Reason-Based Choice: A Bargaining Rationale for the Attraction and Compromise Effects*

Geoffroy de Clippel and Kfir Eliaz†

June 2009

Abstract

Among the most important and robust violations of rationality are the attraction and the compromise effects. The compromise effect refers to the tendency of individuals to choose an intermediate option in a choice set, while the attraction effect refers to the tendency to choose an option that dominates some other options in the choice set. This paper argues that both effects may result from an individual's attempt to overcome the difficulty of making a choice in the absence of a single criterion for ranking the options. Moreover, we propose to view the resolution of this choice problem as a cooperative solution to an *intra*-personal bargaining problem among different selves of an individual, where each self represents a different criterion for choosing. We first identify a set of properties that characterize those choice correspondences that coincide with our bargaining solution, for some pair of preference relations. Second, we provide a revealed-preference foundation to our bargaining solution and characterize the extent to which these two preference relations can be uniquely identified.

Alternatively, our analysis may be reinterpreted as a study of (*inter*-personal) bilateral bargaining over a finite set of options. In that case, our results provide a new characterization, as well as testable implications, of an ordinal bargaining solution that has been previously discussed in the literature under the various names of fallback bargaining, unanimity compromise, Rawlsian arbitration rule and Kant-Rawls social compromise.

*We thank Bart Lipman, Mihai Manea and Roberto Serrano for helpful comments. We also thank Jim Campbell for his research assistance.

†Brown University, Department of Economics, Providence, Rhode Island - declippel@brown.edu, kfir_eliaz@brown.edu.

1. INTRODUCTION

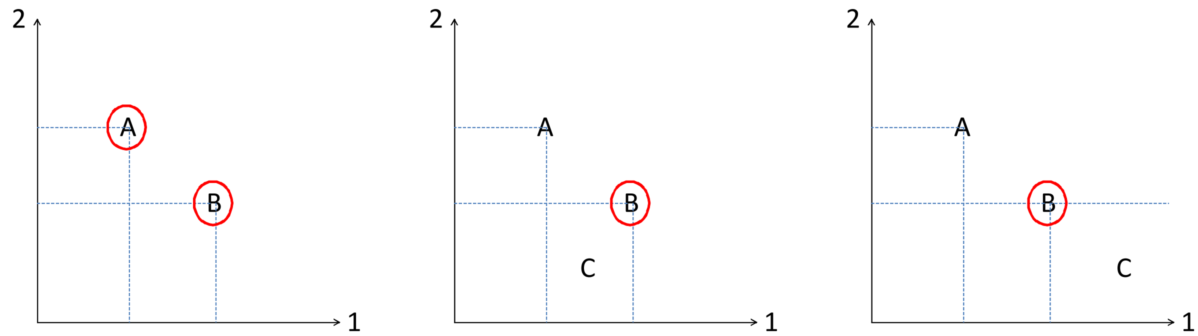
Many of the decision problems we face are complicated by the fact that there is no single dimension or criterion for evaluating the available alternatives. For example, when searching for an apartment or a house, the ranking of the available options may be very different depending on whether the criterion we use is price, size, proximity to work or quality of schools. Similarly, when choosing a car, there are several different criteria or dimensions that one may use such as price, safety, gas efficiency, size, color or esthetics. Also, in deciding between academic job offers there is no one obvious criterion to use as one may consider the ranking of the department, the number of faculty members in one's field, the financial terms, the location, etc.. Often there can be many different dimensions or criteria that one may use, making it difficult, if not impossible, to take all of them into account. This often leads us to focus only on a limited number of dimensions, which we deem most important. However, we are still faced with the difficult task of resolving the trade-off between these dimensions.

If we were fully rational, as is typically assumed in economics, then first, we would be able to take into account all possible dimensions, and second, we would be able to consistenly make the necessary trade-offs across dimensions. However, numerous studies in economics, psychology and marketing provide overwhelming evidence that individuals exhibit systematic departures from rational choice, especially in those situations where there is no obvious single criterion for evaluating the available options. This suggests that individuals often find it difficult to resolve the conflict about how much of one dimension to trade off in favor of another, and hence, they resort to simple heuristics that lead to systematic violations of rationality. Among the most studied and robust violations are the *attraction* and the *compromise* effects.

