Revealed (P)Reference Theory *

Efe A. Ok[†] Pietro Ortoleva[‡] Gil Riella[§]

August 1, 2012

Abstract

The goal of this paper is to develop, axiomatically, a revealed preference theory of reference-dependent choice behavior. Instead of taking the "reference" for an agent as exogenously given in the description of a choice problem, we suitably relax the Weak Axiom of Revealed Preference to obtain, endogeneously, the existence of reference alternatives as well as the structure of choice behavior conditional on those alternatives. We also elaborate on how to make welfare evaluations on the basis of the resulting individual decision-making model, and introduce the notion of "cumulative reference."

JEL Classification: D11, D81.

Keywords: Revealed Preference, Reference-Dependence, Endogeneous Reference Formation, Multi-Utility, Attraction Effect

*An older version of this manuscript was previously circulated under the title "Rational Choice with Endogeneously Determined Reference Points: The Case of the Attraction Effect." Various comments of Heski Bar-Isaac, Jean-Pierre Benoît, Douglas Bernheim, Geoffroy de Clippel, Federico Echenique, Kfir Eliaz, Alessandro Lizzeri, Yusufcan Masatlioglu, Erik Maskin, Debraj Ray, Todd Sarver, Tomasz Stralecki, Leeat Yariv, and the participants at seminars at Brown, Johns Hopkins, Harvard, Michigan, Northwestern, NYU, Oxford, UCLA, UnB, UPenn, have contributed to this work; we gratefuly acknowledge our intellectual debt to them. We are also grateful to the support of the C. V. Starr Center for Applied Economics at New York University.

[†] *Corresponding Author*: Department of Economics, New York University, 19 West 4th Street, New York, NY 10012. E-mail: efe.ok@nyu.edu.

[‡]Division of the Humanities and Social Sciences, California Institute of Technology.

[§]Department of Economics, Universidade de Brasília.

1 Introduction

1.1 Motivation

The canonical model of rational choice maintains that (i) an individual has a well-defined manner of ranking alternatives according to their desirability (independently of any particular choice problem that she might face), and (ii) among *any* collection of feasible alternatives, she chooses the item that she ranks the highest. Despite its various advantages, among which are its unifying structure, universal applicability, tractability, and predictive abilities, this model has long been scrutinized on the basis of its descriptive weaknesses. Indeed, experimental and market evidence points persistently to certain types of choice behavior that are inconsistent with the premises of rational choice theory – human decision processes appear to be more intricate than this theory allows for. In particular, the consensus now seems to be that the presence of certain types of (observable) choice prospects – status quo choices, endowments, default options, etc. – may well affect individual choice behavior.[1]

Perhaps more surprisingly, there is now good evidence that suggests that *reference dependence* may feature in the choice behavior of a decision maker even in choice situations in which *no alternative is predesignated as a reference option*. In certain contexts, it is observed that a seemingly ordinary feasible choice item may, for some reason or another, act as a reference for a decision-maker, thereby affecting her choice behavior. While the behavior of the agent may seem to an outsider hard to "justify," she may actually be acting in a predictable manner relative to her subjectively determined (hence unobservable) reference point. In light of this evidence, the objective of this paper is to modify the classical rational choice theory in a way to account for choice behavior that may depend on such situations.

Among the instances of reference-dependent behavior in contexts where a reference point is not exogeneously identified, one that has received much attention in the literature is the famous *attraction effect phenomenon* (also known as the *asymmetric dominance effect*). Discovered first by Huber, Payne and Puto (1982), and then corroborated by numerous studies in psychology and marketing, this effect may be described as the phenomenon in which, given a choice set of two feasible alternatives, the addition of a third alternative that is clearly inferior (or else strictly dominated) to *one* of the existing alternatives (but not to the other) may induce a shift of preference towards the item that dominates the new alternative.

[1]The experimental and market literature that documents violations of the standard rationality paradigm is too extensive to be discussed here: see Camerer (1995) and Rabin (1998) for surveys. Among the many documented violations, the status quo bias phenomenon has, in particular, received quite a bit of attention in the literature. Among the individual choice models that are developed to represent this phenomenon are those that are based on *loss aversion* (Tversky and Kahneman (1991)), *multi-utility* (Masatlioglu and Ok (2005) and Sagi (2006)), and *multiple priors* (Ortoleva (2010)).

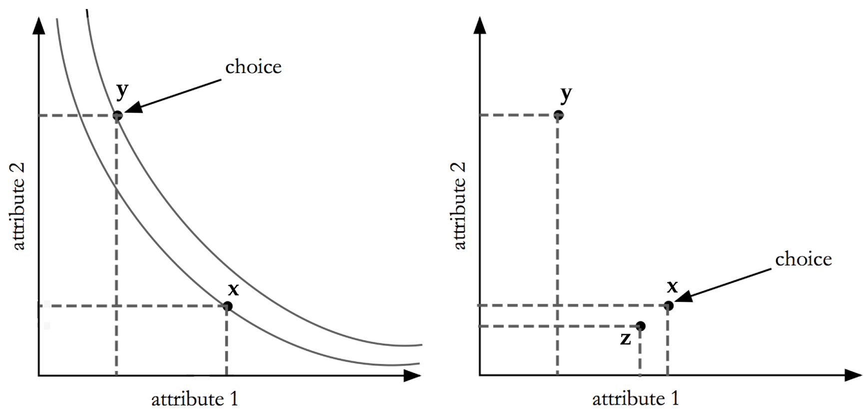

Figure 1

To illustrate, consider two alternatives, x and y, in a world in which each alternative is characterized by exactly two attributes (such as price and quality). Suppose, as shown in Figure 1, y is better (resp., worse) than x relative to the second (resp., first) attribute. Suppose also that the agent chooses y when only x and y are available. Now suppose a third (decoy) alternative z becomes feasible; this alternative is inferior to x relative to *both* attributes, but it is still better than y with respect to the first attribute (Figure 1). The attraction effect phenomenon corresponds to the situation in which the agent chooses x from the set $\{x, y, z\}$, which violates the standard formulation of rationality. The idea is that, somehow, the asymmetrically dominated alternative z acts as a reference for the agent in the problem $\{x, y, z\}$, making the choice prospect that is unambiguously better than z, namely x, more attractive than it actually is in the absence of z.[2,3]

[2]To hint at the potential consequences of the attraction effect in the market place, let us mention the findings of Doyle *et al.* (1999) whose field experiments took place in a local grocery store. First, the authors recorded the weekly sales of the Brands X (Spar (420 g) baked beans) and Y (Heinz (420 g) baked beans) in the store, and observed that Brand X had gotten 19% of the sales in a given week, and Y the rest, even though Brand X was cheaper. Doyle *et al.* then introduced a third Brand Z (Spar (220 g) baked beans) to the supermarket, which was identical to Brand X in all attributes (including the price) except that the size of Brand Z was visibly smaller. The idea is that Brand Z was asymmetrically dominated; it was dominated by X but not by Y. In accordance with the attraction effect, the authors observed in the following week that the sales of Brand X had increased to 33% (while nobody had bought Brand Z).

[3]The attraction effect is demonstrated in the contexts of choice over political candidates (Pan, O'Curry and Pitts (1995)), choice over risky alternatives (Wedell (1991) and Herne (1997)), medical decision-making (Schwartz and Chapman (1999) and Redelmeier and Shafir (1995)), investment decisions (Schwarzkopf (2003)), job candidate evaluation (Highhouse (1996), Slaughter, Sinar and Highhouse (1999), and Slaughter (2007)), and contingent evaluation of environmental goods (Bateman, Munro and Poe (2008)). While most of the experimental findings in this area are through questionnaire studies, some authors have confirmed the attraction effect also through experiments with incentives (Simonson and Tversky (1992) and Herne (1999)).

In the psychological literature, it is argued that the attraction effect may be due to simplifying decision heuristics (Wedell (1991)), or due to one's need to justify his/her decisions (Simonson (1989), and Simonson and Nowlis (2000)), or due to the ambiguity of the information about the attributes of products (Ratneshwar, Given the amount of evidence on the presence of the attraction effect, and its obvious importance for marketing, it seems desirable to provide a universally applicable model of individual decision-making that incorporates this phenomenon. For instance, to study the optimal choices of a firm who wishes to exploit the attraction effect in a market environment, and hence to assess the implications of this phenomenon for market segmentation, one would need to utilize this sort of a model. Similarly, to examine the potential implications of this phenomenon for the determination of political candidates and platforms, presentation of portfolios, etc., one needs a general model of individual choice which allows for the presence of this effect. In a nutshell, the primary goal of the present paper is to develop a model of reference-dependent choice that is suitable for this purpose, and one that deviates from the standard rational choice model in a parsimonious manner.[4]

1.2 Overview

The primitives of our construction are identical to the classical choice theory, namely, a collection of feasible sets of alternatives (choice problems), and a choice correspondence defined on this collection.[5] We study choice correspondences that violate the classical formulation of rationality, namely, the *Weak Axiom of Revealed Preference* (WARP), but we focus only on those violations that could be ascribed to the presence of a certain form of (endogenous) reference effect. To this end, we first introduce some terminology to classify the violations of WARP (Section 2.2.1). More important, we develop a method which extends that of the classical theory of revealed preference to allow not only for revealed *preferences*, but also for revealed *references* (Section 2.2.2). Put differently, this method aims at identifying from an individual's (observable) choice behavior those choice items that may act as (unobservable) reference points for her in some choice situations as well as one's preference relation that

Shocker and Stewart (1987) and Mishra, Umesh and Stem (1993)), or due to the comparative evaluation of goods (Simonson and Tversky (1992) and Bhargava, Kim and Srivastava (2000)), or dynamic formation of preferences in a dominance-seeking manner (Ariely and Wallsten (1995)), or evolutionary pressures (Shafir, Waite and Smith (2002)). In the marketing literature see, *inter alia*, Burton and Zinkhan (1987), Lehmann and Pan (1994), Sivakumar and Cherian (1995), Sen (1998), Kivetz (1999), and Doyle, O'Connor, Reynolds and Bottomley (1999).

[4]There are a few recent studies that investigate the modeling of the attraction effect phenomenon. For instance, Lombardi (2009), Gerasimou (2012) and de Clippel and Eliaz (2012) have characterized models that incorporate this phenomenon to some extent, and Kamenica (2008) has studied the issue by means of an information-based choice model in a particular market environment. However, as we discuss at length in Section 5.3, the first three models capture only a special case of the attraction effect, and the latter work does not provide a general model of choice that is applicable outside of certain market contexts.

[5]In contrast with the experimental literature discussed before, our model does not prescribe an alternative in terms of a given set of attributes: while the exogeneous description of a given alternative might include such attributes, it might also not contain any. (For example, the alternatives could simply be described as objects such as "ice cream," "luxury car A," "presidential candidate B," and so on.) In fact, we will *derive* in this paper the standard "multi-attribute" choice model as a consequence of behavioral postulates.

represents her tastes at large.[6]

Our main contribution is the derivation of a reference-dependent choice model on the basis of (falsifiable) behavioral postulates (Sections 2.3 and 3.1). This model can be described as follows. The decision-maker has a *utility function* over alternatives, a *reference map* $\mathbf{r}$ that maps any given feasible set S to either nothing or a reference alternative $\mathbf{r}(S)$ in that set, and an *attraction correspondence* Q that maps any given alternative z to those prospects the agent is attracted to from the perspective of z. By way of interpretation, we think of $Q(z)$ as the set of all alternatives whose value seem elevated to the agent when compared to z (perhaps because everything in $Q(z)$ dominates z with respect to all attributes this person deems relevant to choice and no such dominance works for alternatives outside $Q(z)$).

In a given feasible set S, a decision maker may or may not single out an alternative as a "reference point." If she does not discern such an alternative – we denote this situation by writing $\mathbf{r}(S) = \diamondsuit$ – her behavior is entirely standard: she chooses any one item in S that maximizes her utility function. On the other hand, if she identifies a reference alternative in S – in this case we have $\mathbf{r}(S) \in S$ – rather than maximizing her utility function over the entire S, she focuses on those members of S which the presence of $\mathbf{r}(S)$ "attracts" her to. Put differently, once $\mathbf{r}(S)$ is determined as her reference point, the decision-maker concentrates only on those outcomes in S that belong to the mental attraction region $Q(\mathbf{r}(S))$ generated by that alternative. That is, in this case, she views her choice problem effectively not as S, but $S \cap Q(\mathbf{r}(S))$. Under this psychological constraint, however, she acts rationally, that is, she finalizes her choice by maximizing her utility function over $S \cap Q(\mathbf{r}(S))$.

This is a (boundedly rational) reference-dependent choice model that allows for the presence of the attraction effect. The choice behavior exhibited in Figure 1 is, for instance, duly consistent with this model. Specifically, the model would "explain" this behavior by saying that it is "as if" the agent views z as a reference point in $\{x, y, z\}$, and the attraction region $Q(z)$ includes x but not y. So, in the problem $\{x, y\}$, where there is no alternative to "use" as a reference for the agent to contrast x and y, the choice is y (as the utility of y exceeds that of x). But when faced with the problem $\{x, y, z\}$, the agent reduces this problem in her mind to $\{x, y, z\} \cap Q(z) = \{x, z\}$, and consequently, chooses x. (See Figure 2.)

We also show how to impose further discipline to our representation by endowing the attraction correspondence Q with some structure (Section 3.4). Put precisely, we show what sort of behavior would allow us to think as if the agent were endowed not only with one utility

[6]Needless to say, "reference dependence" is a multifarious concept. A "reference" may take the form of one's unattainable aspiration, or one that is attainable but is desirable only according to a social norm. By contrast, the notion of "reference" that we focus on here relates to undesirable choice prospects that, when attainable, alter the decision maker's views about the relative appeal of *other* feasible choice options. This view point will be clarified in Section 2.

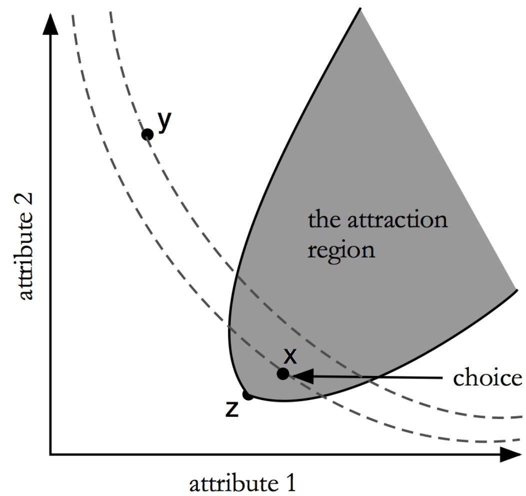

Figure 2

function, but with several of them. (For example, each of these utilities could represent one of the attributes of the goods under consideration.) In the absence of a reference point the agent maximizes some aggregation of these utilities. When she discerns a reference option, however, she focuses only on those items that are superior to this option with respect to each of her utilities. (That is, the attraction region $Q(x)$ of x is found as the set of those options that dominate x in terms of all utility functions.) This way we obtain a representation which embodies the spirit of the attraction effect, but one in which the attributes of the choice prospects are determined endogeneously by the behavior of the agent.

Subsequently, after the formal presentation of the underlying behavioral postulates and our choice model, we discuss how one may carry out a welfare analysis in terms of this model (Section 4), and investigate some potential applications and shortcomings of it thereof (Sections 5.1 and 5.2). Finally, we discuss the related literature (Section 5.3) and conclude with an appendix that contains the proofs of our main results.

