ by the ordinal $\sigma_{\{x^*, y\}}$. Since yRx for all $x \in \Psi(A)$ such that $x \neq x^*$, $\mathcal{R}_{\sigma_{\{x^*, y\}}}(A) \cap \Psi(A) = \{x^*\}$. Since $\mathcal{R}_\sigma \subseteq \mathcal{R}_0 := \mathcal{R}' \cap \Psi$ for all σ, $\mathcal{R}_{\sigma_{\{x^*, y\}}}(A) = \{x^*\}$. Hence x^* uniquely appears in the image of $\mathcal{R}_{\sigma_{\{x^*, y\}}}$ evaluated at some superset of $\{x^*, y\}$, and the recursion procedure sets, ultimately, $\mathcal{R}^*(\{x^*, y\}) = \{x^*\}$. But this implies $x^* Ry$, a contradiction. □

Finally, consider the set $R^{\downarrow}(x) := \{y \in X : xRy\}$. This is a set of alternatives that are, according to our binary relation R, reference dominated by x. For any finite subset $A \subseteq R^{\downarrow}(x)$ such that $x \in A$, we have $x \in \mathcal{R}^*(A) \subseteq \mathcal{R}(A) \subseteq \mathcal{R}'(A)$, which by definition implies x is a reference alternative of A. Using point 2 in Remark 2, we conclude that x is reference alternative for $R^{\downarrow}(x)$, which need not be finite.

To summarize, we have effectively created a partition of $\mathcal{A}$ where the parts are characterized by $\left\{\mathbb{A}^x_{R^\downarrow(x)}\right\}_{x\in X}$. To see this, take any $A\in\mathcal{A}$, since R is a linear order, there is a unique $z\in A$ such that zRy for all $y\in A$, and so $A\in\mathbb{A}^z_{R^\downarrow(z)}$ and $A\notin\mathbb{A}^y_{R^\downarrow(y)}$ for any $y\neq z$. Furthermore for each part $\mathbb{A}^x_{R^\downarrow(x)}$, $\left(c,\mathbb{A}^x_{R^\downarrow(x)}\right)$ is in $\mathcal{T}$. Since $\left\{B\in\mathcal{A}:\underset{y\in B}{\arg\max}R=z\right\}$ is simply $\mathbb{A}^z_{R^\downarrow(z)}$, the proof is complete.

Proof of Proposition 1, Part 1 (without the use of Lemma 2)

Suppose X is finite. We provide an independent proof that a choice correspondence c that satisfies RD (Axiom 1) has an ORDU representation.

1. Let $\Gamma(A)$ be the set of reference alternatives for A. We create a list of alternatives in the following way; list $\Gamma(X)$ with an arbitrary order. Since $X\backslash\Gamma(X)$ is again finite, list $\Gamma(X\backslash\Gamma(X))$ with an arbitrary order; and continue until all $x\in X$ are listed. Finally, let i_x denote the position of x in the list. For any $x,y\in X$, construct xRy if $i_x>i_y$ and xRx.

2. We now construct $\succsim_x$ for each $x\in X$. Consider the set $R^\downarrow(x):=\{y:xRy\}$. Consider c on $\mathbb{A}^x_{R^\downarrow(x)}:=\left\{A\subseteq R^\downarrow(x)\cap\mathcal{A}:x\in A\right\}$, which by construction satisfies WARP.

3. First we set $x\succsim_x x$ for all $x\in X$.

4. Using the doubletons in $\mathbb{A}^x_{R^\downarrow(x)}$, all of which would contain x, we set, for all $y\in R^\downarrow(x)$, either $y\succsim_x x$, or $x\succsim_x y$, or both, according to $c(\{x,y\})$.

5. Now for all $y_1,y_2\succsim_x x$, we set either $y_1\succsim_x y_2$, or $y_2\succsim_x y_1$, or both, according to $c(\{x,y_1,y_2\})$, using tripletons in $\mathbb{A}^x_{R^\downarrow(x)}$. Due to WARP (of c on $\mathbb{A}^x_{R^\downarrow(x)}$), $\succsim_x$ is now complete on the set $\mathbb{P}^x:=\{y:y\succsim_x x\}$, which we call the *prediction set* of x, containing alternatives that are both reference dominated by x (i.e. xRy) and are weakly better than x in binary comparison (i.e. $y\in c(\{y,x\})$).

6. Now consider $X\backslash\mathbb{P}^x=\{y:yRx\text{ or }x\succ_x y\}$. We set $y_1\sim_x y_2$ for all $y_1,y_2\in X\backslash\mathbb{P}^x$, and $y_1\succ_x y_2$ for all $y_1\in\mathbb{P}^x$, $y_2\in X\backslash\mathbb{P}^x$. Our constructed $\succsim_x$ is now complete (on X).[39]

Using quadrupletons in $\mathbb{A}^x_{R^\downarrow(x)}$, we show that $\succsim_x$ constructed above is transitive: Suppose $y_1\succsim_x y_2$ and $y_2\succsim_x y_3$, and that $y_1,y_2,y_3\in\mathbb{P}^x$ (if any of them is in $X\backslash\mathbb{P}^x$ then the argument is straightforward by $\sim_x$), hence $y_1\in c(\{x,y_1,y_2\})$ and $y_2\in c(\{x,y_2,y_3\})$. Furthermore, since $y_1,y_2,y_3\in\mathbb{P}^x$, we have

[39] That is, for any $y_1,y_2\in X$, either $y_1\succsim_x y_2$, or $y_2\succsim_x y_1$, or both.

$\{x, y_1, y_2, y_3\} \in \mathbb{A}^x_{R^{\downarrow}(x)}$, and c on $\mathbb{A}^x_{R^{\downarrow}(x)}$ satisfies WARP implies $y_1 \in c(\{x, y_1, y_2, y_3\})$, and hence $y_1 \in c(\{x, y_1, y_3\})$, which implies $y_1 \succsim_x y_3$.