The attraction effect was first demonstrated by Huber, Payne and Puto (1982), while the compromise effect was introduced by Simonson (1989).[1] The attraction effect refers to the ability of an asymmetrically dominated or relatively inferior alternative, when added to a set, to increase the choice probability of the dominating alternative. The compromise effect refers to the ability of an "extreme" (but not inferior) alternative, when added to a set, to increase the choice probability of an "intermediate" alternative. To illustrate these two effects, consider two options, A and B. Suppose there are two dimensions or criteria for evaluating these options such that A is better than B along the first dimension while

[1] These studies have sprung a whole literature devoted to replicating and extending these effects to various decision problems, including real, monetary choices. For references see Shafir, Simonson and Tversky (1993), Kivetz, Netzer and Srinivasan (2004) and Ariely (2008).

B is better than A along the second dimension (see Figure 1). For example, suppose A and B are two equally priced apartments, but one is closer to work while the other has better schools.



The control treatment   An attraction effect   A compromise effect

Figure 1

In a typical experimental study (which uses a between-subjects design), both A and B are chosen - usually in equal proportions - by a control group of subjects. The attraction effect is observed when a third alternative, C, is added to the set such that it is dominated by only one of the other two options (say, B, as in Figure 1). When subjects are asked to choose from $\{A, B, C\}$, the vast majority of them tend to choose B. The compromise effect occurs when C is added such that it is even better than B along the first dimension but worse than it along the second dimension (i.e., according to the first dimension, C is better than B, which is better than A, while the opposite ranking is obtained according to the second dimension). In such a case, most subjects again tend to pick B. These findings may be interpreted as systematic violations of the Weak Axiom of Revealed Preferences (WARP) by considering a choice correspondence that selects both A and B from $\{A, B\}$, but chooses B alone from $\{A, B, C\}$.[2]

The introduction of these two effects has generated a huge literature in marketing aimed at understanding the source of the effects and their implications for positioning, branding and advertising (see Kivetz, Netzer, and Srinivasan (2004)). One important question that arises is whether the two effects may be viewed as just "snapshots" of a more general choice procedure, which may lead to more significant violations of WARP across various decision problems. This paper attempts to address this question by proposing and characterizing a choice procedure that generates both the attraction and the compromise effects. Our choice procedure is motivated by the interpretation of these two effects

[2]More specifically, this is a violation of the β-axiom proposed by Sen (1971).

as instances of "reason-based choice" (see Simonson (1989), Tversky and Shafir (1992) and Shafir, Simonson and Tversky (1993)). According to this interpretation, in the absence of a single criterion for ranking available options (what is often referred to as "choice under conflict"), choices may be explained "in terms of the balance of reasons for and against the various alternatives" (see Shafir, Simonson and Tversky (1993)).[3] To formalize this interpretation, we envision the decision-maker as trying to reach a compromise between conflicting "inner selves", representing the different attributes or dimensions of the available options. We then propose to view the final choice (i.e., the "balancing of reasons for and against") as a *cooperative* solution to a bargaining among the different selves. In the spirit of the literature on dual-selves (e.g., the $\beta - \delta$ models of present bias, Benhabib and Bisin (2004), Eliaz and Spiegler (2006), Fudenberg and Levine (2006)), we focus our analysis on decision problems that give rise to two selves.

We start by considering the two relevant criteria or dimensions, and their associated rankings, as primitives of the model. This allows us to focus attention on how conflict could be resolved in the mind of a decision maker who is subject to both the attraction and compromise effects, while still satisfying some consistency properties. Also, it may be reasonable in some applications to consider that these primitives are indeed known to the modeler. As illustration, one may think for instance of the choice of product with two attributes such as price and quality, price and size, shipping rate and date of arrival, sugar and fat content, etc. Formally, our first model consists of a finite set of options X and a pair of linear orderings on this set, $\succ = (\succ_1, \succ_2)$. Each ordering is interpreted as the (known) preference relation of one of the individual's dual selves. A bargaining problem is defined to be a non-empty subset of options S. For a given preference profile $\succ$, a bargaining solution is a correspondence $C_\succ$ that associates with every bargaining problem S a subset of S.