2 Reference-Dependent Choice

2.1 Preliminaries

We work with an arbitrarily fixed separable metric space X, which we think of as the universal set of available alternatives. (For concreteness, the reader may view X as an arbitrary finite set or a subset of a Euclidean space.) We let $\mathfrak{X}$ stand for the set of all nonempty compact subsets of X. The elements of $\mathfrak{X}$ are viewed as feasible sets that a decision maker may need to choose an alternative from; they are henceforth referred to as *choice problems*.

A set-valued map $\mathbf{c} : \mathfrak{X} \rightrightarrows X$ is said to be a **choice correspondence** on $\mathfrak{X}$ if $\emptyset \neq \mathbf{c}(S) \subseteq S$ for every $S \in \mathfrak{X}$. (If $\mathbf{c}$ is single-valued, we consider it as a function from $\mathfrak{X}$ into X, and refer to it as a **choice function** on $\mathfrak{X}$.)[7] We say that this correspondence is **continuous** if for any two convergent sequences (x_m) and (y_m) in X, we have $\lim x_m \in \mathbf{c}\{\lim x_m, \lim y_m\}$ provided that $x^m \in \mathbf{c}\{x^m, y^m\}$ for each m. Clearly, when X is finite, every choice correspondence on $\mathfrak{X}$ is continuous.

2.2 Some Foundations for Reference-Dependent Choice

2.2.1 Violations of WARP

The classical assumption of rational choice theory is the so-called *Weak Axiom of Revealed Preference*. Following the formulation of Arrow (1959), we state this property as follows:

Weak Axiom of Revealed Preference (WARP). For every $S \in \mathfrak{X}$ and $T \subseteq S$ with $\mathbf{c}(S) \cap T \neq \emptyset$, we have $\mathbf{c}(S) \cap T = \mathbf{c}(T)$.

The *fundamental theorem of revealed preference* says that for any choice correspondence $\mathbf{c}$ on $\mathfrak{X}$ that satisfies WARP there exists a complete preorder (that is, a complete and transitive binary relation) on X such that, for any $S \in \mathfrak{X}$, the choices from S are the maximum elements in S according to that preorder. This result is usually viewed as a foundation for rational choice theory, for it maintains that, as long as one's choices do not violate WARP, they can be modeled "as if" they stem from the maximization of a complete preference relation.

To generalize the classical choice theory to allow for endogenous reference dependence, therefore, one has to relax WARP. To this end, we first give a name to those choice problems on which the choice behavior of the agent is observed to violate WARP.

Definition. Let $\mathbf{c}$ be a choice correspondence on $\mathfrak{X}$. For any given $S \in \mathfrak{X}$, we say that S is **c-awkward** if there exist some doubleton subset T of S with $c(S) \cap T \neq \emptyset$ and $c(S) \cap T \neq c(T)$.

In words, we say that a choice problem S is "awkward" if the choice behavior of the decision-maker in the context of S is not compatible with how she ranks alternatives in pairwise comparisons. This happens when either (i) the agent chooses an alternative x from S which is revealed to be less desirable than another alternative y in S when she compares these alternatives in isolation, or (ii) the agent deems two alternatives x and y equally desirable when she compares them against each other, even though she reveals a

[7] *Notation.* For enumerated finite sets such as $\{x_1, ..., x_k\}$, we write $\mathbf{c}\{x_1, ..., x_k\}$ instead of $\mathbf{c}(\{x_1, ..., x_k\})$ for simplicity. Similarly, if $\mathbf{c}$ is a choice function, we write $x = \mathbf{c}(S)$ instead of $\{x\} = \mathbf{c}(S)$ for any $S \in \mathfrak{X}$.

strict preference for x over y in the context of S. (Clearly, if $\mathbf{c}$ is a choice *function*, then S is $\mathbf{c}$-awkward iff there are two alternatives x and y in S such that $x = \mathbf{c}(S)$ and $y = \mathbf{c}\{x, y\}$.)

2.2.2 Revealed References

By far the most forthcoming arguments against WARP are based on the tendency of decision-makers to form their choices by using certain alternatives in a feasible set as *references*. For instance, the violation of WARP in the case of the attraction effect phenomenon appears due to such reference-dependent decision making. To study such violations systematically, we now introduce two notions that will help deducing the role of reference alternatives (if any) from one's choice behavior. (This is reminiscent of how one deduces the preferences of a decision-maker from her choice behavior in revealed preference theory.)

Consider a pair of alternatives x and z in X, and suppose there is some $y \in X$ such that x is not chosen over y in the pairwise context (that is, $\{y\} = \mathbf{c}\{x, y\}$), but x is deemed choosable in the problem $\{x, y, z\}$ (that is, $x \in \mathbf{c}\{x, y, z\}$). In other words, x alone is not able to "beat" y, but it does so with the "help" of z. Similarly, if y is deemed choosable from $\{x, y\}$ (that is, $y \in \mathbf{c}\{x, y\}$) while x, but not y, is deemed choosable from $\{x, y, z\}$ (that is $\{x, y\} \cap \mathbf{c}\{x, y, z\} = \{x\}$), it would be reasonable to conclude that the appeal of x against y is somewhat "enhanced" when z is also present in the choice problem. This prompts:

Definition. Let $\mathbf{c}$ be a choice correspondence on $\mathfrak{X}$ and $x, z \in X$. We say that z is a **revealed reference** for x relative to $\mathbf{c}$ (or more simply, a **revealed c-reference** for x) if there is an alternative $y \in X$ such that either (i) $x \in \mathbf{c}\{x, y, z\} \backslash \mathbf{c}\{x, y\}$ or (ii) $y \in \mathbf{c}\{x, y\}$ and $\{x, y\} \cap \mathbf{c}\{x, y, z\} = \{x\}$.

In the case of a choice *function* $\mathbf{c}$, this definition says simply that z is a revealed $\mathbf{c}$-reference for x iff there is a $y \in X$ such that $y = \mathbf{c}\{x, y\}$ and $x = \mathbf{c}\{x, y, z\}$.

2.2.3 Potential References

The notion of revealed reference is somewhat demanding, for it requires that there be a choice problem from which one element is actually *shown* to "help" the other to be chosen. One can think of a related notion according to which the introduction of some alternative z in a feasible set does not "reduce the appeal of x" even though it may not ensure x to be chosen. It is as if z does not really "help" x, but it does not "help" anything else *against* x either. That is to say, if x is chosen against some y when z was not present, then x must continue to be chosen when z is added to $\{x, y\}$, unless z is the unique choice in the new choice problem. Similarly, if y is not chosen against x, then the addition of z should not render y choosable against x. If this is the case, we say that z is a *potential reference* for x.

Definition. Let $\mathbf{c}$ be a choice correspondence on $\mathfrak{X}$ and $x, z \in X$. We say that z is a **potential reference** for x relative to $\mathbf{c}$ (or more simply, a **potential c-reference** for x) if, for every $y \in X$ such that $\mathbf{c}\{x, y, z\} \neq \{z\}$,

$$x \in \mathbf{c}\{x, y\} \text{ implies } x \in \mathbf{c}\{x, y, z\} \quad \text{and} \quad y \notin \mathbf{c}\{x, y\} \text{ implies } y \notin \mathbf{c}\{x, y, z\}.$$

In the case of a choice *function* $\mathbf{c}$, then, z is a potential $\mathbf{c}$-reference for x iff, for every y in X, $x = \mathbf{c}\{x, y\}$ and $z \neq \mathbf{c}\{x, y, z\}$ imply $x = \mathbf{c}\{x, y, z\}$.

We emphasize that the notion of potential reference is a fairly weak one. For example, for a rational decision maker (that is, when $\mathbf{c}$ satisfies WARP), z is a potential $\mathbf{c}$-reference for x, for any $x, z \in X$. On the other hand, for such a choice correspondence $\mathbf{c}$, no element of X qualifies as a revealed $\mathbf{c}$-reference for any alternative in X.

2.3 Behavioral Foundations

Here we outline the axiomatic structure of the present analysis. Our first postulate is basic. Recall that our goal here is not to investigate violations of WARP in general, but rather to study those that are due to endogenous reference effects. Consequently, we wish to model those decision-makers who act rationally in the traditional sense in the absence of referential considerations. In particular, as there could be no such effects in pairwise choice situations – in such problems there is no room for a third alternative to 'help' another – we would like to posit the following property:

No-Cycle Condition (No-Cycle). For every $x, y, z \in X$, if $x \in \mathbf{c}\{x, y\}$ and $y \in \mathbf{c}\{y, z\}$, then $x \in \mathbf{c}\{x, z\}$.

No-Cycle separates our analysis from that of most of the literature on violations of WARP, where the presence of cycles from sets of two elements is an essential component of the analysis.[8,9]

[8]Our imposition of No-Cycle also highlights the fact that the notion of "reference-dependence" we consider here is not related to, say, the status quo phenomenon. The latter notion would necessitate a default option to be thought of as a "reference" in dichotomous choice problems as well. To reiterate, our focus is on the notion of reference alternatives that are not desirable in themselves, but rather, affect the comparative desirability of other alternatives. Thus, the reference notion becomes meaningful in the present setup only when, in addition to the alternative designated as the reference point, there are at least two alternatives in the choice situation at hand. (More on this in Section 2.5.)

[9]Recently, Manzini and Mariotti (2010) have performed an experiment to test how people violate WARP.

We next turn to restrict the way reference effects can take place by connecting the notions of revealed and potential reference. Formally, there is no reason for a revealed reference to be a potential reference relative to an arbitrary choice correspondence. However, the very motivation of these concepts suggest that those choice correspondences for which such an implication is true are of particular interest. Consequently, we posit next that if z is a revealed reference for x, that is, z is shown to actually "help" x be chosen in *some* situation, then z must also be a potential reference for x, that is, z cannot "help" another element *against* x in some other situation. This is but asking for coherence of the evaluation of choice items by using a reference alternative.[10]

Reference Coherence (RCoh). For every $x, z \in X$, if z is a revealed $\mathbf{c}$-reference for x, then z is a potential $\mathbf{c}$-reference for x.

We now come to our most important postulate, *Reference Consistency.* To formulate this, we recall that a collection $\mathcal{T}$ of subsets of a set S is said to be a **cover** of S if the union of the members of $\mathcal{T}$ equals S, that is, $\bigcup\{T : T \in \mathcal{T}\} = S$. (Notice that $\mathcal{T}$ is not necessarily a partition of S; its members are allowed to overlap.) In turn, given a choice correspondence $\mathbf{c}$ on $\mathfrak{X}$, and a choice problem $S \in \mathfrak{X}$, we say that a collection $\mathcal{T}$ of subsets of S is a $\mathbf{c}$**-cover** of S if $\mathcal{T}$ is a cover of S such that each member of $\mathcal{T}$ contains at least one choice from S with respect to $\mathbf{c}$, that is, $\mathbf{c}(S) \cap T \neq \emptyset$ for each $T \in \mathcal{T}$. Loosely speaking, our next postulate states that, for a $\mathbf{c}$-awkward choice problem S, WARP should hold for at least one member of any given $\mathbf{c}$-cover of S.

Reference Consistency (RCon). For any $\mathbf{c}$-awkward $S \in \mathfrak{X}$ and $\mathbf{c}$-cover $\mathcal{T}$ of S, we have $\mathbf{c}(T) = \mathbf{c}(S) \cap T$ for some $T \in \mathcal{T}$.

RCon is a form of restriction on the kinds of violations of WARP we allow for. While it may not be evident at first stroke, this property is actually quite reasonable to impose on choice behavior that departs from rationality only due to reference-dependence. To see this, consider a decision-maker (with choice correspondence $\mathbf{c}$) who acts in a reference-dependent manner in the context of some choice problem S. Let $\{A, B\}$ be a $\mathbf{c}$-cover of S. Since $A \cup B = S$, the reference point that the agent uses in S must belong to either A or B

They have classified the violations of WARP in two categories: *pairwise inconsistencies* - choices violate the No Cycle Condition - and *menu effects* - choices from doubletons do not induce the choices from larger menus. Their conclusion was that menu effects are the main reason for violations of WARP.

[10] Implicit in this formulation is the idea that "being a reference" is an all-or-nothing phenomenon. Loosely speaking, if z is a reference for both x and y, we rule out the possibility that z is 'more of a reference' for x than for y. On the one hand, this simplifies the revealed preference theory that we are about to sketch. On the other, we are not aware of any evidence that motivates the modeling of the notion of "being a reference" as a graded phenomenon.

(or both). Suppose it belongs to A. Then A contains both the reference point used in S and at least one of the choices of the agent from S (because $\mathbf{c}(S) \cap A \neq \emptyset$). Therefore, it stands to reason that the reference point should retain its power in this smaller set A, and lead the agent once again to those choices (from S) that are still available in A, that is, $\mathbf{c}(A) = \mathbf{c}(S) \cap A$. This is the very gist of RCon.

Remark. In the case where X is finite (or more generally, when $\mathfrak{X}$ consists of all nonempty finite subsets of X), the statement of RCon can be simplified considerably, for in that case we can work with doubleton $\mathbf{c}$-covers. That is, in that case, $\mathbf{c}$ satisfies RCon iff, for any $\mathbf{c}$-awkward $S \in \mathfrak{X}$ and any $A, B \subseteq S$ with $A \cup B = S$ and $\mathbf{c}(S) \cap A \neq \emptyset \neq \mathbf{c}(S) \cap B$, we have either $\mathbf{c}(A) = \mathbf{c}(S) \cap A$ or $\mathbf{c}(B) = \mathbf{c}(S) \cap B$. This observation, which may be proved by induction, makes RCon relatively easy to test in experiments.

Our final axiom looks at the situation of indifference, and is required only if we focus on choice correspondences, and not on choice *functions*. Consider the case in which $\{x, y\} \subseteq \mathbf{c}(S)$ for some $S \in \mathfrak{X}$. This could happen in two ways. The first is when the agent is actually indifferent between x and y, that is, when $\{x, y\} = \mathbf{c}\{x, y\}$. The second is when the agent prefers one of the two elements to the other, say x to y, but there is some third element in S which leads the agent to choose y *as well*. That is, in this latter case the reference effect acts by increasing the appeal of y "just enough" to make it indifferent to x. The next postulate rules out this case, thereby imposing that reference effects occur in a procedural manner – either they come into play, leading the agent to choose one option, or they do not, and the agent chooses the same option she would choose from doubletons.

Rationality of Indifference (RI). For every $x, y \in X$, if $\{x, y\} \subseteq \mathbf{c}(S)$ for some $S \in \mathfrak{X}$, then $\{x, y\} = \mathbf{c}\{x, y\}$.

We conclude this section by noting that every one of the axioms we have introduced above corresponds to a weakening of WARP. Indeed, while No-Cycle, RCon and RI are trivially weaker than WARP, the RCoh property is weaker than WARP because, when $\mathbf{c}$ satisfies WARP, there does not exist an alternative that can act as a revealed $\mathbf{c}$-reference.

Remark. The four behavioral postulates we have formulated above constitute a logically independent set of axioms. This is proved in the Appendix.

2.4 Two Examples

We have argued above that our axiomatic framework is weaker than that allowed in the standard choice theory that is based on WARP. In particular, this framework lets one model reference-dependent behavior that pertains to the attraction effect, as we illustrate next.

Example 1. (*Choice with attraction effect*) Let $X := \{x, y, z\}$, and consider the choice function $\mathbf{c}$ on $2^X\backslash\{\emptyset\}$ depicted below:

$$\mathbf{c}\{x,y\} = y, \quad \mathbf{c}\{x,z\} = x, \quad \mathbf{c}\{y,z\} = y \quad \text{and} \quad \mathbf{c}\{x,y,z\} = x.$$

Obviously, this choice behavior is in violation of WARP – the alternative x is not revealed preferred to y even though it is deemed to be the choice from $\{x, y, z\}$. (Thus $\{x, y, z\}$ is $\mathbf{c}$-awkward.) This is the typical behavior of the attraction effect phenomenon. The alternative y emerges as the best alternative with respect to pairwise comparisons, but presumably because z acts as a reference that "attracts" the attention of the individual to x when these two alternatives are both feasible, the agent chooses x from the set $\{x, y, z\}$. It is easily checked that this behavior is consistent with all four of the axioms we considered above; we find here that z is a revealed $\mathbf{c}$-reference for x.