Finally, we show that (R, X) and $\{(\succsim_x, X)\}_{x \in X}$ explain c. Take any $A \in \mathcal{A}$, since A is finite, and R is antisymmetric, there is a unique $R-$maximizer $x \in A$ (i.e., xRy for all $y \in A$), hence $A \subseteq R^{\downarrow}(x)$. Suppose for contradiction $c(A) \nsubseteq \{y \in A : y \succsim_x z \, \forall z \in A\}$; so for some $a \in c(A)$, $a' \succ_x a$ for some $a' \in A$. Then $a \notin c(\{x, a', a\})$. Since $\{x, a', a\}$ is a subset of A, and both choice problems are elements of $\mathbb{A}^x_{R^{\downarrow}(x)}$, this is a violation of the statement c satisfies WARP on $\mathbb{A}^x_{R^{\downarrow}(x)} := \left\{A \subseteq R^{\downarrow}(x) \cap \mathcal{A} : x \in A\right\}$, hence a contradiction. Suppose for contradiction $c(A) \nsupseteq \{y \in A : y \succsim_x z \, \forall z \in A\}$, so for some $a \in A$, $a \succsim_x z$ for all $z \in A$, but $a \notin c(A)$. Take $a' \in c(A)$; since $a \succsim_x a'$, $a \in c(\{x, a', a\})$. Since $\{x, a', a\}$ is a subset of A, and both choice problems are elements of $\mathbb{A}^x_{R^{\downarrow}(x)}$, a contradiction of the statement c satisfies WARP on $\mathbb{A}^x_{R^{\downarrow}(x)}$ is reached. Hence $c(A) = \{y \in A : y \succsim_x z \, \forall z \in A\}$.

It remains to show that for each alternative-indexed preference relation defined, we can construct a utility function representing it. Since X is finite, and each $\succsim_x$ is a complete and transitive preference relation, this is standard.

Proof of Proposition 1, Part 2

We invoke Lemma 2 to prove the intermediary result that, if c satisfies Reference Dependence (Axiom 1) and Continuity (Axiom 2), then there exists a linear order (R, X) and a set of complete preference relations $\{(\succsim_x, X)\}_{x \in X}$ such that for all $A \in \mathcal{A}$, we have $c(A) = \left\{y \in A : y \succsim_{r(A)} x \, \forall x \in A\right\}$, where $r(A) = \underset{z \in A}{\arg\max} R$:

Using the notation in Subsection **??**, define $\mathcal{T}$ as the property WARP. By Lemma 2, there exists a Ψ-consistent linear order (R, X) such that c on $\left\{S \in \mathcal{A} : \underset{z \in A}{\arg\max} S = x\right\}$ satisfies $\mathcal{T}$ for all $x \in X$. Notice that $\left\{S \in \mathcal{A} : \underset{z \in A}{\arg\max} S = x\right\} = \mathbb{A}^x_{R^{\downarrow}(x)}$, and so we conclude that for all $T, S \in \mathbb{A}^x_{R^{\uparrow}(x)}$, $c(S) \cap T = c(T)$ whenever $T \subset S \subseteq A$ and $c(S) \cap T \neq \emptyset$. We proceed to build $\{(\succsim_x, X)\}_{x \in X}$ using the method outlines in the special case proof above, which gives us a complete and transitive $\succsim_x$ for all x, as well as $c(A) = \left\{y \in A : y \succsim_{r(A)} z \, \forall z \in A\right\}$ where $r(A) = \underset{x \in A}{\arg\max} S := \{x \in A : xRy \, \forall y \in A\}$.

It remains to show that for each alternative-indexed preference relation defined, we can construct a utility function representing it. [To be completed, but essentially just Efe Ok Order 9 Pg 18]

Proof of Proposition 2

Remark. (Notational) Currently, this older version of the proof reverses, without loss, the order R. That is, $r(A) = \underset{p\in A}{\arg\min} R$ as opposed to $\arg\max$. The proof remains valid, and readers are advised to simply, at the very end, "reverse" the order R constructed here.

We define $\Delta(X)$ as a $|X|-1$ dimensional simplex, as is conventional, and hence *full-dimensional* means $|X|-1$ dimensional. First, we split $r \in \Delta(X)$ into two groups, $E = \{r \in \Delta(X) : r = (\delta_b)^{\alpha}(\delta_w)\}$, and $I = \Delta(X)\backslash E$, the "exterior" and "interior" sets. Set $\Psi(A) = A\backslash\Phi(A)$, it is easy to check that $a \in \Psi(B)$ if $a \in B \subseteq A$ and $a \in \Psi(A)$. Applying Lemma 2, we get a linear order $(R, \Delta(X))$ that gives a partition of $\mathcal{A}$, $\left\{\mathbb{A}^r_{R^{\uparrow}(r)}\right\}_{r\in\Delta X}$, such that c on $\mathbb{A}^r_{R^{\uparrow}(r)}$ satisfies WARP and Independence for all $r \in \Delta(X)$. Furthermore, since R is Ψ-consistent, or $\min(A, R) \in A\backslash\Phi(A)$, we have pRq for all $p \in \Phi(\{p,q\})$.

Lemma 5. *For $r \in I$ and any open ball B_r around r, $B_r \cap R^{\uparrow}(r)$ contains a full-dimensional convex subset of $\Delta(X)$.*

Proof. Take $r \in I$. By definition, $r(x) \neq 0$ for some $x \neq b, w$ ($r(x)$ is the probability that lottery r gives prize x). Consider all mean-preserving spread of r, $MPS(r) \subseteq \Delta A$, this is a $|X-2|$ dimensional convex set. Since $q \in MPS(r)$ implies $q \in \Phi(\{r,p\})$, we have that qRr and hence $MPS(r) \subseteq R^{\uparrow}(r)$. Consider the set $\mathbb{S}(r) := \cup_{q\in MPS(r)\cup\{r\}} \{e \in \Delta X : e \text{ is an extreme spread of } q\}$, this is an interval on the line connecting δ_b and δ_w. Consider the convex hull $\mathbb{C}(r) := \operatorname{conv}(MPS(r) \cup \mathbb{S}(r) \cup \{r\})$. Clearly, $\mathbb{C}(r)$ is a convex set. Furthermore, since $\mathbb{S}(r)$ is not contained in $MPS(r) \cup \{r\}$ (otherwise lotteries in $\mathbb{S}$ has the same mean, but this is not possible), $\mathbb{C}(r)$ is full dimensional. Since $e \in \mathbb{S}(r)$ only if e is an extreme spread of q for some $q \in MPS(r) \cup \{r\} \subseteq R^{\uparrow}(r)$, and $e \in \Phi(\{e,q\})$, we have $eRqRr$, hence $\mathbb{S}(r) \subseteq R^{\uparrow}(r)$. Finally, for $p \in \mathbb{C}(r) \setminus (MPS(r) \cup \mathbb{S}(r) \cup \{r\})$, it must be that $p = e^{\alpha}q$ for some $q \in MPS(r) \cup \{r\}$ and e an extreme spread of q, hence again $p \in \Phi(\{p,q\})$, so $pRqRr$, so $\mathbb{C}(r) \subseteq R^{\uparrow}(r)$. Since B_r is also a full-dimensional and convex set, $B_r \cap \mathbb{C}(r)$ is a full-dimensional convex set in $B_r \cap R^{\uparrow}(r)$. □