Which cooperative bargaining solution can capture our dual-self interpretation of reason-based choice? This solution should first of all exhibit properties that capture the attraction and compromise effects. We interpret an attraction effect as the following property (ATT): whenever a set of options is expanded by adding an alternative that is Pareto dominated by some previously chosen element, then only those chosen alternatives that dominate the new alternative are selected from the new set. We view a compromise as an attempt to resolve conflicting preferences over a pair of alternatives by selecting an outcome that is ranked in between the two by both bargainers. A bargaining solution, therefore, exhibits a compromise effect, or what we call the "No Better Compromise"

[3] Note that reasons involving relationships to other alternatives may lead to violations of WARP.

property (NBC), if whenever x and y are chosen from a set, then there cannot be an element in that set that both bargainers rank between x and y.

Our first main result (Theorem 1) establishes the existence of a unique bargaining solution that satisfies the above properties, in addition to a number of other properties that capture a notion of consistency across decision problems, the cooperative nature of the bargaining, immunity to framing and symmetry. To describe this solution, imagine that for every bargaining problem, each bargainer assigns each option a score equal to the number of elements in its lower contour set. Hence, each option is associated with a pair of scores. The bargaining solution selects the options whose minimal score is highest. This solution has been previously discussed in the literature under various names: "Rawlsian arbitration rule" (Sprumont (1993)), "Kant-Rawls Social Compromise" (Hurwicz and Sertel (1997)), "fallback bargaining" (Brams and Kilgour (2001)), as well as "unanimity compromise" (Kibris and Sertel (2007)). An appealing feature of this bargaining solution is that it is purely ordinal and applies to any arbitrary finite set of options (in contrast to the Nash or Kalai-Smordinsky solutions).[4]

Next we consider an environment in which there is no obvious way to rank the options along two dimensions. We interpret our focus on only two dimensions as an assumption that the decision-maker can process only a limited number of dimensions or attributes. Thus, if the options are characterized by a large number of attributes, it may not be clear which two dimensions the decision-maker focuses on. Hence, an outside observer may not be able to infer what rankings the decision-maker uses to evaluate the options. Alternatively, there may be only two salient dimensions or attributes, but it is not obvious how a decision-maker would rank the options along each dimension (consider, for example, attributes such as color, taste, smell). In such an environment the only observations we may have about the decision-maker are the choices he makes (i.e., his choice correspondence). We ask the following question: what are the necessary and sufficient conditions for representing the decision-maker *as if* he has two selves (each characterized by a linear ordering on X), which make a choice according to the fallback bargaining solution?

Our second main result (Theorem 2) identifies these conditions. This result relies on the notions of "revealed Pareto dominance" and "revealed compromises". An option x is revealed to be Pareto superior to y if it is chosen over y in a pairwise comparison. An option y is revealed to be a compromise between x and z if no option in this triplet is revealed to be Pareto superior over another, and y is chosen uniquely from $\{x, y, z\}$. The

[4]Mariotti (1998) proposes an extension of the Nash bargaining solution to finite environments. However the extended solution still uses cardinal information as it is defined over sets of payoff vectors.

necessary and sufficient conditions identified in Theorem 2 include the revealed versions of the relevant properties characterized in Theorem 1, in addition to properties that capture the consistency of the revealed Pareto relation and the consistency of revealed compromises.

Because we need to simultaneously recover *two* preference relations, proving Theorem 2 requires a different approach than the one that is typically used in the choice theoretic literature. The difficulty arises when we observe that both x and y were chosen from $\{x, y\}$ and that both y and z were chosen from $\{y, z\}$. These choices reveal to us that the two selves disagree on the rankings of $\{x, y\}$ and $\{y, z\}$. The challenge we face is to determine whether the self who ranks x above y also ranks y above z.[5] We overcome this difficulty by constructing an induction argument in which the elements of X are added in a particular order. In particular, we partition the set of options into "revealed Pareto layers", and the elements in each Pareto layer are further partitioned into "revealed extreme layers" (where the most extreme layer includes elements that are never revealed to be compromises in that Pareto layer).

This refined induction proves useful, not only in defining the selves' orderings and showing that they are transitive, but also in addressing the question of "identifiability": to what extent can we identify the set of preference profiles that are compatible with the observed choices? Clearly, exchanging the rankings between the two selves does not affect the bargaining solution. Theorem 3 argues that there is a sense in which any further multiplicity is with respect to "irrelevant alternatives". This means that for any given bargaining problem S, we can pin down the pair of preferences over the minimal set of options that Pareto dominate any option outside this set.