At the other end, one may be worried that the discipline of our axiomatic model is perhaps too lax, thereby leaving us with a theory with little predictive power. It is indeed not transparent what sort of behavior is ruled out by the conjunction of the behavioral postulates of Section 2.3. This will be best understood through the representation theorems we present in the next section, but if only to give an idea about this, and illustrate how these axioms work together, we present an example of a choice correspondence that is ruled out by the properties of No-Cycle and RCon.

Example 2. (*Choice that is not rationalized by reference-dependence*) Let $X := \{w, x, y, z\}$, and consider a choice correspondence $\mathbf{c}$ on $2^X\backslash\{\emptyset\}$ such that

$$\mathbf{c}\{x,y\} \ni x, \quad \mathbf{c}\{x,y,z\} = \{z\}, \quad \mathbf{c}\{w,x,y\} = \{y\} \quad \text{and} \quad \mathbf{c}(X) = \{w\}.$$

Then, $\mathbf{c}$ is ruled out by the present axiomatic model. To see this, suppose $\mathbf{c}$ satisfies No-Cycle and RCon. As $\{\{w, y\}, \{x, y\}\}$ is a $\mathbf{c}$-cover of $\{w, x, y\}$ and $\{y\} \neq \mathbf{c}\{x, y\}$, RCon entails that $\{y\} = \mathbf{c}\{w, y\}$. Similarly, as $\{\{w, x, y\}, \{w, z\}\}$ is a $\mathbf{c}$-cover of X and $\{w\} \neq \mathbf{c}\{w, x, y\}$, we have $\{w\} = \mathbf{c}\{w, z\}$. It then follows from No-Cycle that $\{x\} = \mathbf{c}\{x, z\}$. Combining this observation with the fact that $\{\{x, z\}, \{y, z\}\}$ is a $\mathbf{c}$-cover of $\{x, y, z\}$, RCon entails that either $\{z\} = \mathbf{c}\{y, z\}$. It follows that we have $\{z\} = \mathbf{c}\{y, z\}$, $\{w\} = \mathbf{c}\{w, z\}$ and $\{y\} = \mathbf{c}\{w, y\}$, contradicting No-Cycle.

2.5 Reference Dependence versus Menu Dependence

If one's choice behavior is reference-dependent, then it cannot be identified solely on the basis of her pairwise comparisons of the alternatives. As such, reference-dependent choice behavior is a special case of menu-dependent choice behavior. The converse is false, however. There are menu-dependent behaviors that cannot be attributed to reference-dependence. Before we move to the representation of the latter type of behavior, we use the No-Cycle condition to demonstrate this point.

Let us consider the following scenario. An individual is in an (unusual) restaurant which serves only two dishes, steak tartare and chicken. The individual loves steak tartare, but is cautious about ordering it, because if it is not prepared well and fresh, eating steak tartare can be hazardous to health. In addition, there seems to be hardly any reason to expect that a restaurant that serves only steak tartare and grilled chicken would be a good place to eat steak tartare at. Thus, she orders the chicken dish. But now consider the alternate scenario in which at the same (or some other restaurant unknown to her), the menu consists of steak tartare, grilled chicken and frog legs. The individual does not fancy frog legs in particular, but she is well aware that this dish is a delicacy of French cuisine. She thus reasons that steak tartare should be good here, and thus orders the steak tartare.

This is one of the classic examples in which perfectly reasonable choice behavior leads to a violation of WARP.[11] While its justification is through an information-based scenario, the choice behavior itself resembles the attraction effect phenomenon closely. After all, where we denote steak tartare by x, grilled chicken by y and frog legs by z, the choice function $\mathbf{c}$ of the individual maintains that $\mathbf{c}\{x,y\} = y$ and $\mathbf{c}\{x,y,z\} = x$. The upshot here is that whether or not this behavior conforms with the attraction effect depends on the other pairwise comparisons. Indeed, the story suggests that we have $\mathbf{c}\{x,z\} = x$; after all, the information content of the set $\{x,z\}$ with regard to the cook's proficiency in French cooking is exactly the same as the set $\{x,y,z\}$. But then, if $\mathbf{c}$ were to satisfy No-Cycle, we would have to have $\mathbf{c}\{y,z\} = y$, which is rather awkward as there is nothing in the story that tells us that this individual likes grilled chicken better than frog legs. Put differently, where $X := \{x,y,z\}$, the choice function $\mathbf{c}$ on $2^X\backslash\{\emptyset\}$ with

$$\mathbf{c}\{x,y\} = y, \quad \mathbf{c}\{x,z\} = x, \quad \mathbf{c}\{y,z\} = z \quad \text{and} \quad \mathbf{c}\{x,y,z\} = x$$

fits perfectly with the story above, and yet it fails to satisfy No-Cycle. This shows that

[11] This example is due to Luce and Raiffa (1957) who "explain" the involved behavior by saying that one may maximize a preference relation if a certain alternative (in this example, frog legs) is on the menu, and a different preference relation otherwise. This is a prototypical example of menu-dependent rational behavior.

WARP is violated here due to a form of "menu dependence" that goes beyond reference-dependence. (Notice that the choice y in the menu $\{x, y\}$ is justified not by using x as a reference that attracts the agent to y, but rather using the content of the entire menu $\{x, y\}$.)

3 Representation Theorems

3.1 The Reference-Dependent Choice Model

Our objective now is to provide a representation for choice correspondences that satisfy the four behavioral properties we have considered above. First, let us fix a symbol $\diamondsuit$. (Formally speaking, $\diamondsuit$ can be thought of as any object that does not belong to X.) The following auxiliary definition is useful for streamlining the presentation.

Definition. A function $r \colon \mathfrak{X} \to X \cup \{\diamondsuit\}$ is said to be a **reference map** on $\mathfrak{X}$ if, for any $S \in \mathfrak{X}$, we have $r(S) \in S$ whenever $r(S) \neq \diamondsuit$, and $r(S) = \diamondsuit$ whenever $|S| \leq 2$.

The following is our first main result.

Theorem 1. *A continuous choice correspondence* $\mathbf{c}$ *satisfies No-Cycle, RCoh, RCon and RI if, and only if, there exist a continuous function* $u : X \to \mathbb{R}$, *a correspondence* $Q : X \cup \{\diamondsuit\} \rightrightarrows X$ *with* $Q(\diamondsuit) = X$, *and a reference map* $\mathbf{r}$ *on* $\mathfrak{X}$ *such that* $\mathbf{r}(S) \neq \diamondsuit$ *iff* S *is* $\mathbf{c}$-awkward,

$$\mathbf{c}(S) = \arg\max u(S \cap Q(\mathbf{r}(S))) \tag{1}$$

for every $S \in \mathfrak{X}$, *and*

$$\arg\max u(T \cap Q(\mathbf{r}(T))) = \arg\max u(T \cap Q(\mathbf{r}(S))) \tag{2}$$

for every $S, T \in \mathfrak{X}$ *with* $\mathbf{r}(S) \in T \subseteq S$ *and* $\arg\max u(S \cap Q(\mathbf{r}(S))) \cap T \neq \emptyset$.[12]

In what follows, we refer to any triplet $\langle u, \mathbf{r}, Q \rangle$ in which u is a continuous real map on X, $\mathbf{r}$ is a reference map on $\mathfrak{X}$, and $Q : X \cup \{\diamondsuit\} \rightrightarrows X$ is a correspondence such that $Q(\diamondsuit) = X$ and (2) holds (for every $S, T \in \mathfrak{X}$ with $\mathbf{r}(S) \in T \subseteq S$ and $\arg\max u(S \cap Q(\mathbf{r}(S))) \cap T \neq \emptyset$), as a **reference-dependent choice model** on $\mathfrak{X}$. In the context of such a model, we call u the *utility function*, $\mathbf{r}$ the *reference map*, and Q the *attraction correspondence* of the model. Finally, if $\mathbf{c}$ is a choice correspondence on $\mathfrak{X}$ such that (1) holds and $\mathbf{r}(S) = \diamondsuit$ iff S is $\mathbf{c}$-awkward, we say that $\langle u, \mathbf{r}, Q \rangle$ **represents** $\mathbf{c}$.

[12] *Notation.* For any nonempty sets S and X such that $S \subseteq X$, and real-valued function f on X, by $\arg\max f(S)$, we mean the set of all maximizers of f on S, that is, $\{x \in S : f(x) \geq f(\omega) \text{ for all } \omega \in S\}$.

In this jargon, Theorem 1 says that a continuous choice correspondence $\mathbf{c}$ on $\mathfrak{X}$ satisfies the four behavioral properties of Section 2.3 iff it can be represented by a reference-dependent choice model $\langle u, \mathbf{r}, Q \rangle$ on $\mathfrak{X}$. Here, u is interpreted as a standard utility function for the decision-maker, *free of any referential considerations.* In turn, the reference map $\mathbf{r}$ tells us whether the agent uses a reference point in a given choice problem S or not. In the former case, $\mathbf{r}(S)$ stands for the reference alternative in S. By definition of a reference map, such a reference alternative must be feasible in S, that is, $\mathbf{r}(S) \in S$. The case where the agent does not discern a reference item in S, is captured by the statement $\mathbf{r}(S) = \diamondsuit$. Again by definition of a reference map, this situation is sure to occur when S contains at most two alternatives, for then, as we have noted earlier, there is no room for an alternative to change the comparative evaluation of two other choice items. More generally, $\mathbf{r}(S) = \diamondsuit$, that is, referential considerations do no affect one's choice behavior over S (and its subsets), iff S is not $\mathbf{c}$-awkward.

For any $\omega \in X \cup \{\diamondsuit\}$, we interpret the set $Q(\omega)$ as the *attraction region* of ω; this corresponds to the set of all alternatives in X that "look better" to the decision-maker when compared to ω. (For instance, if the agent deems a number of attributes of the alternatives as relevant for her choice, then $Q(\omega)$ may be thought of as the set of all alternatives that dominate ω with respect to all attributes.) In line with this interpretation, the condition $Q(\diamondsuit) = X$ posits that "nothing" (i.e., $\diamondsuit$) does not enhance the appeal of any alternative – it does not attract the agent's attention to any particular set of alternatives.

Given these considerations, we interpret the behavior of a decision-maker in terms of a given reference-dependent choice model $\langle u, \mathbf{r}, Q \rangle$ on $\mathfrak{X}$ as follows. For any choice problem $S \in \mathfrak{X}$, the agent either evaluates what to choose in a reference-independent manner, or identifies a reference point in S and uses this point to finalize her choice. In the former case we have $\mathbf{r}(S) = \diamondsuit$ and $Q(\mathbf{r}(S)) = X$, and hence (1) takes the form $\mathbf{c}(S) = \arg\max u(S)$, in concert with the standard theory of rational choice.[13] In the latter case, $\mathbf{r}(S)$ is an alternative in S, say z, and the agent is "mentally attracted" to the elements of S that belong also to $Q(z)$. It is "as if" she faces the *mental constraint* that her choices from S must belong to $Q(z)$. (See Figure 3.) (1) says that, within this constraint, then, the agent is fully rational: she solves her problem upon the maximization of u, that is, $\mathbf{c}(S) = \arg\max u(S \cap Q(z))$.

Finally, (2) imposes some consistency between the references and choices of an individual from nested sets. To wit, take a problem $S \in \mathfrak{X}$ in which the agent uses a reference point, that is, $\mathbf{r}(S) \neq \diamondsuit$. Now consider another choice problem $T \subseteq S$ which contains some of the choices from S (that is, $\mathbf{c}(S) \cap T \neq \emptyset$) as well as the reference used in S (that is, $\mathbf{r}(S) \in T$).

[13] The standard theory is therefore a special case of our model; it is captured by $\langle u, \mathbf{r}, Q \rangle$ upon setting $r(S) = \diamond$ for all $S \in \mathfrak{X}$.

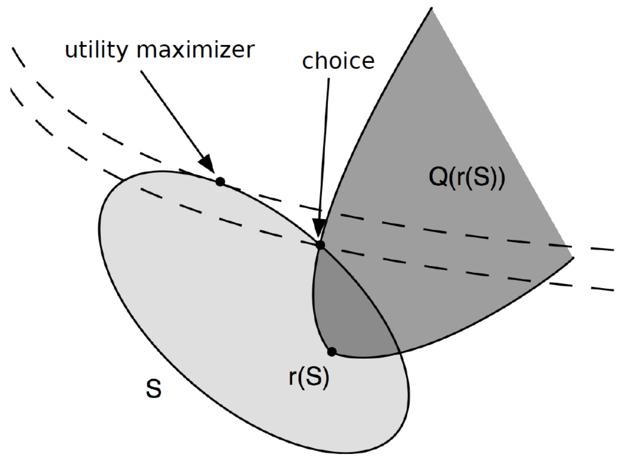

Figure 3

There does not appear to be a compelling reason for the agent to change her referential assessment in this instance; it seems requiring $\mathbf{r}(T) = \mathbf{r}(S)$ is in the nature of things. Yet, while one may choose to impose such a restriction in the case of a specific application, this is really not a behavioral assumption – it cannot possibly be deduced from postulates on a given choice correspondence on $\mathfrak{X}$. After all, $\mathbf{c}$ cannot identify $\mathbf{r}$ uniquely, as we shall see in Section 3.3.. What one can guarantee, however, is the consistency of choices from S and T in this case, and condition (2) does exactly this. That is, when T is nested in S, and both a choice from S and the reference of S remain feasible in T, then, even if a different reference point could be used in T, the agent's choices would be identical to the ones she would have picked if she used $\mathbf{r}(S)$ as the reference point in T. Indeed, if a reference-dependent choice model $\langle u, \mathbf{r}, Q\rangle$ represents $\mathbf{c}$, then (2) ensures that

$$\mathbf{r}(S) \in T \subseteq S \text{ and } \mathbf{c}(S) \cap T \neq \emptyset \quad \text{imply} \quad \mathbf{c}(T) = \mathbf{c}(S) \cap T \tag{3}$$

for any $S, T \in \mathfrak{X}$.[14]

3.2 Properties of Reference-Dependent Choice Models

Using the terminology of Section 2.2, we now identify various properties of choice correspondences that are represented by reference-dependent choice models. First, let us characterize the choice problems that are rendered "awkward" by such correspondences.

[14] *Proof.* By (2), we have $\mathbf{c}(T) = \arg\max u(T \cap Q(\mathbf{r}(S)))$. So, if $x \in \mathbf{c}(T)$, then we have $x \in Q(\mathbf{r}(S))$, while by picking any $y \in \mathbf{c}(S) \cap T$, and using the representation, we find $u(x) \geq u(y) \geq u(\omega)$ for all $\omega \in S \cap Q(\mathbf{r}(S))$. Thus, again by the representation, $\mathbf{c}(T) \subseteq \mathbf{c}(S) \cap T$. The converse containment is similarly established.

Proposition 2. *Let* $\mathbf{c}$ *be a choice correspondence on* $\mathfrak{X}$ *that is represented by a reference-dependent choice model* $\langle u, \mathbf{r}, Q\rangle$*.Then, a choice problem* $S \in \mathfrak{X}$ *is* $\mathbf{c}$*-awkward if, and only if,* $\mathbf{c}(S) \neq \arg\max u(S)$.