Define for each $r \in \Delta X$ the *strict prediction set* $\mathbb{P}^+_r := \left\{p \in R^{\uparrow}(r) : r \notin c(\{p,r\})\right\}$. There are lotteries that are both reference dominated by r and is chosen over r in a binary decision.

Lemma 6. *For $r \in I$, $\mathbb{P}^+_r$ contains a full-dimensional convex subset of $\Delta(X)$.*

Proof. Take $r \in I$. Suppose for contradiction $r \in c(\{e,r\})$ for all e an extreme spread of r; then since the lottery $(\delta_w)^{r(w)}(\delta_b)$ is in the closure of the extreme spread of r, continuity of c implies

$r \in c\left(\left\{r, (\delta_w)^{r(w)}(\delta_b)\right\}\right)$, which is a violation of FOSD. Hence there is an extreme spread of r, e, such that $r \notin c(\{r, e\})$. Since $r^\alpha e \in R^\uparrow(r)$ and c on $\mathbb{A}^r_{R^\uparrow(r)}$ satisfies Independence, we can find $p := r^\alpha e \in \mathbb{P}^+_r$ where $\alpha \in (0,1)$, hence $p \in I$. By continuity of c, there exists an open ball B_p around p such that $r \notin c(\{r, q\})$ for all $q \in B_p$. By Lemma 5, $B_p \cap R^\uparrow(p)$ contains a full-dimensional convex subset of $\Delta(X)$. Since pRr, $B_p \cap R^\uparrow(p) \subseteq B_p \cap R^\uparrow(r)$, hence $\mathbb{P}^+_r$ contains a full-dimensional convex subset of $\Delta(X)$. $\square$

Immediately, this implies that for $r \in I$, we can build an increasing $u_r : X \to \mathbb{R}$, unique up to an affine transformation, such that $c(A) = \arg\max_{p \in A} \mathbb{E}_p u_r(x)$ if $A \in \mathbb{A}^r_{R^\uparrow(r)}$. The technique is standard. Let $\mathbb{P}$ be a full-dimensional convex subset of $\mathbb{P}^+_r$. First, notice that for all $p, q \in \mathbb{P}$, we have $\{r, p, q\} \in \mathbb{A}^r_{R^\uparrow(r)}$ and $r \notin c(\{r, p, q\})$. Recall that c on $\mathbb{A}^r_{R^\uparrow(r)}$ satisfies WARP and Independence. By define $p \succsim_r q$ if $p \in c(\{r, p, q\})$, we get a binary relation $(\succsim_r, \mathbb{P})$ that is complete, transitive, continuous, and satisfies independence. Since $\mathbb{P}$ is full-dimensional and convex, it contains a subset that is essentially a linear transformation of a $|X| - 1$ dimensional simplex. Since $(\succsim_r, \mathbb{P})$ satisfies FOSD, an increasing utility function $u_r : X \to \mathbb{R}$, unique up to an affine transformation, such that $c(A) = \arg\max_{p \in A} \mathbb{E}_p u_r(x)$ for all $A \in \mathbb{A}^r_{\mathbb{P}}$. We normalize this function to $u_r : X \to [0,1]$, where $u_r(w) = 0$ and $u_b(b) = 1$.

We now show that this utility function works for $\mathbb{A}^r_{R^\uparrow(r)}$. First, for any two lotteries $p, q \in \Delta X$, there exist $p', q' \in \mathbb{P}$ such that $p' = (p)^\alpha s$ and $q' = (q)^\alpha s$ for some $s \in \Delta X$ and $\alpha \in [0,1]$; we call p', q' $\mathbb{P}$-*common mixtures* of p, q. This can be done by using an arbitrary point $s \in \operatorname{int}(\mathbb{P})$ and take α small enough until both p' and q' enter $\mathbb{P}$. Take any $p \in R^\uparrow(r)$ and let r', p' be $\mathbb{P}$-common mixtures of r, p. Since c on $\mathbb{A}^r_{R^\uparrow(r)}$ satisfies Independence, $i' \in c(\{r, r', p'\})$ if and only if $i \in c(\{r, p\})$, for $i = r, p$. Now take any $p, q \in R^\uparrow(r)$ such that $p \in c(\{r, p\})$ and $q \in c(\{r, q\})$, then again by Independence on $\mathbb{A}^r_{R^\uparrow(r)}$, $p' \in c(\{r, p', q'\})$ if and only if $p \in c(\{r, p, q\})$, where p', q' are $\mathbb{P}$-common mixtures of p, q.

We have thus shown that $c(\{r, p\}) = \arg\max_{s \in \{r,p\}} \mathbb{E}_s u_r(x)$ for all $\{r, p\} \in \mathbb{A}^r_{R^\uparrow(r)}$ and $c(\{r, p, q\}) = \arg\max_{s \in \{r,p,q\}} \mathbb{E}_s u_r(x)$ for all $\{r, p, q\} \in \mathbb{A}^r_{R^\uparrow(r)}$ where $p \in c(\{r, p\})$ and $q \in c(\{r, q\})$. Since c on $\mathbb{A}^r_{R^\uparrow(r)}$ satisfies WARP, showing $c(A) = \arg\max_{p \in A} \mathbb{E}_p u_r(x)$ for all $A \in \mathbb{A}^r_{R^\uparrow(r)}$ is straightforward from here.