Proof. If S is not $\mathbf{c}$-awkward, then $Q(\mathbf{r}(S)) = Q(\diamondsuit) = X$, and we obtain $\mathbf{c}(S) = \arg\max u(S)$ from (1). Conversely, suppose S is $\mathbf{c}$-awkward but we have $\mathbf{c}(S) = \arg\max u(S)$. By definition of $\mathbf{c}$-awkwardness, there exists a doubleton subset T of S with $\emptyset \neq \mathbf{c}(S) \cap T \neq \mathbf{c}(T)$. But as $\mathbf{r}(T) = \diamondsuit$ (because $|T| = 2$), this means that $\emptyset \neq (\arg\max u(S)) \cap T \neq \arg\max u(T)$, which is impossible.

Thus, the choice behavior induced by a reference-dependent choice model $\langle u, \mathbf{r}, Q\rangle$ deems a given choice problem "awkward" iff the choices from that problem are not made solely on the basis of the maximization of the utility function u. In the context of such a model, "awkwardness" is the same thing as departing from "choice by utility maximization."

Given a $\mathbf{c}$-awkward problem S, the representation by $\langle u, \mathbf{r}, Q\rangle$ entails that $\mathbf{r}(S) \neq \diamondsuit$, that is, in this case, the model designates a reference alternative $\mathbf{r}(S)$ in S. We now show that this alternative cannot itself be chosen in S, and it is indeed viewed as a "reference" by $\mathbf{c}$ in the sense that $\mathbf{r}(S)$ is a revealed $\mathbf{c}$-reference for every choice from S.

Proposition 3. *Let* $\mathbf{c}$ *be a choice correspondence on* $\mathfrak{X}$ *that is represented by a reference-dependent choice model* $\langle u, \mathbf{r}, Q\rangle$*. Then, for any choice problem* $S \in \mathfrak{X}$ *with* $\mathbf{r}(S) \neq \diamond$:

(a) $\mathbf{r}(S) \in S \backslash \mathbf{c}(S)$;

(b) $\mathbf{r}(S)$ *is a revealed* $\mathbf{c}$*-reference for each* $x \in \mathbf{c}(S)$.

Proof. Suppose $\mathbf{r}(S) \in \mathbf{c}(S)$. Then, by (3), we have $\mathbf{c}\{x, \mathbf{r}(S)\} = \{x, \mathbf{r}(S)\} \cap \mathbf{c}(S)$ for every $x \in S$, which contradicts the $\mathbf{c}$-awkwardness of S. Thus: $\mathbf{r}(S) \in S \backslash \mathbf{c}(S)$ whenever $\mathbf{r}(S) \in S$, and (a) is proved.

Now, put $z := \mathbf{r}(S)$. By Proposition 2, $\mathbf{c}(S) \neq \arg\max u(S)$. By (1), this implies that there exists $y \in \arg\max u(S) \setminus \mathbf{c}(S)$. Now pick any $x \in \mathbf{c}(S)$ and use (3) to find that $y \in \mathbf{c}\{x, y\}$ and $\{x, y\} \cap \mathbf{c}\{x, y, z\} = \{x\}$. That is, z is a revealed $\mathbf{c}$-reference for x.

The next observation shows that an individual whose choice behavior is represented by a reference-dependent choice model would always choose an alternative over any revealed reference for that alternative.

Proposition 4. *Let* $\mathbf{c}$ *be a choice correspondence on* $\mathfrak{X}$ *that is represented by a reference-dependent choice model* $\langle u, \mathbf{r}, Q\rangle$*. Then,* $u(z) < u(x)$ *for any* $x, z \in X$ *such that* z *is a revealed* $\mathbf{c}$*-reference for* x*. In particular,* $u(\mathbf{r}(S)) < u(x)$ *for any* $S \in \mathfrak{X}$ *with* $\mathbf{r}(S) \in S$ *and* $x \in \mathbf{c}(S)$.

Proof. Take any x and z in X such that z is a revealed $\mathbf{c}$-reference for x. Then, there is a $y \in X$ such that either (i) $x \in \mathbf{c}\{x,y,z\} \backslash \mathbf{c}\{x,y\}$ or (ii) $y \in \mathbf{c}\{x,y\}$ and $\{x,y\} \cap \mathbf{c}\{x,y,z\} = \{x\}$. In either of these cases $\{x,y,z\}$ is $\mathbf{c}$-awkward, so $\mathbf{r}\{x,y,z\} \neq \diamondsuit$ by Proposition 2. On the other hand, if either x or y is $\mathbf{r}\{x,y,z\}$, then (3) readily yields a contradiction. Thus: $\mathbf{r}\{x,y,z\} = z$. It then follows from Proposition 3.(a) that $z \notin \mathbf{c}\{x,y,z\}$, and hence, whether we are in case (i) or (ii), $\{x\} = \{x,z\} \cap \mathbf{c}\{x,y,z\} = \mathbf{c}\{x,z\}$. Then $u(z) < u(x)$ by the representation of $\mathbf{c}$, establishing the first assertion of the present proposition. The second assertion obtains by combining what we have just showed with Proposition 3.(b).

Finally, we highlight the connection between reference alternatives and attraction correspondence for choice correspondences that are represented by $\langle u, \mathbf{r}, Q \rangle$ as in Theorem 1.

Proposition 5. *Let* $\mathbf{c}$ *be a choice correspondence on* $\mathfrak{X}$ *that is represented by a reference-dependent choice model* $\langle u, \mathbf{r}, Q \rangle$*. Then, for every* $x, z \in X$*:*

(a) *If* z *is a revealed* $\mathbf{c}$*-reference for* x*, then* $x \in Q(z)$*;*

(b) *If* $x \in Q(z)$*, then* z *is a potential* $\mathbf{c}$*-reference for* x*.*

Proof. Let z be a revealed $\mathbf{c}$-reference for x. Then, as in the proof of Proposition 4, there is a $y \in X$ such that $\mathbf{r}\{x,y,z\} = z$ and $x \in \mathbf{c}\{x,y,z\}$. The representation now ensures that $x \in Q(z)$. This proves part (a) of the proposition. The proof of part (b) is similar.

In words, Proposition 5 says that every alternative attracts the decision-maker to those items that it is revealed to be a reference for, and conversely, if an alternative happens to attract the agent to x, then that alternative must at least be a potential reference for x.

3.3 Uniqueness of Reference-Dependent Choice Models

Suppose two reference-dependent choice models $\langle u, \mathbf{r}_1, Q_1 \rangle$ and $\langle v, \mathbf{r}_2, Q_2 \rangle$ on $\mathfrak{X}$ are *behaviorally equivalent*, that is, they represent the same choice correspondence on $\mathfrak{X}$. What can we say about the relation between the ingredients of these two models? The answer is straightforward in the case of u and v. These functions are continuous, and they represent the same preference relation on X, and hence, they must be continuous and strictly increasing transformations of each other. The situation is more subtle for the reference maps and attraction correspondences, however. We begin with the former.

Example 3. (*Non-Uniqueness of reference maps*) Let $X := \{x, y, z_1, z_2\}$, and consider a choice correspondence $\mathbf{c}$ on $2^X \backslash \{\emptyset\}$ such that $\mathbf{c}\{x,y\} = \{y\}$ and $\mathbf{c}(S) = \{x\}$ for every non-singleton subset S of X that contains x and is distinct from $\{x,y\}$. A behavior of this kind is typical of the attraction effect phenomenon – compare Figures 1 and 4 – and naturally, it

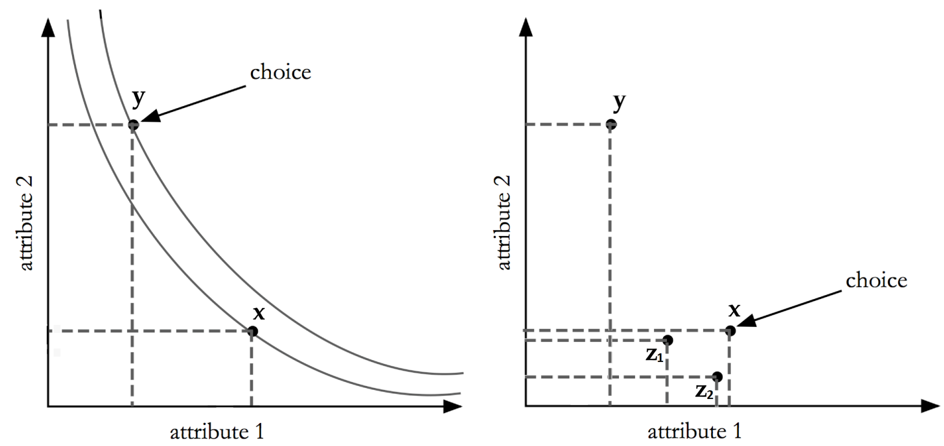

Figure 4

can be captured by a reference-dependent choice model. Suppose $\langle u, \mathbf{r}, Q \rangle$ is one such model. Then, clearly, $u(y) > u(x) > \max\{u(z_1), u(z_2)\}$ and $\mathbf{r}(S) \in \{z_1, z_2\}$ for every subset S of X that contains x, y and at least one of the alternatives z_1 and z_2. The reference map $\mathbf{r}$ is, therefore, uniquely identified at any nonempty proper subset of X. However, insofar as the choice behavior is concerned, the model cannot possibly distinguish between the referential attributes of z_1 and z_2; putting either $\mathbf{r}(X) = z_1$ or $\mathbf{r}(X) = z_2$ makes no behavioral difference. Put differently, if $\mathbf{r}(X) = z_1$, and $\mathbf{r}'$ is a reference map on $\mathfrak{X}$ that agrees with $\mathbf{r}$ at every nonempty proper subset of X but has $\mathbf{r}'(X) = z_2$, then $\langle u, \mathbf{r}', Q \rangle$ represents $\mathbf{c}$ as well.

This example suggests that there is a sense of "arbitrariness," in the specification of a reference map in general. Consequently, the uniqueness properties of a reference-dependent model in terms of its reference map would best be identified by focusing on *all* options that can act as references in a given choice problem simultaneously. To this end, we introduce the notion of a reference *correspondence.*

Definition. Let $\langle u, \mathbf{r}, Q \rangle$ be a reference-dependent choice model that represents a choice correspondence $\mathbf{c}$ on $\mathfrak{X}$. We define the **reference correspondence** $\mathbf{R} : \mathfrak{X} \to 2^X \cup \{\diamondsuit\}$ associated with this model by setting $\mathbf{R}(S) := \{\diamondsuit\}$ if $\mathbf{r}(S) = \diamondsuit$, and taking $\mathbf{R}(S)$ as the set of all z in S such that

$$\mathbf{c}(T) = \arg\max u(T \cap Q(z)) \quad \text{for every } T \in \mathfrak{X} \text{ with } z \in T \subseteq S \text{ and } \mathbf{c}(S) \cap T \neq \emptyset$$

if $\mathbf{r}(S) \neq \diamondsuit$.[15]

[15] In view of Theorem 1, $\mathbf{c}$ plays only an auxiliary role here. That is, $\mathbf{R}$ is determined uniquely once $\langle u, \mathbf{r}, Q \rangle$

In words, given a reference-dependent choice model $\langle u, \mathbf{r}, Q\rangle$ that represents a choice correspondence $\mathbf{c}$ on $\mathfrak{X}$, the correspondence $\mathbf{R}$ maps any $\mathbf{c}$-awkward feasible set S to the collection of *all* alternatives that could be used as a reference point in the set S; $\mathbf{r}(S)$ is only one of these alternatives. On any non-$\mathbf{c}$-awkward feasible set the choice behavior is reference-free, so $\mathbf{R}$ maps any such set to $\diamondsuit$. (For instance, if $\mathbf{c}$ is the choice correspondence of Example 3 here, we have $\mathbf{R}(X) = \{z_1, z_2\}$.)

In the case of the attraction correspondences the uniqueness issue is not so much about "arbitrariness," but about "redundancy." Put more precisely, it is about the potentially redundant alternatives that may be contained in Q. To illustrate, suppose $\langle u, \mathbf{r}, Q\rangle$ is a reference-dependent choice model and z is an alternative in X with $\{z\} = \arg\min u(X)$. Now take any $Q' : X \cup \{\diamondsuit\} \rightrightarrows X$ with $Q(x)\backslash\{z\} \subseteq Q'(x) \subseteq Q(x) \cup \{z\}$ for each $x \in X \cup \{\diamondsuit\}$. It is plain that $\langle u, \mathbf{r}, Q'\rangle$ is then a reference-dependent choice model that represents exactly the same choice correspondence that $\langle u, \mathbf{r}, Q\rangle$ represents. As z is never chosen in any choice problem (other than in $\{z\}$), it can never influence the choice behavior of an agent whether or not it passes the attraction test imposed by Q.

The uniqueness properties of attraction correspondences may be identified by eliminating all redundant elements such a correspondence (in a given reference-dependent choice model) may contain. As we shall see, an additional merit of doing this is to obtain an "optimal" representative from any given collection of all behaviorally equivalent reference-dependent choice models. The following concept will be instrumental for this purpose.

Definition. Let $\langle u, \mathbf{r}, Q\rangle$ be a reference-dependent choice model on $\mathfrak{X}$. We define the correspondence $Q^* : X \cup \{\diamondsuit\} \rightrightarrows X$ by $Q^*(\diamondsuit) := X$, and

$$Q^*(z) := \{x \in X : \mathbf{r}(S) = z \text{ and } x \in \arg\max u(S \cap Q(z)) \text{ for some } S \in \mathfrak{X}\}$$

for each $z \in X$.

Given a reference-dependent choice model $\langle u, \mathbf{r}, Q\rangle$ that represents $\mathbf{c}$ on $\mathfrak{X}$, the correspondence Q^* maps any alternative z to the set of all x that would be choosable (according to $\mathbf{c}$) in a feasible set for which z acts as a reference. Thus, intuitively, Q^* is obtained from Q by pruning the latter so as to get rid of the redundant elements $Q(z)$ might possess (for each z in X). What is more, this correspondence has behavioral content:

$$Q^*(z) = \{x \in X : z \text{ is a revealed } \mathbf{c}\text{-reference for } x\} \quad \text{for all } z \in X. \tag{4}$$

is specified; we use the choice correspondence represented by $\langle u, \mathbf{r}, Q\rangle$ in the definition only to simplify the involved notation.

This is an easy consequence of Proposition 5.(a) and the definitions of $Q^*(z)$ and being a revealed $\mathbf{c}$-reference.

We are now ready to characterize the uniqueness properties of reference-dependent choice models.

Proposition 6. *Let $\mathbf{c}$ be a choice correspondence on $\mathfrak{X}$ that is represented by a reference-dependent choice model $\langle u_1, \mathbf{r}_1, Q_1 \rangle$. Then, a reference-dependent choice model $\langle u_2, \mathbf{r}_2, Q_2 \rangle$ represents $\mathbf{c}$ if, and only if, $u_2 = f \circ u_1$ for some continuous and strictly increasing $f : u_1(X) \to \mathbb{R}$, $\mathbf{R}_1 = \mathbf{R}_2$, and $Q_1^* = Q_2^*$.*

The utility functions of two behaviorally equivalent reference-dependent choice models must be ordinally equivalent. Moreover, they must induce the same reference correspondence, which means that the difference between the involved reference maps must be due only to the "arbitrariness" matter we have discussed above. Finally, the difference between the attraction correspondences of these two models must be due only to the potential redundancies; removing these redundancies yields the same attraction correspondence.