Corollary 1. *For $r \in \Delta(X)$ and $p \in R^\uparrow(r) \cap I$ such that $r \notin c(\{r, p\})$, there exists $q \in R^\uparrow(r) \cap I$ such that $\{q\} = c(\{r, p, q\})$. Furthermore, $\mathbb{P}^{+p}_r := \left\{q \in R^\uparrow(r) : \{q\} = c(\{r, p, q\})\right\}$ contains a full-dimensional convex subset of $\Delta(X)$.*

Proof. The proof utilizes techniques in the proofs of 5 and 6. First, we show the existence of $q \in R^{\uparrow}(r) \cap I$ such that $\{q\} = c(\{r,p,q\})$. Consider the set of extreme spread of of p, we know that this set is a subset of $R^{\uparrow}(p)$, and is hence a subset of $R^{\uparrow}(r)$. Notice that $r \notin c(\{r,p,e\})$ for any extreme spread e of p since c on $\mathbb{A}^{r}_{R^{\uparrow}(r)}$ satisfies WARP and $r \notin c(\{r,p\})$. Using the technique in the proof of Lemma 6, it must be that for some extreme spread e^* of p, we have $p \notin c(\{r,p,e^*\})$, otherwise by continuity of c we have $p \in c\left(\left\{r, p, (\delta_w)^{p(w)}(\delta_b)\right\}\right)$, a violation of FOSD. Take any non-trivial convex combination $p^{\alpha}e^*$, this is in $R^{\uparrow}(p) \subseteq R^{\uparrow}(r)$, in I, and $\{p^{\alpha}e^*\} = c(\{r,p,p^{\alpha}e^*\})$, so let $q = p^{\alpha}e^*$. Finally, by continuity of c, take an open ball B_q such that $q' \in B_q$ implies $\{q'\} = c(\{r,p,q'\})$. By Lemma 5, $B_q \cap R^{\uparrow}(q)$ contains a full-dimensional convex subset of $\Delta(X)$. Moreover, $B_q \cap R^{\uparrow}(q) \subseteq B_q \cap R^{\uparrow}(r) \subseteq \mathbb{P}^{+p}_{r}$. So $\mathbb{P}^{+p}_{r}$ contains a full-dimensional convex subset of $\Delta(X)$. $\square$

Lemma 7. *Consider $r_1, r_2 \in I$ and $r_2 R r_1$. Then $u_{r_1} = f \circ u_{r_2}$ for some concave and increasing $f : [0,1] \to [0,1]$.*

Proof. This proof uses Axiom 4. Take any $r_1, r_2 \in I$ such that $r_2 R r_1$. u_{r_1} and u_{r_2} are defined above, let $\bar{f}$ be defined on the utility numbers $u_{r_2}(x)$, $x \in X$, such that $u_{r_1}(x) = \bar{f} u_{r_2}(x)$. Since u_{r_1} and u_{r_2} are strictly increasing, $\bar{f}$ is strictly increasing in its domain. We show that for any $x_1, x_2, x_3 \in X$ such that $x_1 < x_2 < x_3$, we have $\bar{f}(\alpha u_2(x_1) + (1-\alpha) u_2(x_3)) \geq \alpha \bar{f}(u_2(x_1)) + (1-\alpha)\bar{f}(u_2(x_3))$, where $\alpha u_2(x_1) + (1-\alpha) u_2(x_3) = u_2(x_2)$. Suppose not, then for some $\beta < \alpha$, we have $\bar{f}(\alpha u_2(x_1) + (1-\alpha) u_2(x_3)) < \beta \bar{f}(u_2(x_1)) + (1-\beta)\bar{f}(u_2(x_3)) < \alpha \bar{f}(u_2(x_1)) + (1-\alpha)\bar{f}(u_2(x_3))$. Consider lotteries $\delta = \delta_{x_2}$ and $p = (\delta_{x_1})^{\beta}(\delta_{x_3})$. The previous equation shows that $\mathbb{E}_{\delta} u_{r_1}(x) < \mathbb{E}_{p} u_{r_1}(x)$ and $\mathbb{E}_{\delta} u_{r_2}(x) > \mathbb{E}_{p} u_{r_2}(x)$. Let δ_1, p_1 be $\mathbb{P}$-common mixtures of δ, p, where $\mathbb{P}$ here is a full-dimensional convex subset of $\mathbb{P}^{+r_2}_{r_1}$ if $r_1 \notin c(\{r_1, r_2\})$, and of $\mathbb{P}^{+}_{r_1}$ otherwise. Let δ_2, p_2 be $\mathbb{P}$-common mixtures of δ, p, where $\mathbb{P}$ here is a full-dimensional convex subset of $\mathbb{P}^{+}_{r_2}$. Since u_{r_1} and u_{r_2} are Bernoulli utility functions for r_1 and r_2 respectively, we have $\{p_1\} = c(\{r_1, \delta_1, p_1\})$ and $\{\delta_2\} = c(\{r_2, \delta_2, p_2\})$. Notice that $A := \{r_1, r_2, \delta_1, \delta_2, p_1, p_2\} \in \mathbb{A}^{r_1}_{R^{\uparrow}(r_1)}$, so $c(A) = \arg\max_{q \in A} \mathbb{E}_q u_{r_1}(x)$. We established that $\mathbb{E}_{r_1} u_{r_1}(x) < \mathbb{E}_{p_1} u_{r_1}(x)$, $\mathbb{E}_{r_2} u_{r_1}(x) < \mathbb{E}_{p_1} u_{r_1}(x)$, and $\mathbb{E}_{\delta_i} u_{r_1}(x) < \mathbb{E}_{p_i} u_{r_1}(x)$ for $i = 1, 2$, so $c(\{r_1, r_2, \delta_1, \delta_2, p_1, p_2\}) \subseteq \{p_1, p_2\}$. But this and $\{\delta_2\} = c(\{r_2, \delta_2, p_2\})$ violates Axiom 4. Finally, it is straightforward that one can extend $\bar{f}$ to a concave function $f : [0,1] \to [0,1]$ (for example, by connecting the points with straight lines). $\square$

At this point we are almost done with proving the representation, less $r \in E$.