We may use Proposition 6 to identify a representation with the finest possible attraction correspondence. Given two behaviorally equivalent reference-dependent choice models $\langle u_1, \mathbf{r}_1, Q_1 \rangle$ and $\langle u_2, \mathbf{r}_2, Q_2 \rangle$ on $\mathfrak{X}$, let us say that the former is **tighter than** the latter if $Q_1(z) \subseteq Q_2(z)$ for every $z \in X$. Clearly, the tighter a reference-dependent choice model, the fewer are the redundancies allowed in the involved attraction correspondence. The most efficient representative of a given collection of all behaviorally equivalent reference-depencent choice models is thus the tightest among them.

Clearly, if $\langle u_1, \mathbf{r}_1, Q_1 \rangle$ is a reference-dependent choice model that represents a given choice correspondence $\mathbf{c}$, then $\langle u_1, \mathbf{r}_1, Q_1^* \rangle$ is another reference-dependent choice model that represents $\mathbf{c}$. Moreover, by (4), Proposition 5.(a) and Proposition 6, if $\langle u_2, \mathbf{r}_2, Q_2 \rangle$ is another representation of $\mathbf{c}$, we necessarily have $Q_1^*(z) \subseteq Q_2(z)$ for every $z \in X$ and $Q_2^* = Q_1^*$. Thus:

Corollary 7. *Let $\mathbf{c}$ be a choice correspondence on $\mathfrak{X}$ that is represented by a reference-dependent choice model $\langle u_1, \mathbf{r}_1, Q_1 \rangle$. Then, $\langle u_1, \mathbf{r}_1, Q_1^* \rangle$ is a tightest reference-dependent choice model that represents $\mathbf{c}$.*

3.4 Reference Dependence and Asymmetric Dominance

A reference-dependent choice model $\langle u, \mathbf{r}, Q \rangle$ found in Theorem 1 posits hardly any structure on the attraction correspondence Q. As a final order of business in this section, we discuss how one may sharpen that characterization by determining when we can guarantee that the attraction region of each alternative (according to Q) would consist of the options that dominate it in terms of a *set* of utilities. This is very much in line with the empirical evidence on the attraction effect phenomenon.

The key to this is to strengthen the RCoh property on a given choice correspondence $\mathbf{c}$ on $\mathfrak{X}$. In particular, consider finitely many alternatives $x_1, \ldots, x_n$ (with $n \geq 2$) such that x_1 is a revealed $\mathbf{c}$-reference for x_2, x_2 is a revealed $\mathbf{c}$-reference for x_3, ..., and x_{n-1} is a revealed $\mathbf{c}$-reference for x_n. In this instance it would be too much to require that x_1 be a revealed reference for x_n. After all, it may be the case that the choice of x_n may never require the "help" of x_1 for it is already a rather desirable alternative. It stands to reason, however, that x_1 would never "help" an alternative "against" x_n; this would point to an inconsistency in the way references arise for the decision-maker. Consequently, it makes sense to presume that x_1 must at least be a potential $\mathbf{c}$-reference for x_n.

Reference Acyclicity (RA). For any integer $n \geq 2$ and $x_1, ..., x_n \in X$, if x_i is revealed $\mathbf{c}$-reference for x_{i+1} for each $i = 1, ..., n-1$, then x_1 is a potential $\mathbf{c}$-reference for x_n.

It is plain that RA is stronger than RCoh. Our next result demonstrates how Theorem 1 alters if we replace the latter property with RA.

Theorem 8. *A continuous choice correspondence* $\mathbf{c}$ *satisfies No-Cycle, RCon, RA and RI if, and only if, there is a reference-dependent choice model* $\langle u, \mathbf{r}, Q \rangle$ *on* $\mathfrak{X}$ *and a nonempty set* $\mathcal{U}$ *of real maps on* X *such that* $\langle u, \mathbf{r}, Q \rangle$ *represents* $\mathbf{c}$ *and*

$$Q(z) = \{y \in X : U(y) \geq U(z) \text{ for each } U \in \mathcal{U}\}, \quad z \in X. \tag{5}$$

Thus, if a choice correspondence $\mathbf{c}$ satisfies the axioms of Theorem 8, then, not only the associated choice behavior can be represented by means of a reference-dependent choice model $\langle u, \mathbf{r}, Q \rangle$, but we have a specific interpretation for the attraction correspondence Q in this case. It is as if the agent were endowed with a collection of utility functions, and in a problem where she discerns a reference point, she were attracted to those options that happen to dominate that reference in terms of each of these utilities. In turn, one may interpret these utilities as measuring the performance of alternatives in X with respect to a certain attribute. According to this interpretation, then, the individual behaves exactly as prescribed by the attraction effect phenomenon: the presence of asymmetrically dominated alternatives (relative to a reference item) attracts her to choose among those options that dominate the reference item. We emphasize, however, that Theorem 8 is more general than the usual presentation of this phenomenon. This theorem obtains *endogenously*, and on the basis of purely behavioral postulates, not only the reference points of a decision-maker, but also her *subjective* attributes in terms of which she evaluates the alternatives. (By contrast, the attributes in the classical case of the attraction effect phenomenon are given as part of the description of the problem.) Theorem 8, therefore, identifies how one can check for the presence of attraction effect in contexts where attributes of the choice prospects are not explicitly given, and provides a choice model that accounts for this effect across any choice domain.

Remark. If "being potential $\mathbf{c}$-reference for" is a transitive relation on X which is closed in $X \times X$, then we may take $\mathcal{U}$ in Theorem 8 as a countable collection of continuous real maps on X. Moreover, when X is finite, $\mathcal{U}$ can be taken as a finite collection in this theorem.

4 Welfare Analysis

It is well known that making welfare comparisons is not a trivial matter when subjects violate WARP, for then the revealed preference criterion may fail to give us a consistent ranking across choice problems. In this section we discuss a few criteria that could be used to make such comparisons when the violations of WARP arise due to referential considerations.

An immediate possibility is to regard an individual with a choice correspondence $\mathbf{c}$ "better off" with an alternative x as opposed to y – we denote this situation by $x \succsim^{\mathbf{c}} y$ – if the individual would choose x over y in a pairwise choice problem, that is, $x \in \mathbf{c}\{x, y\}$. This criterion would not be suitable in the case of arbitrary violations of WARP, as it may well yield a cylic method of making welfare comparisons. This difficulty, however, does not arise in the present context where violations of rationality occur only due to reference effects. Indeed, the No-Cycle assumption makes the above welfare ranking a well-defined, complete and transitive binary relation on X. In fact, if $\mathbf{c}$ is a choice correspondence that is represented by a reference-dependent choice model $\langle u, \mathbf{r}, Q \rangle$ on $\mathfrak{X}$, then the pairwise welfare criterion is captured by the utility function u, that is, $x \succsim^{\mathbf{c}} y$ iff $u(x) \geq u(y)$, for every $x, y \in X$. In addition, it also follows readily from the representation that this criterion is used by the agent not only in pairwise choice situations, but also when she evaluates options that her reference point renders attractive.

The pairwise welfare ranking criterion ignores the potential effect of referential considerations on one's welfare judgements, and there does not seem to be an *a priori* reason for this. A more comprehensive approach would regard an individual with choice correspondence $\mathbf{c}$ "better off" with an alternative x as opposed to y – we denote this situation by $x \rhd^{\mathbf{c}} y$ – if the individual would never choose y in a choice problem that contains x, that is, $y \in X \backslash \mathbf{c}(S)$ for every $S \in \mathfrak{X}$ with $x \in S$. This is precisely the criterion suggested by Bernheim and Rangel

(2009) to make welfare comparisons when one's choice behavior violates WARP.

In terms of choice correspondences that admit a representation such as the one given in Theorem 1, the Bernheim-Rangel welfare criterion can be characterized as follows:

Proposition 9. *Let* $\mathbf{c}$ *be a choice correspondence on* $\mathfrak{X}$ *that is represented by a reference-dependent choice model* $\langle u, \mathbf{r}, Q\rangle$*. Then, for any* x *and* y *in* X*, the following are equivalent:*

(a) $x \rhd^{\mathbf{c}} y$;

(b) $y \notin \mathbf{c}\{x, y, z\}$ *for any* $z \in X$;

(c) $u(x) > u(y)$*, and* $\mathbf{r}\{x, y, z\} \neq z$ *for any* $z \in X$.

Proof. It is immediate that (a) implies (b). Suppose (b) holds. Then, setting $z := y$ in (b) yields $\{x\} = \mathbf{c}\{x, y\}$, that is, $u(x) > u(y)$. So, if (c) were false, there would exist a $z \in X$ such that $\mathbf{r}\{x, y, z\} = z$. By Proposition 3, $\mathbf{c}\{x, y, z\} \subseteq \{x, y\}$. Then, Proposition 4 and the fact that $u(x) > u(y)$ imply that $u(x) > u(z)$. Since $\mathbf{r}\{x, y, z\} = z \neq \diamondsuit$, the definition of a reference-dependent choice model now gives us that $\mathbf{c}\{x, y, z\} \neq \arg\max u\{x, y, z\} = \{x\}$. Consequently, $\mathbf{c}\{x, y, z\} = \{y\}$, which contradicts (b). Conclusion: (b) implies (c). Finally, suppose $(x, y) \in S \times \mathbf{c}(S)$ for some $S \in \mathcal{X}$. By the represention of $\mathbf{c}$, Proposition 3.(a) and Proposition 4, this can only happen if either (i) $u(y) \geq u(x)$ or (ii) $u(x) > u(y)$ and $\mathbf{r}(S) \notin \{\diamond, x, y\}$. In case (i), we obviously contradict (c). In case (ii), setting $z := \mathbf{r}(S)$ and using (3), we find $y \in \mathbf{c}\{x, y, z\}$ and the same reasoning used above shows that this can happen only if $\mathbf{r}\{x, y, z\} = z$, which contradicts (c). Conclusion: (c) implies (a).

For choice correspondences with a representation such as the one given in Theorem 8, we may formulate sharper sufficient conditions for the Bernheim-Rangel welfare criterion to apply.

Proposition 10. *Let* $\mathbf{c}$ *be a choice correspondence on* $\mathfrak{X}$ *that is represented by a reference-dependent choice model* $\langle u, \mathbf{r}, Q\rangle$ *such that* (5) *holds for some nonempty* $\mathcal{U} \subseteq \mathbb{R}^X$*. Then:*

(a) $x \rhd^{\mathbf{c}} y$ *for any* $x, y \in X$ *such that* $u(x) > u(y)$ *and* $U(x) \geq U(y)$ *for all* $U \in \mathcal{U}$;

(b) $x \rhd^{\mathbf{c}} z$ *for any* $x, z \in X$ *such that* z *is a revealed* $\mathbf{c}$*-reference for* x.

Proof. Part (a) follows readily from the representation and (5). On the other hand, if z is a revealed $\mathbf{c}$-reference for x, then by Proposition 4, we have $u(x) > u(z)$ and $U(x) \geq U(z)$ for all $U \in \mathcal{U}$, so we may invoke part (a) to obtain (b).

5 Further Considerations

5.1 On Potential Applications

The prevalence of the attraction effect phenomenon in a variety of contexts suggests that reference-dependent choice models may find applications in a diverse set of situations. Especially in contexts where it makes sense to view the choice items as possessing multiple attributes relevant for choice (may they be objective or subjective), the two models developed here may be appropriate for representing the individual choice behavior.

Industrial organization provides many settings in which one can examine the market-related consequences of the attraction effect. Indeed, a companion paper, Ok, Ortoleva and Riella (2011), already provides one such application in the context of monopolistic vertical product differentiation. It considers the standard screening problem of a monopolist who chooses price/quality bundles to offer to consumers whose quality valuation is private information; this is the classical model of Mussa and Rosen (1978). The model of Ok, Ortoleva and Riella (2011) allows a fraction of the customers to be reference-dependent decision makers whose choices are modeled by an instant of the reference-dependent choice model characterized in Theorem 8.[16] It is found that, under some parametric restrictions, the attraction effect would in equilibrium be exploited by the monopolist to better segment the market. If sufficiently many consumers are subject to the attraction effect, the firm may even reach to the first best solution and extract all the surplus, despite the presence of asymmetric information. (One interesting prediction that emanates from this analysis is that the monopolist would produce decoy products only for high quality goods.) Recently, Spiegler (2012) has also utilized the reference-dependent choice model of Theorem 8 to examine how one may capture the strategic considerations that guide firms when they attempt to influence the reference points of the consumers who are subject to the attraction effect.

One can also use the reference-dependent choice models of Theorems 1 and 8, in the contexts of investment decisions, such as portfolio choice, to investigate the potential financial implications of the attraction effect phenomenon. Similarly, in multi-dimensional (spatial) voting problems (with commitment) in which there are at least three candidates, it would be interesting to investigate how the equilibrium choice of platforms by the candidates would be affected if it is known that a part of the voters are subject to the attraction effect. This problem too can be formally defined by using a reference-dependent choice model.

[16]This is in the same spirit with the work of Esteban and Miyagawa (2006) and Esteban, Miyagawa and Shum (2007) who also study the Mussa-Rosen model by modeling the consumer preferences in a non-standard manner, namely, by means of using the self-control preferences of Gul and Pesendorfer (2001).

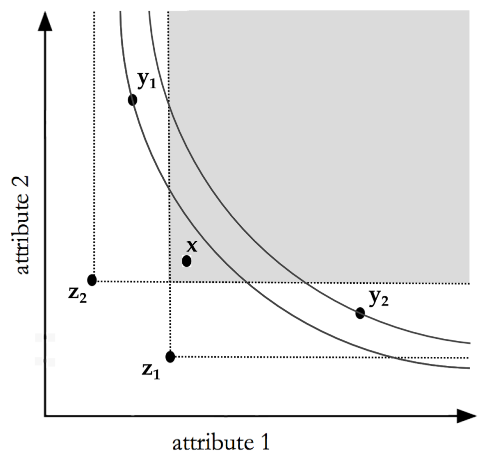

Figure 5

5.2 Caveat: Cumulative Reference-Dependence

The revealed preference theory that we have developed in this paper models the behavior of a decision-maker who may condition her choices from a feasible set on a reference alternative that she may identify in that set. Implicit in this behavior is that one utilizes a single reference option, but not multiple reference alternatives *jointly.* As we elaborate now, this points to a shortcoming of the choice models characterized in Theorems 1 and 8, as well as to a natural direction that the present work should be extended in future research.

Example 4. Let $X := \{x, y, z, w\}$, and consider a choice function $\mathbf{c}$ on $2^X \backslash \{\emptyset\}$ such that $\mathbf{c}\{x, y\} = \mathbf{c}\{y, z, w\} = \mathbf{c}\{x, y, z\} = \mathbf{c}\{x, y, w\} = y$ and $\mathbf{c}\{x, z, w\} = \mathbf{c}(X) = x$. Such a choice function cannot be represented by a reference-dependent choice model – it violates RCon – but there is reason to view it as arising from reference-dependence. The idea is that both z and w are options that "help" x be chosen, but, alone, the reference effect of neither is powerful enough to sway the attention of the decision-maker from y to x in the problems $\{x, y, z\}$ and $\{x, y, w\}$, respectively. However, the *cumulative* referential effect created by observing the feasibility of both z and w together in the problem X is enough to divert the attention of the individual to x.

Example 5. Let $X := \{x, y_1, y_2, z_1, z_2\}$, and consider a choice function $\mathbf{c}$ on $2^X \backslash \{\emptyset\}$ such that $\mathbf{c}\{x, y_1, y_2, z_1\} = y_2$, $\mathbf{c}\{x, y_1, y_2, z_2\} = y_1$ and $\mathbf{c}(X) := x$. Again, such a choice function violates RCon, so it cannot be represented by a reference-dependent choice model. But, intuitively, the violation of WARP by this choice function arises due to reference-dependence. As Figure 4 illustrates, the alternative z_1 may "help" x against y_1, and z_2 may "help" x

against y_2, much in the same spirit with the attraction effect phenomenon. However, z_1 does not eliminate y_2 in the problem $\{x, y_1, y_2, z_1\}$, and z_2 does not eliminate y_1 in $\{x, y_1, y_2, z_2\}$, from consideration, so x is not chosen from either of these choice situations. Yet, in the context of the problem X, it seems reasonable that the agent may feel the reference effects of *both* z_1 *and* z_2, the former diverting attention to x and y_2, and the latter to x and y_1, and hence the two referential evaluations point *jointly* to choosing the option x.

These examples suggest that it may be a good idea to extend reference-dependent choice models to account for cumulative reference effects. We now outline a way of doing this.

Definition. A function $r\colon \mathfrak{X} \to 2^X \cup \{\diamondsuit\}$ is said to be a **multi-reference map** on $\mathfrak{X}$ if, for any $S \in \mathfrak{X}$, we have $\emptyset \neq \mathbf{r}(S) \subseteq S$ whenever $r(S) \neq \diamondsuit$, and $r(S) = \diamondsuit$ whenever $|S| \leq 2$.

We say that a triplet $\langle u, \mathbf{r}, Q \rangle$ in which u is a continuous real map on X, $\mathbf{r}$ is a multi-reference map on $\mathfrak{X}$, and $Q : (2^X \backslash \{\emptyset\}) \cup \{\diamondsuit\} \rightrightarrows X$ is a correspondence such that $Q(\diamondsuit) = X$ and (2) holds (for every $S, T \in \mathfrak{X}$ with $\mathbf{r}(S) \subseteq T \subseteq S$ and $\arg\max u(S \cap Q(\mathbf{r}(S))) \cap T \neq \emptyset$), is a **multi-reference-dependent choice model** on $\mathfrak{X}$. Such a model is said to **represent** a choice correspondence $\mathbf{c}$ on $\mathfrak{X}$ if (1) holds for each $S \in \mathfrak{X}$.

Provided that their definitions are suitably completed, the choice functions considered in Examples 4 and 5 can be represented by multi-reference-dependent choice models. (In the first model we would have $\mathbf{r}(X) = \{z, w\}$ and $Q\{z, w\} = \{x\}$, and in the second model we would set $\mathbf{r}(X) = \{z_1, z_2\}$ and $Q\{z_1, z_2\} = \{x\}$.) However, the notion of a multi-reference-dependent choice model is a bit too general in that it allows sets of reference options to be assigned attraction regions essentially arbitrarily. By contrast, Example 4, and especially Example 5, suggest that the reference effect of a collection of options may be regarded as the reference effects of each of these options put together. This leads us to:

Definition. A multi-reference-dependent choice model $\langle u, \mathbf{r}, Q \rangle$ on $\mathfrak{X}$ is said to be a **cumulative reference-dependent choice model** on $\mathfrak{X}$ if

$$Q(\mathbf{r}(S)) = \bigcap \{Q\{z\} : z \in \mathbf{r}(S)\} \quad \text{for every } S \in \mathfrak{X}.$$

A choice function as in Example 4 (suitably defined) can trivially be represented by a cumulative reference-dependent choice model $\langle u, \mathbf{r}, Q \rangle$ on $2^X \backslash \{\emptyset\}$ where $\mathbf{r}(X) = \{z, w\}$ and $Q\{z\} = Q\{w\} = \{x\}$. More interestingly, this model can be used to represent a choice function that is consistent with the choice data of Example 5 by setting $\mathbf{r}(X) = \{z_1, z_2\}$, $Q\{z_1\} = \{x, y_2\}$ and $Q\{z_2\} = \{x, y_1\}$. Then, the joint effect of z_1 and z_2 is found as $Q\{z_1\} \cap Q\{z_2\} = \{x\}$, as illustrated in Figure 5.

A cumulative reference-dependent choice model captures the attraction effect phenomenon both in the standard and the cumulative sense. Identification of the behavioral basis of such models is thus of interest. This is an open problem at present.

5.3 On the Related Literature

The notion of "reference dependence" has been extensively investigated in economics. Indeed, there is a sizable literature on modeling choice problems for which reference points are *exogenously* given – such choice problems are special cases of what Rubinstein and Salant (2008) call "problems with frames."[17] The main difference between this literature and the present approach toward modeling reference-dependent choice lies in our treatment of "references" as being *endogenous* to choice problems. The type of choice problems we study here do not come with any preassigned frames; the presence/absence of a reference option for a decision-maker is determined by observing choices across various situations, very much in the tradition of the revealed preference theory.

To our knowledge, the only other papers that focus on the notion of endogenous reference formation are Köszegi and Rabin (2006, 2007). These papers develop a model in which the agent's reference point is determined as her rational expectations about the outcome she will receive given her behavior, which in turn must be optimal in terms of a classical gain-loss utility function conditional on this reference point. This approach, while interesting, differs from ours in several aspects. First, in the Köszegi-Rabin approach one's reference point is really beliefs about future returns, and hence it is, at least conceptually speaking, not a choice item. Because this approach models both the choice items and beliefs in terms of lotteries over outcomes, however, such references may happen to belong to a given choice problem (at the time of choice). When they do, they emerge as desirable alternatives; the individual may well choose this object in a choice problem. By contrast, in our model a reference point in a choice problem, if it arises, is an option that is surely *feasible* in that problem, but one that the agent *will never choose* (as in the attraction effect); recall Proposition 3.(a). Second, the model of Köszegi and Rabin (2006) coincides with that of rational choice if there is no underlying uncertainty, which means that this model is not suitable to address the instances of reference-dependent behavior like the attraction effect in enviroments in which no uncertainty is present. By contrast, our model allows for uncertainty as a special case, but also addresses reference-dependence under certainty as well.

Also related to the present work are the recent papers that introduce choice models

[17] Among the choice models with exogneously given references are the behavioral (loss aversion) models of Kahneman and Tversky (1979) and Tversky and Kahneman (1991), and the axiomatic choice models of Chateauneuf and Wakker (1999), Masatlioglu and Ok (2005, 2012), and Diecidue and de Ven (2008).

compatible with the attraction effect phenomenon. For instance, de Clippel and Eliaz (2012) axiomatizes a model where one's choices are the cooperative solution of a bargaining problem between two preference relations. Depending on the specific relations, their model may be consistent with certain instances of the attraction effect (as well as other choice anomalies, such as the compromise effect). Another paper that introduces a choice model for the attraction effect phenomenon is Lombardi (2009). Lombardi's model is of the following form: First, from any choice problem S, the agent selects the elements of S that are maximal according to some acyclic binary relation, and second, she eliminates from the resulting maximal set those alternatives whose lower contour sets are strictly contained in that of some other maximal alternative. A similar model is obtained by Gerasimou (2012) as well; the chosen alternatives in Gerasimou's model are those that are undominated, but that dominate at least one other alternative, in terms of a partial order.

These models do not analyze the attraction effect phenomenon as a consequence of reference-dependent choice. Even more important, any one of these models are compatible with only a "weak" form of this phenomenon. In particular, these models allow for the addition of an asymmetrically dominated option to help the agent decide between two alternatives to which she reveals to be *indifferent* in a pairwise comparison, but none of them allows for this addition to lead the agent to alter her strict ranking of choice items. Put more precisely, all of these models are compatible with the following choice data: $\mathbf{c}\{x, y\} = \{x, y\}$ and $\mathbf{c}\{x, y, z\} = \{y\}$ – the introduction of z breaks the indifference between x and y –, but none of them is compatible with the choice data $\mathbf{c}\{x, y\} = \{x\}$ and $\mathbf{c}\{x, y, z\} = \{y\}$ – the introduction of z leads the agent to alter her choice from x to y. (In particular, if we specialize these models to the case of choice *functions,* none of them remain compatible with the attraction effect phenomenon.) Yet, an overwhelming part of the data on the attraction effect is of the latter form (and is in fact presented in terms of choice functions). The choice models that are developed in this paper are fully compatible with this version of the attraction effect phenomenon, and thus conform with the related empirical evidence.

Kamenica (2008) studies a clever model in which there is a market with rational consumers some of whom are informed and some of whom are uninformed, with uninformed consumers exhibiting a behavior in equilibrium that conform with the attraction effect (as well as the compromise effect) phenomenon. In this model, choice anomalies are not seen as violations of "rationality," but they rather emerge as equilibrium behavior in a specific market environment. By contrast, in the present paper, we do regard the attraction effect phenomenon as a violation of WARP. This is motivated by the presence of this phenomenon in very different sorts of economic environments, as well as in laboratory experiments where the informational structure does not seem to be the source of the problem.

In a nutshell, the choice models developed in this paper have the structure of rationalization by mentally-constrained optimization, that is, they posit that the decision-maker focus in a given choice situation S only on $S \cap Q(\mathbf{r}(S))$, the set of options that dominate the (endogenously determined) reference point. Masatlioglu, Nakajima and Ozbay (2011) present a related model in which there are finitely many alternatives, choice behavior is modeled through choice *functions*, and subjects focus only on a subset $\Gamma(S)$ of the choice problem S, finalizing their choices by maximizing utility over those elements. Their interpretation is that $\Gamma(S)$ is the subset of the alternatives in S on which the agent focuses her attention. They posit that the self-map Γ on $2^X \backslash \{\emptyset\}$ is such that $\Gamma(S) = \Gamma(S \backslash \{x\})$ holds whenever $x \in S \backslash \Gamma(S)$, and refer to any such map as an *attention filter*. If we restrict our attention to the case where X is finite, and work only with a choice function $\mathbf{c}$, then the model characterized in Theorem 1 can be seen as a special case of that model. For, if $\mathbf{c}$ is a choice function on $2^X \backslash \{\emptyset\}$ represented by a reference-dependent choice model $\langle u, \mathbf{r}, Q \rangle$, then by setting $\Gamma(S) := \{\mathbf{c}(S)\} \cup \{x \in X : u(\mathbf{c}(S)) > u(x)\}$, we find that Γ is an attention filter and $\{\mathbf{c}(S)\} = \arg\max u(\Gamma(S))$ for every nonempty $S \subseteq X$. However, by contrast to our model, that of Masatlioglu, et al. (2011) is not tailored toward addressing the issue of reference-dependent choice; that model is simply too general to be useful for this purpose; it violates, in general, the properties of No-Cycle, RCoh and RCon altogether.[18]

The model of Cherepanov, Feddersen and Sandroni (2010) is also built around the idea of a mentally-constrained optimization choice procedure. In that paper, where X is again assumed to be finite, a pair $(\succsim, \psi)$ is referred to as a *model of behavior* when $\succsim$ is a binary relation on X, and ψ a map such that $\emptyset \neq \psi(S) \subseteq S$ for all choice problems S. Cherepanov, et al. say that a choice correspondence $\mathbf{c}$ is represented by the model of behavior $(\succsim, \psi)$ if, for every choice problem S, $\mathbf{c}(S)$ is the set of the alternatives that maximize $\succsim$ in $\psi(S)$. In turn, Cherepanov, et al. (2010) provides an axiomatic foundation for the case in which ψ is induced by a collection Λ of rationales (binary relations on X) in the sense that, for each choice problem S, $\psi(S)$ is found as the set of elements of S that are maximal with respect to at least one of the rationales in Λ. Our reference-dependent choice model is clearly related to this representation. If $\mathbf{c}$ is a choice correspondence on $2^X \backslash \{\emptyset\}$ that is represented by a reference-dependent choice model $\langle u, \mathbf{r}, Q \rangle$, there is a natural way of defining a model of behavior $(\succsim, \psi)$, where $\succsim$ is the preference relation on X that is represented by u, and $\psi(S) := Q(\mathbf{r}(S))$ for each choice problem S. However, the two models are not nested. On the one hand, there are instances that are allowed by the Cherepanov-Feddersen-Sandroni

[18] Lleras et al. (2010) study a related choice model where the defining property of the map Γ is that $x \in \Gamma(S)$ implies $x \in \Gamma(T)$ for every $T \subseteq S$ with $x \in T$. Simple examples would show that the resulting choice model and that of Theorem 1 are not nested.

model but not by ours – like cyclic choice in pairwise choice situations. On the other hand, it is in general not true that the map $S \mapsto Q(\mathbf{r}(S))$ is induced by a set of rationales; in fact, the former model violates both RCoh and RCon.[19]

Finally, we recall that our representation of a choice correspondence is built on a two-stage decision making procedure: in the first stage the decision-maker eliminates all alternatives that are not highlighted by the reference point, and in the second stage she selects the alternatives that yield the highest utility among the highlighted ones. There are other papers that have investigated such sequential decision making procedures. For example, Manzini and Mariotti (2007) and Rubinstein and Salant (2006) axiomatize choice *functions* that can be represented as if the agent applies two binary relations, one after the other:[20] first, she eliminates all elements in a problem that are not maximal relative to the first relation, and second, among the remaining elements chooses that which is optimal according to the second relation. The connection between these models and ours is not tight, however. Both of these models are primed to capture cycles that may emanate in pairwise choice situations. By No-Cycle, we rule out such violations of WARP, while it is known that either of these models reduce to the standard rational choice model if they were to satisfy No-Cycle. Put differently, the intersection of the class of choice correspondences that we have characterized here and those of Manzini and Mariotti (2007) and Rubinstein and Salant (2006) contains only the classical rational choice model (for choice functions).

A Proofs

A.1 Proof of Theorem 1

We begin with a preliminary observation. Let $\mathbf{c}$ be a choice correspondence on $\mathfrak{X}$ that satisfies RCon, and take any $\mathbf{c}$-awkward $S \in \mathfrak{X}$. Let $\mathcal{T}_S$ stand for the collection of all T in $\mathfrak{X} \cap 2^S$ such that $\mathbf{c}(S) \cap T \neq \emptyset$ and $\mathbf{c}(T) \neq \mathbf{c}(S) \cap T$. As S is $\mathbf{c}$-awkward, $\mathcal{T}_S \neq \emptyset$. Besides, if $S = \bigcup\{T : T \in \mathcal{T}_S\}$, then $\mathcal{T}_S$ is a $\mathbf{c}$-cover of S, so by RCon, there is a $T \in \mathcal{T}$ with $\mathbf{c}(T) = \mathbf{c}(S) \cap T$, which is impossible in view of the definition of $\mathcal{T}_S$. It follows that $\bigcup\{T : T \in \mathcal{T}_S\}$ is a proper subset of S, that is, there is some $z \in S$ such that $\mathbf{c}(T) = \mathbf{c}(S) \cap T$ for all $T \in \mathfrak{X} \cap 2^S$ with $\mathbf{c}(S) \cap T \neq \emptyset$ and $z \in T$.

We have proved the "only if" part of the following characterization, whose "if" part is straightforward.

Lemma A.1. Let $\mathbf{c}$ be a choice correspondence on $\mathfrak{X}$. Then, $\mathbf{c}$ satisfies RCon if, and only if, for any $\mathbf{c}$-awkward set S in $\mathfrak{X}$, there exists a $z \in S$ such that $\mathbf{c}\,(T) = \mathbf{c}\,(S) \cap T$ for all $T \in \mathfrak{X} \cap 2^S$ with $\mathbf{c}\,(S) \cap T \neq \emptyset$ and $z \in T$.

[19] Let $X := \{x, y, z, w\}$, and consider the choice function $\mathbf{c}$ on $2^X \backslash \{\emptyset\}$ defined by $\mathbf{c}\{x, y\} = \mathbf{c}\{x, z\} = \mathbf{c}\{x, w\} = \mathbf{c}\{x, y, z\} = \mathbf{c}\,\{x, z, w\} := x$, $\mathbf{c}\{y, z\} = \mathbf{c}\,\{y, w\} = \mathbf{c}\,\{x, y, w\} = \mathbf{c}\,\{y, z, w\} := y$, $\mathbf{c}\,\{z, w\} := z$ and $\mathbf{c}(X) := x$. This function cannot be represented as in the Cherepanov-Feddersen-Sandroni model, but it can be represented by a reference-dependent choice model (in fact, by one that is represented as in Theorem 8).