Lemma 8. *For $r \in E$ and $p \in R^{\uparrow}(r)$, $p \neq r$, either p first order stochastically dominates r or the converse.*

Proof. Take $r \in E$ and $p \in R^{\uparrow}(r)$, $p \neq r$. Let $\alpha = r(b)$, then $r(w) = 1 - \alpha$. If $p(b) < \alpha$ and $p(w) < (1-\alpha)$, then r is an extreme spread of p and rRp, so $p \notin R^{\uparrow}(r)$. Furthermore, it is not possible that $p(b) \geq \alpha$ and $p(w) \geq (1-\alpha)$ if $p \neq r$. Hence either $p(b) \geq \alpha$ and $p(w) \leq (1-\alpha)$ with at least one strict inequality, or $p(b) \leq \alpha$ and $p(w) \geq (1-\alpha)$ with at least one strict inequality. If the earlier, p FOSD r; if the later r FOSD p. $\square$

With this observation in mind, we construct u_r for $r \in E$. Define $E_1 := \{r \in E : r \notin c(\{r,p\}) \text{ for some } p \in R^{\uparrow}(r) \cap I\}$. and $E_2 := E \backslash E_1$. For $r \in E_1$, $\mathbb{P}^{+}_{r}$ contains a full-dimensional convex subset of $\Delta(X)$, and so we will build u_r using the same method we used to build u_r for $r \in I$. We will construct u_r for $r \in E_2$ after the following result.

Corollary 2. *Consider $r_1, r_2 \in I \cup E_1$ and $r_2 R r_1$. Then $u_{r_1} = f \circ u_{r_2}$ for some concave and increasing $f : [0,1] \to [0,1]$.*

Proof. Consider the proof in Lemma 7, but that when $r_2 \in E_1$, we simply let δ_1, p_1 be $\mathbb{P}$-common mixtures of δ, p, where $\mathbb{P}$ here is a full-dimensional convex subset of $\mathbb{P}^{+}_{r_1}$. Before, we let $\mathbb{P}$ here be a full-dimensional convex subset of $\mathbb{P}^{+r_2}_{r_1}$ when $r_1 \notin c(\{r_1, r_2\})$, but now such subset need not exist as $r_2 \notin I$. To compensate for this, since $\delta_2, p_2 \in \mathbb{P}^{+}_{r_2}$ implies that δ_2, p_2 FOSD r_2 due to Lemma 8, we replace the argument "$\mathbb{E}_{r_2} u_{r_1}(x) < \mathbb{E}_{p_1} u_{r_1}(x)$" with "$\mathbb{E}_{r_2} u_{r_1}(x) < \mathbb{E}_{p_2} u_{r_1}(x)$". Everything else goes through as in the proof in Lemma 7, giving us the desired result. $\square$

For $r \in E_2$, given Lemma 8, any increasing utility function $u_r : X \to [0,1]$ will accomplish $c(A) = \arg\max_{p \in A} \mathbb{E}_p u_r(x)$ for all $A \in \mathbb{A}^{r}_{R^{\uparrow}(r)}$. With this freedom, we construct u_r in the following way. Consider for an increasing utility function u_p, the object $\rho^p = \left(\rho^p_2, ..., \rho^p_{|X|-1}\right) \in (0,1)^{|X|-2}$ where $\rho^p_i := \frac{u_p(x_i) - u_p(x_{i-1})}{u_p(x_{i+1}) - u_p(x_{i-1})}$ (that is, ρ^p_i satisfies $u_p(x_i) = \rho^p_i u_p(x_{i+1}) + (1-\rho^p_i) u_p(x_{i-1})$). There is a one-to-one relationship between u_p and ρ^ρ. It is an algebraic exercise to show that $u_p = f \circ u_q$ for some concave and increasing $f : [0,1] \to [0,1]$ if and only if $\rho^p_i \geq \rho^q_i$ for all $i \in \{2, ..., |X|-1\}$. Take $r \in E_2$ and define $\rho^r := \left(\inf_{p \in K}(\rho^p_2), ..., \inf_{p \in K}\left(\rho^p_{|X|-1}\right)\right)$, where $K_r := (I \cup E_2) \cap \{p : rRp\} \subseteq \Delta(X)$, and subsequently construct u_r using ρ^r. It is easy to show that R being risk-consistent implies K_r is nonempty for all $r \in E_2 \backslash \{\delta_b, \delta_w\}$, and so u_r is defined other than when $r \in \{\delta_b, \delta_w\}$.

For the non-generic case where for some $j \in \{b, w\}$ we have $\delta_j \in E_2$ and K_{δ_j} is not defined, this implies $pR\delta_j$ for all $p \in \Delta(X) \setminus \{\delta_b, \delta_w\}$. Then, we define $\rho^{\delta_j}_i = \frac{1}{2}(1) + \frac{1}{2}\sup_{p \in \Delta X \setminus \{\delta_b, \delta_w\}} \rho^p_i$

for all i and construct u_{δ_j} correspondingly. Utility functions indexed by such a δ_j and that by any $p \in \Delta X \backslash \{\delta_j\}$ now satisfy $\rho_i^{\delta_j} \geq \rho_i^p$, with equality when p also is a δ_j falling into this special case (there are at most two of them, δ_b and δ_w).

We now show that for $r_1, r_2 \in \Delta(X)$ where $r_2 R r_1$, we have $\rho^{r_1} \geq \rho^{r_2}$. This is already shown for any $r_1, r_2 \in I \cup E_1$ by Corollary 2. It also is shown for the special cases in the previous paragraph. Henceforth we restrict attention to the remaining cases. Say $r_1 \in E_2$, $r_2 \in I \cup E_1$, but $\rho_i^{r_1} < \rho_i^{r_2}$ for some i. Then $\inf_{p \in K_{r_1}} (\rho_i^p) < \rho_i^{r_2}$, so $\rho_i^p < \rho_i^{r_2}$ for some $p \in K_{r_1}$. However, $p \in K_{r_1}$ implies $r_2 R p$ since R is transitive; since $p \in I \cup E_2$, this contradicts Corollary 2. Say $r_1 \in I \cup E_1$, $r_2 \in E_2$, but $\rho_i^{r_1} < \rho_i^{r_2}$ for some i. Then $\rho_i^{r_1} < \inf_{p \in K_{r_2}} (\rho_i^p)$, so $\rho_i^{r_1} < \rho_i^p$ for all $p \in K_{r_2}$. But $r_1 \in K_{r_2}$, a contradiction. Finally, for $r_1, r_2 \in E_2$ and $r_2 R r_1$, either $K_{r_1} = K_{r_2}$ or $K_{r_1} \subsetneq K_{r_2}$. If the earlier, it is immediately that $\rho^{r_1} = \rho^{r_2}$. If the later, then $\rho_i^{r_1} = \inf_{p \in K_{r_1}} (\rho_i^p) \leq \inf_{p \in K_{r_2}} (\rho_i^p) = \rho_i^{r_2}$ for all i, as desired.