[20] For papers that work with sequential procedures with more than two stages, see Apestegia and Ballester (2012) and Manzini and Mariotti (2012).

We now move to the proof of Theorem 1.

[*Necessity*] Let $\mathbf{c}$ be a continuous choice correspondence on $\mathfrak{X}$ which is represented by a reference-dependent choice model $\langle u, \mathbf{r}, Q\rangle$ on $\mathfrak{X}$. It readily follows from this representation that, for any $x, y \in X$, we have $x \in \mathbf{c}\{x,y\}$ iff $u(x) \geq u(y)$. It is also plain that $u(x) = u(y)$ for any $x, y \in X$ with $\{x, y\} \subseteq \mathbf{c}(S)$ for some $S \in \mathfrak{X}$. Thus $\mathbf{c}$ satisfies No-Cycle and RI. On the other hand, Proposition 5 entails that $\mathbf{c}$ satisfies RCoh. Finally, suppose S is a $\mathbf{c}$-awkward set in $\mathfrak{X}$ so that $\mathbf{r}(S) \neq \diamondsuit$, by Proposition 2. Put $z := \mathbf{r}(S)$, and notice that (3) entails that $\mathbf{c}(T) = \mathbf{c}(S) \cap T$ for all $T \in \mathfrak{X} \cap 2^S$ with $\mathbf{c}(S) \cap T \neq \emptyset$ and $z \in T$. In view of the arbitrariness of S and Lemma A.1, therefore, we may conclude that $\mathbf{c}$ satisfies RCon.

[*Sufficiency*] Suppose that $\mathbf{c}$ is a continuous choice correspondence on $\mathfrak{X}$ that satisfies the four axioms in the statement of the theorem. Define the binary relation $\succsim \subseteq X \times X$ by $x \succsim y$ iff $x \in \mathbf{c}\{x, y\}$. Since $\mathbf{c}$ is a continuous choice correspondence that satisfies No-Cycle, $\succsim$ is a complete and continuous preorder on X. Given that X is a separable metric space, therefore, we may invoke Debreu's Utility Representation Theorem to find a continuous function $u : X \to \mathbb{R}$ such that $x \succsim y$ iff $u(x) \geq u(y)$ for every $x, y \in X$. This implies $\mathbf{c}\{x, y\} = \arg\max u(\{x, y\})$ for every $x, y \in X$.

Claim A.1. A set $S \in \mathfrak{X}$ is not $\mathbf{c}$-awkward if and only if $\mathbf{c}(S) = \arg\max u(S)$.

Proof. Assume $S \in \mathfrak{X}$ is not $\mathbf{c}$-awkward, and pick any $(x, y) \in \arg\max u(S) \times \mathbf{c}(S)$. Then, $x \in \mathbf{c}\{x, y\}$ and $y \in \mathbf{c}(S) \cap \{x, y\}$, while $\mathbf{c}\{x, y\} = \mathbf{c}(S) \cap \{x, y\}$ (because S is not $\mathbf{c}$-awkward). It follows that $x \in \mathbf{c}(S)$ and $y \in \arg\max u(S)$. Thus: $\mathbf{c}(S) = \arg\max u(S)$. The converse assertion follows readily from the definition of $\mathbf{c}$-awkwardness and the choice of u.

Claim A.2. For each $\mathbf{c}$-awkward $S \in \mathfrak{X}$, there exists a $z \in S$ such that z is a revealed $\mathbf{c}$-reference for every element of $\mathbf{c}(S)$ and $\mathbf{c}(T) = \mathbf{c}(S) \cap T$ for all $T \in \mathfrak{X} \cap 2^S$ with $\mathbf{c}(S) \cap T \neq \emptyset$ and $z \in T$.

Proof. Take any $\mathbf{c}$-awkward S in $\mathfrak{X}$. It follows from RI and the definition of u that $u(x) = u(w)$ for all $x, w \in \mathbf{c}(S)$. Thus, by Claim 1, there exists an alternative y in $\arg\max u(S) \backslash \mathbf{c}(S)$. By RCon and Lemma A.1, we know that there is a $z \in S$ such that $\mathbf{c}(T) = \mathbf{c}(S) \cap T$ for all $T \in \mathfrak{X} \cap 2^S$ with $\mathbf{c}(S) \cap T \neq \emptyset$ and $z \in T$. This implies that y cannot belong to $\mathbf{c}\{x, y, z\}$ for any $x \in \mathbf{c}(S)$. Since $u(y) \geq u(x)$ for all $x \in \mathbf{c}(S)$, we conclude that z is a revealed $\mathbf{c}$-reference for every element of $\mathbf{c}(S)$.

For each $z \in X$, define $A(z) := \{x \in X : z$ is a revealed $\mathbf{c}$-reference for $x\}$ and $B(z) := \{x \in X : z$ is a potential $\mathbf{c}$-reference for $x\}$. Clearly, RCoh implies $A(z) \subseteq B(z)$ for all $z \in X$. Now, for any $z \in X$, pick any $Q : \mathfrak{X} \cup \{\diamondsuit\} \rightrightarrows X$ with $Q(\diamondsuit) = X$ and

$$A(z) \subseteq Q(z) \subseteq B(z).$$

Next, we define the map $\mathbf{r} : \mathfrak{X} \to X \cup \{\diamondsuit\}$ as follows: (i) If $S \in \mathfrak{X}$ is not $\mathbf{c}$-awkward, then $\mathbf{r}(S) := \diamondsuit$; and (ii) if $S \in \mathfrak{X}$ is $\mathbf{c}$-awkward, then $\mathbf{r}(S)$ is set to be any one $z \in S$ such that z is a revealed $\mathbf{c}$-reference for all $x \in \mathbf{c}(S)$ and $\mathbf{c}(T) = \mathbf{c}(S) \cap T$ for all $T \in \mathfrak{X} \cap 2^S$ with $\mathbf{c}(S) \cap T \neq \emptyset$ and $z \in T$. By Claim A.2, and the Axiom of Choice, $\mathbf{r}$ is well-defined, and it is, thus, a reference map on $\mathfrak{X}$. Furthermore, the following observation is immediate from these definitions:

Claim A.3. If $S \in \mathfrak{X}$ is not $\mathbf{c}$-awkward, then $\mathbf{c}(S) = \arg\max u(S \cap Q(\mathbf{r}(S)))$.

Finally:

Claim A.4. If $S \in \mathfrak{X}$ is $\mathbf{c}$-awkward, then $\mathbf{c}(S) = \arg\max u(S \cap Q(\mathbf{r}(S)))$.

Proof. Let $S \in \mathfrak{X}$ be $\mathbf{c}$-awkward. By definition of $\mathbf{r}$, we have $\mathbf{r}(S) \in S$ and $\mathbf{c}(S) \subseteq Q(\mathbf{r}(S))$. Pick any $x \in \mathbf{c}(S)$ and $y \in Q(\mathbf{r}(S)) \cap S$. Again by definition of $\mathbf{r}$, we have $x \in \mathbf{c}\{x, y, \mathbf{r}(S)\} = \mathbf{c}(S) \cap \{x, y, \mathbf{r}(S)\}$. It then follows from the definition of Q that $\mathbf{r}(S)$ is a potential $\mathbf{c}$-reference for y, which implies that we cannot have $\{y\} = \mathbf{c}\{x, y\}$. That is, $u(x) \geq u(y)$, and we conclude that $\mathbf{c}(S) \subseteq \arg\max u(S \cap Q(\mathbf{r}(S)))$. Now pick any $y \in \arg\max u(S \cap Q(\mathbf{r}(S)))$ and $x \in \mathbf{c}(S)$. By the previous observation, we have $\{x, y\} = \mathbf{c}\{x, y\}$. Since $\mathbf{r}(S)$ is a potential $\mathbf{c}$-reference for y and $x \in \mathbf{c}\{x, y, \mathbf{r}(S)\}$, this implies $y \in \mathbf{c}\{x, y, \mathbf{r}(S)\}$. By definition of $\mathbf{r}$, therefore, $y \in \mathbf{c}(S)$. Conclusion: $\mathbf{c}(S) = \arg\max u(S \cap Q(z))$.

It remains to show that u, $\mathbf{r}$ and Q satisfy (2), and, for every $S \in \mathfrak{X}$, $\mathbf{r}(S) \neq \diamond$ only if $\arg\max u(S) \neq \arg\max u(S \cap Q(\mathbf{r}(S)))$. The former is an easy consequence of the definitions of Q and $\mathbf{r}$, while the later is an immediate consequence of the definition of $\mathbf{r}$ and Claim 1.

A.2 Proof of Proposition 6

For any given reference-dependent choice model $\langle u, \mathbf{r}, Q \rangle$, (4) ensures that Q^* depends only on the choice correspondence $\mathbf{c}$ that $\langle u, \mathbf{r}, Q \rangle$ represents. We next show that the same is true for the reference correspondence $\mathbf{R}$ associated with this model.

Lemma A.2. Let $\mathbf{c}$ be a choice correspondence on $\mathfrak{X}$ that is represented by a reference-dependent choice model $\langle u, \mathbf{r}, Q \rangle$, and let S be a $\mathbf{c}$-awkward choice problem in $\mathfrak{X}$. Then, $z \in \mathbf{R}(S)$ iff (i) $z \in S$, (ii) z is a revealed $\mathbf{c}$-reference for each $x \in \mathbf{c}(S)$, and (iii) $\mathbf{c}(T) = \mathbf{c}(S) \cap T$ for every $T \in \mathfrak{X} \cap 2^S$ with $\mathbf{c}(S) \cap T \neq \emptyset$ and $z \in T$. (In particular, $\mathbf{c}$ determines $\mathbf{R}$ completely, since for non $\mathbf{c}$-awkward choice problems Proposition 2 guarantees that $\mathbf{R}(S) = \diamond$.)

Proof. Take any $z \in S$ that satisfies the properties (ii) and (iii). As S is $\mathbf{c}$-awkward, it follows from these properties and the construction of the reference-dependent choice model given in the proof of Theorem 1 that there is a reference map $\mathbf{r}\prime$ on $\mathfrak{X}$ such that $\mathbf{r}\prime(S) = z$ and $\langle u, \mathbf{r}\prime, Q \rangle$ also represents $\mathbf{c}$. Then, by (2), we have $\mathbf{c}(T) = \arg\max u(T \cap Q(z))$ for any $T \in \mathfrak{X} \cap 2^S$ with $\mathbf{c}(S) \cap T \neq \emptyset$ and $z \in T$. Conclusion: $z \in \mathbf{R}(S)$. Conversely, assume that z is an element of $\mathbf{R}(S)$. Obviously, $z \in S$. Besides, as S is $\mathbf{c}$-awkward and $\langle u, \mathbf{r}, Q \rangle$ represents $\mathbf{c}$, we may use Proposition 2 to find a y in $\arg\max u(S) \backslash \mathbf{c}(S)$. Since $z \in \mathbf{R}(S)$, we have $\mathbf{c}(S) = \arg\max u(S \cap Q(z))$, so, we find that y does not belong to $Q(z)$. Pick any $x \in \mathbf{c}(S)$. As $u(y) \geq u(x)$, we have $y \in \mathbf{c}\{x, y\}$, but $z \in \mathbf{R}(S)$ implies

$$\{x, y\} \cap \mathbf{c}\{x, y, z\} = \{x, y\} \cap \arg\max u(\{x, y, z\} \cap Q(z)) = \{x\}.$$

Thus: z is a revealed $\mathbf{c}$-reference for x. On the other hand, that z satisfies property (iii) is an immediate consequence of the fact that $\langle u, \mathbf{r}, Q \rangle$ represents $\mathbf{c}$. Our lemma is thus proved.

The following is a useful consequence of this observation.

Lemma A.3. Let $\mathbf{c}$ be a choice correspondence on $\mathfrak{X}$ that is represented by a reference-dependent choice model $\langle u, \mathbf{r}, Q \rangle$. Let $f : u(X) \to \mathbb{R}$ be a continuous and strictly increasing map, and take any $\boldsymbol{\rho} : \mathfrak{X} \to X \cup \{\diamondsuit\}$

such that $\boldsymbol{\rho}(S) \in \mathbf{R}(S)$ for all $S \in \mathfrak{X}$. Then, $\langle f \circ u, \boldsymbol{\rho}, Q^* \rangle$ is a reference-dependent choice model on $\mathfrak{X}$ that represents $\mathbf{c}$.

Proof. In the notation of the proof of Theorem 1, we have here $A(z) = Q^*(z)$ for all $z \in X$; recall (4). Furthermore, for any $S \in \mathfrak{X}$, we have $\boldsymbol{\rho}(S) = \diamondsuit$ (because $\mathbf{R}(S) = \{\diamondsuit\}$) if S is not $\mathbf{c}$-awkward, while, by Lemma A.2, when S is $\mathbf{c}$-awkward, $\boldsymbol{\rho}(S)$ is a revealed $\mathbf{c}$-reference for all $x \in \mathbf{c}(S)$ and $\mathbf{c}(T) = \mathbf{c}(S) \cap T$ for every $T \in \mathfrak{X} \cap 2^S$ with $\mathbf{c}(S) \cap T \neq \emptyset$ and $\boldsymbol{\rho}(S) \in T$. The argument given in the proof of Theorem 1 then applies to show that $\langle u, \boldsymbol{\rho}, Q^* \rangle$, and hence, $\langle f \circ u, \boldsymbol{\rho}, Q^* \rangle$, represents $\mathbf{c}$.

We now move to the proof of Proposition 6. To streamline the argument, let us define the binary relation $\approx$ on the set of all reference-dependent choice models on $\mathfrak{X}$ by $\langle u, \mathbf{r}, Q \rangle \approx \langle u', \mathbf{r}', Q' \rangle$ iff $\langle u, \mathbf{r}, Q \rangle$ and $\langle u', \mathbf{r}', Q' \rangle$ represent the same choice correspondence on $\mathfrak{X}$. (Clearly, $\approx$ is an equivalence relation.) We can then restate Proposition 6 as follows: Let $\langle u_1, \mathbf{r}_1, Q_1 \rangle$ and $\langle u_2, \mathbf{r}_2, Q_2 \rangle$ be two reference-dependent choice models on $\mathfrak{X}$. We have $\langle u_1, \mathbf{r}_1, Q_1 \rangle \approx \langle u_2, \mathbf{r}_2, Q_2 \rangle$ iff $u_2 = f \circ u_1$ for some continuous and strictly increasing $f : u_1(X) \to \mathbb{R}$, $\mathbf{R}_1 = \mathbf{R}_2$ and $Q_1^* = Q_2^*$.

[*Necessity*] Suppose $\langle u_1, \mathbf{r}_1, Q_1 \rangle \approx \langle u_2, \mathbf{r}_2, Q_2 \rangle$. Then, u_1 and u_2 are continuous real functions on X that represent the same complete preorder on X (that arises from choices over pairwise choice problems), and hence, they must be continuous and strictly increasing transformations of each other. Furthermore, for each $i = 1, 2$, Lemma A.2 and (4) ensure that $\mathbf{R}_i$ and Q_i^* depend only on the choice correspondence that $\langle u_i, \mathbf{r}_i, Q_i \rangle$ represents. As these models represent the same choice correspondence, therefore, we must have $\mathbf{R}_1 = \mathbf{R}_2$ and $Q_1^* = Q_2^*$.

[*Sufficiency*] Suppose that $\langle u_1, \mathbf{r}_1, Q_1 \rangle$ and $\langle u_2, \mathbf{r}_2, Q_2 \rangle$ are two reference-dependent choice models on $\mathfrak{X}$ such that u_2 is a continuous and strictly increasing transformation of u_1, $\mathbf{R}_1 = \mathbf{R}_2$ and $Q_1^* = Q_2^*$. Since $\mathbf{r}_2(S) \in \mathbf{R}_2(S) = \mathbf{R}_1(S)$, for every $S \in \mathfrak{X}$, Lemma A.3. implies that $\langle u_1, \mathbf{r}_1, Q_1 \rangle \approx \langle u_2, \mathbf{r}_2, Q_2^* \rangle$. But, applying Lemma A.3 again, we find $\langle u_2, \mathbf{r}_2, Q_2^* \rangle \approx \langle u_2, \mathbf{r}_2, Q_2 \rangle$. As $\approx$ is transitive, therefore, $\langle u_1, \mathbf{r}_1, Q_1 \rangle \approx \langle u_2, \mathbf{r}_2, Q_2 \rangle$, as we sought.

A.3 Proof of Theorem 8

[*Necessity*] By Theorem 1, we know that $\mathbf{c}$ is a continuous choice correspondence on $\mathfrak{X}$ that satisfies No-Cycle, RCon and RI. Moreover, by Proposition 5, if z is a revealed $\mathbf{c}$-reference for x, then $x \in Q(z)$ while if $x \in Q(z)$, then z is a potential reference for x. RA is then a straightforward consequence of these two facts and the structure of Q.

[*Sufficiency*] Define the relation $\trianglerighteq$ on X by $x \trianglerighteq y$ iff either y is a revealed $\mathbf{c}$-reference for x or $x = y$. Let $\succcurlyeq$ stand for the transitive closure of $\trianglerighteq$. From RA we know that $x \succcurlyeq y$ implies that y is a potential $\mathbf{c}$-reference for x. But in the proof of Theorem 1 we have showed that any $Q : \mathfrak{X} \cup \{\diamondsuit\} \rightrightarrows X$ such that (i) $x \in Q(y)$ whenever y is a revealed $\mathbf{c}$-reference for x, and (ii) y is a potential $\mathbf{c}$-reference for x whenever $x \in Q(y)$, can be used for that type of representation. Therefore, if we define $Q(y) := \{x : x \succcurlyeq y\}$ for all $y \in X$, we obtain a representation as in the statement of Theorem 1. To complete the proof of Theorem 8 all we have to do then is to use the fact that any preorder has a multi-utility representation (cf. Evren and Ok (2012)).

A.4 Independence of the Axioms

In this section we show the axioms that we imposed in Theorems 1 and 8 are independent.

Example A.1. Let $X := \{x, y, z\}$, and consider the choice function $\mathbf{c}$ on $2^X \backslash \{\emptyset\}$ defined by $\mathbf{c}\{x, y\} := x$, $\mathbf{c}\{y, z\} := y$, $\mathbf{c}\{x, z\} := z$ and $\mathbf{c}(X) := x$. Clearly, $\mathbf{c}$ violates No-Cycle, but it satisfies RCon and RI. The only revealed $\mathbf{c}$-reference here is y, which is a revealed $\mathbf{c}$-reference for x. It follows that $\mathbf{c}$ satisfies RA (and hence RCoh) as well.

Example A.2. Let $X := \{x, y, z, w\}$, and consider the choice function $\mathbf{c}$ on $2^X \backslash \{\emptyset\}$ defined by $\mathbf{c}\{x, y\} = \mathbf{c}\{x, z\} = \mathbf{c}\{x, w\} = \mathbf{c}\{x, y, z\} = \mathbf{c}\{x, y, w\} = \mathbf{c}\{x, z, w\} := x$, $\mathbf{c}\{y, z\} = \mathbf{c}\{y, w\} = \mathbf{c}\{y, z, w\} := y$, and $\mathbf{c}\{z, w\} = \mathbf{c}(X) := z$. It is easy to see that $\mathbf{c}$ satisfies No-Cycle and RI. Also, there is no revealed $\mathbf{c}$-reference in this example, so $\mathbf{c}$ also satisfies RA (and hence RCoh). On the other hand, $\mathbf{c}$ violates RCon. For instance, $\mathcal{T} := \{\{x, z\}, \{y, z, w\}\}$ is a $\mathbf{c}$-cover of X, but we have $\mathbf{c}(X) \cap T \neq \mathbf{c}(T)$ for each $T \in \mathcal{T}$.

Example A.3. Let $X := \{x, y, z, w\}$, and consider the choice function $\mathbf{c}$ on $2^X \backslash \{\emptyset\}$ defined by $\mathbf{c}\{x, y\} = \mathbf{c}\{x, z\} = \mathbf{c}\{x, w\} = \mathbf{c}\{x, y, z\} := x$, $\mathbf{c}\{y, z\} = \mathbf{c}\{y, w\} = \mathbf{c}\{x, y, w\} := y$, $\mathbf{c}\{z, w\} = \mathbf{c}\{x, z, w\} = \mathbf{c}\{y, z, w\} := z$ and $\mathbf{c}(X) := z$. It is easy to check that $\mathbf{c}$ satisfies No-Cycle, RCon and RI. On the other hand, w is a revealed $\mathbf{c}$-reference for y, but w is not a potential $\mathbf{c}$-reference for y (because $\mathbf{c}\{y, z\} = y$ and yet $\mathbf{c}\{y, z, w\} = z$). Thus, $\mathbf{c}$ violates RCoh (and hence RA).

Example A.4. Let $X := \{x, y, z\}$, and consider the choice correspondence $\mathbf{c}$ on $2^X \backslash \{\emptyset\}$ defined by $\mathbf{c}\{x, y\} = \mathbf{c}\{x, z\} := \{x\}$, $\mathbf{c}\{y, z\} := \{y\}$, and $\mathbf{c}(X) := \{x, y\}$. It is plain that $\mathbf{c}$ satisfies No-Cycle but not RI. On the other hand, the only revealed $\mathbf{c}$-reference here is z which is a revealed $\mathbf{c}$-reference for y. It follows that $\mathbf{c}$ satisfies RA (and hence RCoh). It is also easily checked that $\mathbf{c}$ satisfies RCon.

References

Apestegia, J. and M. Ballester (2012), "Choice by sequential procedures," mimeo, Universitat Pompeu Fabra.

Ariely, D. and T. Wallsten (1995), "Seeking Subjective Dominance in Multi-Dimensional Space: An Explanation of the Asymmetric Dominance Effect," *Organizational Behavior and Human Decision Processes,* 63, 223-232.

Arrow, K. (1959), "Rational Choice Functions and Orderings," *Economica,* 26, 121-127.

Bateman, I., A. Munro, and G. Poe (2008), "Decoy Effects in Choice Experiments and Contingent Valuation: Asymmetric Dominance," *Land Economics*, 84, 115-127.

Bernheim, D. and A. Rangel (2009), "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics," *Quarterly Journal of Economics*, 124, 51-104.

Bhargava, M., J. Kim, and R. Srivastava (2000), "Explaining Context Effects on Choice Using a Model of Comparative Judgement," *Journal of Consumer Psychology*, 9, 167-177.

Burton, S. and G. Zinkhan (1987), "Changes in Consumer Choice: Further Investigation of Similarity and Attraction Effects," *Psychology and Marketing*, 4, 255-266.

Camerer, C. (1995), "Individual Decision Making," in J. Kagel and A. Roth, eds., *Handbook of Experimental Economics*, Princeton: Princeton University Press.

Chateauneuf, A. and P. Wakker (1999), "An Axiomatization of Cumulative Prospect Theory for Decision under Risk," *Journal of Risk and Uncertainty*, 18, 137-145.

Cherepanov, V., T. Federsen and A. Sandroni (2010), "Rationalization," mimeo, Northwestern University.
de Clippel, G. and K. Eliaz (2012), "Reason-Based Choice: A Bargaining Rationale for the Attraction and Compromise Effects," *Theoretical Economics*, 7, 125-162.
Diecidue, E. and J. de Ven (2008), "Aspiration Level, Probability of Success and Failure, and Expected Utility," *International Economic Review*, 49, 683-700.
Doyle J., D. O'Connor, G. Reynolds, and P. Bottomley (1999), "The Robustnes of hte Asymmetrically Dominated Effect: Buying Frames, Phantom Alternatives, and In-Store Purchases," *Psychology and Marketing*, 16, 225-243.
Esteban, S. and E. Miyagawa (2006), "Temptation, Self-Control, And Competitive Nonlinear Pricing," *Economic Letters*, 90, 348-355.
Esteban, S., Miyagawa, E. and M. Shum (2007), "Nonlinear Pricing with Self-Control Preferences," *Journal of Economic Theory*, 135, 306-338.
Evren, O. and E. A. Ok (2011), "On the Multi-Utility Representation of Preference Relations," *Journal of Mathematical Economics*, 47, 554-563.
Gerasimou, G. (2012), "Choice in the Presence of Attractions," mimeo, University of St. Andrews.
Gul, F. and W. Pesendorfer (2001), "Temptation and Self-Control," *Econometrica,* 69, 1403-1435.
Herne, K. (1997), "Decoy Alternatives in Policy Choices: Asymmetric Domination and Compromise Effects," *European Journal of Political Economy*, 13, 575-589.
Herne, K. (1999), "The Effects of Decoy Gambles on Individual Choice," *Experimental Economics*, 2, 31-40.
Highhouse, S. (1996), "Context-Dependent Selection: Effects of Decoy and Phantom Job Candidates," *Organizational Behavior and Human Decision Processes*, 13, 575-589.
Huber, J., J. Payne, and C. Puto (1982), "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *Journal of Consumer Research,* 9, 90-98.
Kahneman, D. and A. Tversky (1979), "Prospect Theory: An Analysis of Choice under Risk," *Econometrica*, 47: 263-291.
Kamenica, E. (2008), "Contextual Inference in Markets: On the Informational Content of Product Lines," *American Economic Review*, 98, 2127-2149.
Kivetz, R. (1999), "Advances in Research on Mental Accounting and Reason-Based Choice," *Marketing Letters*, 10, 249-266.
Köszegi, B. and M. Rabin (2006), "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 121, 1133-1165.
Köszegi, B. and M. Rabin (2007), "Reference-Dependent Risk Attitudes," *American Economic Review*, 97, 1047-1073.
Lehman, D. and Y. Pan (1994), "Context Effects, New Brand Entry and Consideration Sets," *Journal of Marketing Research,* 31, 364-374.
Lleras, J., Y. Masatlioglu, D. Nakajima, and E. Ozbay (2010), "When more is less: Choice by Limited Consideration," mimeo, University of Maryland.
Lombardi, M. (2009). "Reason-Based Choice Correspondences," *Mathematical Social Sciences*, 57, 58–66.
Luce, D. and H. Raiffa, *Games and Decisions*, Wiley, New York, 1957.
Manzini, P. and M. Mariotti (2007), "Sequentially Rationalizable Choice," *American Economic Review*, 97, 1824–1839.
Manzini, P. and M. Mariotti (2010), "Revealed Preferences and Boundedly Rational Choice Procedures: an Experiment," mimeo, University of St. Andrews.

Manzini, P. and M. Mariotti (2012), "Choice by Lexicographic Semiorders," *Theoretical Economics*, 7, 1-23.

Masatlioglu, Y. and E. A. Ok (2005), "Rational Choice with Status-Quo Bias," *Journal of Economic Theory*, 121, 1-29.

Masatlioglu, Y. and E. A. Ok (2012), "A Canonical Model of Choice with Initial Endowments," mimeo, NYU.

Masatlioglu, Y., D. Nakajima, and E. Ozbay (2011), "Revealed attention," *American Economic Review*, forthcoming.

Mishra, S., U. Umesh, and D. Stem (1993), "Antecedents of the Attraction Effect: An Information Processes Approach," *Journal of Marketing Research,* 30, 331-349.

Mussa, M. and S. Rosen (1978), "Monopoly and Product Quality," *Journal of Economic Theory*, 18: 301-317.

Ok, E.A., P. Ortoleva, and G. Riella (2011), "Theory of Product Differentiation in the Presence of the Attraction Effect," *mimeo*, NYU.

Ortoleva, P. (2010), "Status Quo Bias, Multiple Priors and Uncertainty Aversion," *Games and Economic Behavior*, 69, 411-429.

Pan, Y., S. O'Curry, and R. Pitts (1995), "The Attraction Effect and Political Choice in two Elections," *Journal of Consumer Psychology*, 4, 85-101.

Rabin, M., (1998), "Psychology and Economics," *Journal of Economic Literature*, 36: 11-46.

Ratneshwar, S., A. Shocker, and D. Stewart (1987), "Toward Understanding the Attraction Effect: The Implications of Product Stimulus Meaningfulness and Familiarity," *Journal of Consumer Research*, 13: 520-533.

Redelmeier, D. and E. Shafir (1995), "Medical Decision Making in Situations that Offer Multiple Alternatives," *Journal of American Medical Association*, 273: 302-305.

Richter, M. (1966), "Revealed Preference Theory," *Econometrica,* 34, 625-645.

Rubinstein, A. and Y. Salant (2006), "Two Comments on the Principle of Revealed Preference," mimeo, NYU.

Rubinstein, A. and Y. Salant (2008), "(A, f) : Choice with Frames," *Review of Economic Studies*, 75, 1287-1296.

Sagi, J. (2006), "Anchored Preference Relations," *Journal of Economic Theory*, 130, 283-295.

Schwartz, R. and G. Chapman (1999), "Are More Options Always Better? The Attraction Effect in Physicians' Decisions about Medications," *Medical Decision Making*, 19: 315-323.

Schwarzkopf, O. (2003), "The Effects of Attraction on Investment Decisions," *Journal of Behavioral Finance,* 4, 96-108.

Sen, S. (1998), "Knowledge, Information Mode, and the Attraction Effect," *Journal of Consumer Research*, 25: 64-77.

Shafir, S., T. Waite, and B. Smith (2002), "Context Dependent Violations of Rational Choice in Honeybees (Apis Mellilea) and Gray Jays (Perisoreus Canadensis)," *Behavioral Ecology and Sociobiology*, 51, 186-187.

Simonson, I. (1989), "Choice Based on Reasons: The Case of Attraction and Compromise Effects," *Journal of Conflict Resolution*, 16: 158-175.

Simonson, I. and S. Nowlis (2000), "The Role of Explanations and Need for Uniqueness in Consumer Decision Making: Unconventioanal Choices based on Reasons," *Journal of Consumer Research*, 27, 49-68.

Simonson, I. and A. Tversky (1992), "Choice in Context: Tradeoff Contrast and Extremeness Aversion," *Journal of Marketing Research*, 29, 281-295.

Sivakumar, K. and J. Cherian (1995), “Role of Product Entry and Exit on the Attraction Effects,” *Marketing Letters*, 29: 45-51.

Slaughter, J., (2007), “Effects of Two Selection Batteries on Decoy Effects in Job Finalist Choice,” *Journal of Applied Social Psychology*, 37, 76-90.

Slaughter, J., E. Sinar, and S. Highhouse (1999), “Decoy Effects and Attribute Level Inferences,” *Journal of Applied Psychology*, 84, 823-828.

Spiegler, R. (2012), “Competitive Framing,” mimeo, UCL.

Tversky, A. and D. Kahneman (1991), “Loss Aversion in Riskless Choice: A Reference-Dependent Model,” *Quarterly Journal of Economics*, 106, 1039–1061.

Wedell, N. (1991), “Distinguishing Among Models of Contextually Induced Preference Reversals,” *Journal of Experimental Psychology*, 17: 767-778.