Thus, we have now shown that for any $r_1, r_2 \in \Delta(X)$, $\rho^{r_1} \geq \rho^{r_2}$ whenever $r_2 R r_1$, or equivalently $u_{r_1} = f \circ u_{r_2}$ for some concave and increasing $f : [0,1] \to [0,1]$.

Proof of Proposition 4

Suppose $c(\{p,q\}) = \{p,q\}$. Without loss of generality, either $\arg\max_R \{p, q, p^\alpha q\} = p$ or $\arg\max_R \{p, q, p^\alpha q\} = p^\alpha q$. If the former, then $\arg\max_R \{p,q\} = \arg\max_R \{p, p^\alpha q\} = p$, hence the utility functions used in the two choice problems $c(\{p,q\})$ and $c(\{p, p^\alpha q\})$ are both u_p. It is immediately that, since $p^\alpha q$ is a mixture of p and q, we have $c(\{p, p^\alpha q\}) = \{p, p^\alpha q\}$, and by Transitivity we have $c(\{q, p^\alpha q\}) = \{q, p^\alpha q\}$.[40] If the latter, then $\arg\max_R \{p^\alpha q, q\} = \arg\max_R \{p, p^\alpha q\} = p^\alpha q$. Suppose for contradiction $p \notin c(\{p, p^\alpha q\})$, then $p^\alpha q \notin c(p^\alpha q, q)$ since $p^\alpha q$ is a mixture of p and q.[41] But by Transitivity we would have $c(\{p,q\}) = \{q\}$, a contradiction. We have hence proved the second property of Betweenness in Definition 6. The first property is immediate using this second property, Continuity, and FOSD (the latter two are axioms/implications of AREU).

Proof of Proposition 4

We first prove point 2. Using Proposition 4, we know that indifference curves are linear and do not intersect. Take an arbitrary indifference curve and consider two points p, q on it that lie in the interior of the triangle. Let p' and q' be mean-preserving contractions of p and q such that the line

[40]Since $\sum_{x \in X} p(x) u_p(x) = \sum_{x \in X} q(x) u_p(x) = \sum_{x \in X} [\alpha p(x) + (1-\alpha) q(x)] u_p(x)$.

[41]Since $\sum_{x \in X} p(x) u_{p^\alpha q}(x) < \sum_{x \in X} [\alpha p(x) + (1-\alpha) q(x)] u_{p^\alpha q}(x)$ implies $\sum_{x \in X} [\alpha p(x) + (1-\alpha) q(x)] u_{p^\alpha q}(x) < \sum_{x \in X} q(x) u_{p^\alpha q}(x)$.

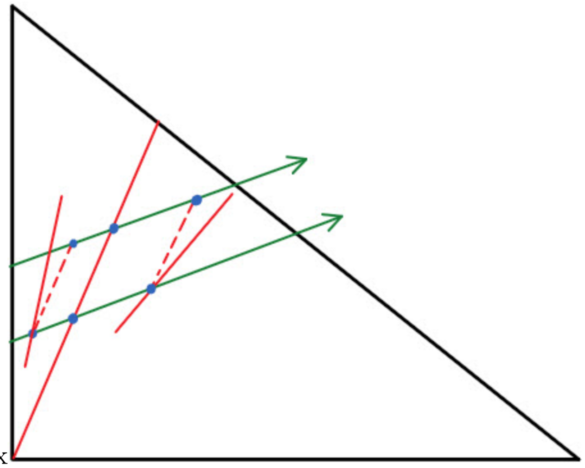

Figure 6.1: Indifference curves fan out when AREU is combined with Transitivity and risk aversion Proposition 4). Arrows correspond to direction of mean-preserving spread.

connecting p', q' is parallel to the line connecting p, q. Since p', q' are mean-preserving contractions, $\min(\{p, q\}, R)$ R $\min(\{p', q'\}, R)$, and so AREU posits that $c(\{p', q'\})$ is explained by a more concave utility function than the one used for $c(\{p, q\})$, corresponding to a weakly steeper indifference curve. Figure 6.1 provides an illustration. Point 3 is proven analogously. The consequence of these unidirectional fanning, along with continuity, rules out the possibility of c being both strictly risk averse and strictly risk loving in this triangle, i.e., point 1 of the proposition.

Proof of Proposition 5

The proof for utility representation is three-fold. First, we show that with Axiom 6 and 7, for each time $r \in T$, the set of all choice problems such that the earliest payment arrives at this time can be explained by a nonempty set of Discounted Utility specifications, where an element of this set is $(\tilde{u}, \delta)$, a utility function and a discount factor. Second, we show that at least one utility function u can be supported for all $r \in T$, and set as δ_r as the corresponding discount factor supporting u for r; this is the more involved part of the proof and uses Axiom 8. Lastly with Axiom 8 again, we show the desired relationship between δ_r and $\delta_{r'}$ for any two r, r'.

By Lemma 1 and Lemma 2, for any $r \in T$ where $r < \bar{t}$, c satisfies WARP and Stationarity over $S := \{A \in \mathcal{A} : \Psi(A) = (\cdot, r)\}$ (S is the collection of choice sets such that the soonest available payment arrives at time r). Take $\epsilon > 0$ such that $r + \epsilon < \bar{t}$. For each (x, r)

Proof of Proposition 7

Fix c. First we show that with Axiom 10 and Axiom 9, for each equity ratio $r \leq 1$, the set of all choice problems where the greatest equity is r can be explained by the maximization of $x + v_r(y)$ for some unique $v_r : \mathbb{R}_+ \to \mathbb{R}$. For each alternative $(x, y) \in X$, the revealed preference relation generated from $c : \mathcal{S} \to \mathcal{A}$, where $\mathcal{S} = \{A \in \mathcal{A} : r(A) = e_{(x,y)} \text{ and } (x, y) \in A\}$, satisfies acyclicity and does not violate quasi-linearity. Combined with Continuity, acyclicity gives us a set of utility functions $u_{(x,y)}(\cdot)$ where for all $A \in \mathcal{S}$, $c(A)$ is the set of maximizers of $u_{(x,y)}(\cdot)$ in A. With non-violation of quasi-linearity, any admissible $u_{(x,y)}(\cdot)$ must be a strictly increasing transformation of $x + v_{(x,y)}(y)$ for some $v_{(x,y)} : \mathbb{R}_+ \to \mathbb{R}$. Otherwise, since for any pair of income distributions $\{(x', y'), (x'', y'')\}$ there are infinitely many shifted copies $\{(x' + a, y'), (x'' + a, y'')\}$ such that $e_{(x'+a,y')}, e_{(x''+a,y'')} \geq e_{(x,y)}$ and $(x, y) \notin c\{(x, y), (x + a, y), (x' + a, y')\}$, a violation of quasi-linearity must occur.

Fix an r, we now show that $v_{(x,y)}$ must coincide for all (x, y) where $e_{(x,y)} = r$. Consider the set of choice problems $\mathcal{S} = \{A \in \mathcal{A} : r(A) = r\}$. Note that c satisfies WARP and Quasi-linearity on $\mathcal{S}$. To see this, take any two choice problems A_1, A_2 in $\mathcal{S}$. For each $i = 1, 2$, there must be an alternative $(x_i, y_i) \in A_i$ such that $e_{(x_i,y_i)} = r$ and $e_{(x',y')} \geq r$ for all other (x', y') in A_i. Consider an income distribution (x^*, y^*) such that $x^* \leq \min\{x_1, x_2\}$ and $y^* \leq \min\{y_1, y_2\}$ and $e_{(x^*,y^*)} = r$. Due to $(x_i, y_i) \in \Psi(A_i \cup \{(x^*, y^*)\})$, Axiom 10 (Fairness Dependence) and Monotonicity (so that (x^*, y^*) is not chosen), $c(A_i) = c(A_i \cup \{(x^*, y^*)\})$. But $(x^*, y^*) \in \Psi(A_1 \cup A_2 \cup \{(x^*, y^*)\})$, so by Axiom 10 again $c(A_1 \cup \{(x^*, y^*)\})$ and $c(A_2 \cup \{(x^*, y^*)\})$, which as established are just $c(A_1)$ and $c(A_2)$, cannot generate a violation of WARP or quasi-linearity. Consequently, $v_{(x,y)}$ must coincide for all (x, y) such that $e_{(x,y)} = r$.

Finally we show that for all $r > r'$, $v_r(y) - v_r(y') \geq v_{r'}(y) - v_{r'}(y')$ for all $y > y'$ (reminder: higher r implies greater attainable equity). Suppose not, our goal is to substantiate a contradiction of Axiom 11 in the choice correspondence. Fix any $y, y' \in \mathbb{R}_+$ such that $y > y'$. Define $\tilde{v}_r = v_r(y) - v_r(y')$ and $\tilde{v}_{r'} = v_{r'}(y) - v_{r'}(y')$. We want to show $\tilde{v}_r \geq \tilde{v}_{r'}$. Suppose for contraction this is not true, let z be any value such that $\tilde{v}_r < z < \tilde{v}_{r'}$. Find a number b such that $\max\{(z - b)/y, (z - b)/y'\} < r'$ and $(z - b) \geq 0$, which is clearly possible for fixed r', y, and y' since $r' > 0$ and b can be arbitrarily close to z from below. Define $x := z - b$, $x' := 2z - b$, let (x_0, y_0) be some income distribution such that $e_{(x_0,y_0)} = r'$ and $x_0 < x$, $y_0 < y$, and similarly let (x_1, y_1) be some income distribution such that $e_{(x_1,y_1)} = r$ and $x_1 < x$, $y_1 < y$ (these are always possible). Consider the set $A := \{(x, y), (x', y'), (x_0, y_0)\}$. $c(A)$ comes from maximizing the utility function $\hat{x} + v_{r'}(\hat{y})$, and (x_0, y_0) will never be chosen since it is strictly less than (x, y) in each component. Likewise, $c(A \cup \{(x_1, y_1)\})$ comes from maximizing the utility function $\hat{x} + v_r(\hat{y})$ and both (x_0, y_0), (x_1, y_1) will not be chosen. We essentially introduced reference points that won't be chosen, forcing the choice to be between (x, y) and (x', y'). Now note that the way z was obtained gives us $\tilde{v}_r + z < 2z < \tilde{v}_{r'} + z$, and so $\tilde{v}_r + z - b < 2z - b < \tilde{v}_{r'} + z - b$. The first and second inequality are equivalent to $v_r(y) + x < v_r(y') + x'$ and $v_{r'}(y) + x > v_{r'}(y') + x'$ respectively. Finally, the latter gives us $c(A) = \{(x, y)\}$ (where $(x_0, y_0) \in A$) and the former gives us $c(A \cup \{(x_1, y_1)\}) = \{(x', y')\}$; since $A \subset A \cup \{(x_1, y_1)\}$, this is a contradiction of Axiom 11. This establishes $v_r(y) - v_r(y') \geq v_{r'}(y) - v_{r'}(y')$ for all $y > y'$, for all $r > r'$.

Remark. Some proofs are not included, and can be requested from me (rc@xzlim.com).

References

Ainslie, G. (1992). *Picoeconomics: The strategic interaction of successive motivational states within the person*. Cambridge University Press.

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

Andreoni, J. & Sprenger, C. (2011). *Uncertainty equivalents: Testing the limits of the independence axiom*. Technical report, National Bureau of Economic Research.

Aumann, R. J. & Serrano, R. (2008). An economic index of riskiness. *Journal of Political Economy*, 116(5), 810–836.

Bardsley, N. (2008). Dictator game giving: altruism or artefact? *Experimental Economics*, 11(2), 122–133.

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations Research*, 30(5), 961–981.

Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations Research*, 33(1), 1–27.

Benhabib, J., Bisin, A., & Schotter, A. (2010). Present-bias, quasi-hyperbolic discounting, and fixed costs. *Games and Economic Behavior*, 69(2), 205–223.

Bleichrodt, H. & Schmidt, U. (2002). A context-dependent model of the gambling effect. *Management Science*, 48(6), 802–812.

Bolton, G. E. & Ockenfels, A. (2000). Erc: A theory of equity, reciprocity, and competition. *American economic review*, 90(1), 166–193.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience theory of choice under risk. *Quarterly Journal of Economics*, 127(3), 1243–1285.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, 121(5), 803–843.

Camerer, C. (1995). Individual decision making. *Handbook of experimental economics*.

Cerreia-Vioglio, S., Dillenberger, D., & Ortoleva, P. (2015). Cautious expected utility and the certainty effect. *Econometrica*, 83(2), 693–728.

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

Chew, S. H. (1983). A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the allais paradox. *Econometrica*, (pp. 1065–1092).

Chew, S. H. & Epstein, L. G. (1989). A unifying approach to axiomatic non-expected utility theories. *Journal of Economic Theory*, 49(2), 207–240.

Chew, S. H., Epstein, L. G., & Segal, U. (1991). Mixture symmetry and quadratic utility. *Econometrica: Journal of the Econometric Society*, (pp. 139–163).

Cox, J. C., List, J. A., Price, M., Sadiraj, V., & Samek, A. (2016). *Moral costs and rational choice: Theory and experimental evidence*. Technical report, National Bureau of Economic Research.

Dana, J., Cain, D. M., & Dawes, R. M. (2006). What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and human decision Processes*, 100(2), 193–201.

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33(1), 67–80.

de Clippel, G. & Rozen, K. (2014). Bounded rationality and limited datasets. *unpublished*.

Dean, M., Kıbrıs, Ö., & Masatlioglu, Y. (2017). Limited attention and status quo bias. *Journal of Economic Theory*, 169, 93–127.

Dekel, E. (1986). An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom. *Journal of Economic theory*, 40(2), 304–318.

Dillenberger, D. (2010). Preferences for one-shot resolution of uncertainty and allais-type behavior. *Econometrica*, 78(6), 1973–2004.

Fehr, E. & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, 114(3), 817–868.

Fehr, E. & Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism–experimental evidence and new theories. *Handbook of the economics of giving, altruism and reciprocity*, 1, 615–691.

Fishburn, P. C. (1983). Transitive measurable utility. *Journal of Economic Theory*, 31(2), 293–317.

Fishburn, P. C. & Rubinstein, A. (1982). Time preference. *International economic review*, (pp. 677–694).

Frederick, S., Loewenstein, G., & O'donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of economic literature*, 40(2), 351–401.

Freeman, D. (2016). Revealing naïveté and sophistication from procrastination and preproperation. *unpublished*.

Gul, F. (1991). A theory of disappointment aversion. *Econometrica*, (pp. 667–686).

Gul, F. & Pesendorfer, W. (2001). Temptation and self-control. *Econometrica*, 69(6), 1403–1435.

Gul, F., Pesendorfer, W., et al. (2006). The revealed preference implications of reference dependent preferences. *manuscript, Princeton University.*

Herne, K. (1999). The effects of decoy gambles on individual choice. *Experimental Economics*, 2(1), 31–40.

Kagel, J. H. & Roth, A. E. (2016). *The handbook of experimental economics*, volume 2. Princeton university press.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic perspectives*, 5(1), 193–206.

Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of choice under risk. *Econometrica*, 47, 263–291.

Kıbrıs, Ö., Masatlioglu, Y., & Suleymanov, E. (2018). A theory of reference point formation. *unpublished*.

Korenok, O., Millner, E. L., & Razzolini, L. (2014). Taking, giving, and impure altruism in dictator games. *Experimental Economics*, 17(3), 488–500.

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

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

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 112(2), 443–478.

Lian, C., Ma, Y., & Wang, C. (2017). Low interest rates and risk taking: Evidence from individual investment decisions. *unpublished*.

Lipman, B. L., Pesendorfer, W., et al. (2013). Temptation. In *Advances in Economics and Econometrics: Tenth World Congress*, volume 1 (pp. 243–288).: by D. Acemoglu, M. Arellano, and E. Dekel, Cambridge University Press.

List, J. A. (2007). On the interpretation of giving in dictator games. *Journal of Political economy*, 115(3), 482–493.

Loewenstein, G. & Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics*, 107(2), 573–597.

Loomes, G. & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, 92(368), 805–824.

Loomes, G. & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *Review of Economic Studies*, 53(2), 271–282.

Machina, M. J. (1982). " expected utility" analysis without the independence axiom. *Econometrica*, (pp. 277–323).

Manzini, P. & Mariotti, M. (2007). Sequentially rationalizable choice. *American Economic Review*, 97(5), 1824–1839.

Masatlioglu, Y., Nakajima, D., & Ozbay, E. Y. (2012). Revealed attention. *American Economic Review*, 102(5), 2183–2205.

Masatlioglu, Y. & Ok, E. A. (2005). Rational choice with status quo bias. *Journal of Economic Theory*, 121(1), 1–29.

Masatlioglu, Y. & Ok, E. A. (2013). A canonical model of choice with initial endowments. *Review of Economic Studies*, 81(2), 851–883.

Masatlioglu, Y. & Raymond, C. (2016). A behavioral analysis of stochastic reference dependence. *American Economic Review*, 106(9), 2760–82.

Nelson, Jr., W. R. (2002). Equity or intention: it is the thought that counts. *Journal of Economic Behavior & Organization*, 48(4), 423–430.

Ok, E. A. & Benoît, J.-P. (2007). Delay aversion. *Theoretical Economics*, 2(1), 71–113.

Ok, E. A., Ortoleva, P., & Riella, G. (2015). Revealed (p) reference theory. *American Economic Review*, 105(1), 299–321.

Ortoleva, P. (2010). Status quo bias, multiple priors and uncertainty aversion. *Games and Economic Behavior*, 69(2), 411–424.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4), 323–343.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American economic review*, (pp. 1281–1302).

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of economic literature*, 38(2), 332–382.

Sutter, M. (2007). Outcomes versus intentions: On the nature of fair behavior and its development with age. *Journal of Economic Psychology*, 28(1), 69–78.

Wakker, P. & Deneffe, D. (1996). Eliciting von neumann-morgenstern utilities when probabilities are distorted or unknown. *Management science*, 42(8), 1131–1150.