So far, we have interpreted our choice procedure as a solution to an *intra*-personal bargaining problem. Alternatively, we may interpret it as a solution to an *inter*-personal bargaining problem where two distinct individuals need to agree on an option. While most of the choice theoretic literature aims to characterize testable implications of models of *individual* decision-making, the same set of tools may be applied to models of *collective* decision-making. Since many collective decisions are achieved through bargaining, it seems important to identify the necessary and sufficient conditions for inferring the bargainers' preferences and for modelling their decisions as an outcome of cooperative bargaining. This paper takes a first step in this direction by studying situations in

[5] Note that this difficulty does not arise in establishing the revealed-preference foundation of *non-cooperative* solution concepts, such as Nash equilibrium (Sprumont (2000)). There, we can isolate the preference relation of each player by fixing the action of the opponent.

which two individuals bargain over some finite, arbitrary set of alternatives. We, therefore, focuses on *ordinal* bargaining solutions on finite domains. Among such solutions, the fallback bargaining solution has received much attention in the literature. Moreoever, this solution has a non-cooperative foundation, which is similar to the real-life bargaining protocol, that the Federal Mediation and Conciliation Service (FMCS) recommends to disputing parties.[6] Theorem 2 and 3 then provide testable implications of the fallback solution and characterize the extent to which the bargainers' preferences may be recovered from the data.

The rest of the paper is organized as follows. The related literature is discussed in the next section. Section 3 defines the basic concepts and notations. This is followed by an axiomatic characterization of the fallback solution for known preferences in Section 4. The revealed-preference analysis of this solution is presented in Section 5. Finally, Section 6 discusses possible extensions and provides some concluding remarks.

2. RELATION TO THE LITERATURE

In relation to the literature, our paper makes the following contributions. First, we propose a *single* model that "explains" both the attraction and the compromise effects and characterize its testable implications. Second, we provide a revealed-preference foundation for a dual-self model in which the selves strive to reach compromise rather than to behave non-cooperatively. Third, our axiomatic characterization also provides a revealed-preference foundation for a cooperative bargaining solution. To better assess these contributions, we discuss below some of the related papers in the literature.

Explaining attraction and compromise

A number of recent papers have proposed formal models that explain either the attraction effect or the compromise effect. However, there is no single model in this literature that generates both effects in a single-person decision problem (such as those encountered in the experiments that document these effects). Ok, Ortoleva and Riella (2008) relax the Weak Axiom of Revealed Preferences to allow for choice behavior that exhibits the attraction effect, but *not the compromise effect.* They propose a reference-dependent choice model in which given a choice problem S, the decision-maker maximizes a real function u

[6] See Section 1404.12, "Selection by Parties and Appointments of Arbitrators" in http://www.fmcs.gov/internet/itemDetail.asp?categoryID=197&itemID=16959

over those options that Pareto dominate a reference point $r(S)$ according to a sequence of utility functions $\mathbf{u}$. This choice procedure may be interpreted as a bargaining problem with a *continuum* of bargainers and a disagreement point $r(S)$, where the solution maximizes a social welfare function (SWF) u over the set of options that are "individually rational". The authors characterize necessary and sufficient conditions on choice data to be consistent with some bargaining model $(r, \mathbf{u}, u)$. One of these conditions, labeled "reference-dependent WARP", rules out the compromise effect.[7]

Kivetz, Netzer and Srinivasan (2004) argue that individuals may exhibit the compromise effect when choosing among multi-attribute options because of the rule they use to aggregate the different subjective values they assign to the attributes. The authors propose several functional forms of aggregation rules that can generate the compromise effect - *but not the attraction effect* - and test the predictions of these functions on experimental data.

Kamenica (2008) argues that in a monopolistic market with enough uninformed but *rational* consumers, there are some conditions that guarantee the existence of equilibria in which the uninformed consumers exhibit the compromise effect, or the attraction effect, with positive probability. While this argument suggests one interpretation of why consumers in a market may exhibit compromise/attraction-like behavior, there are some caveats in adopting this argument as *the* explanation of the attraction and compromise effect. First, Kamenica (2008) studies a signalling game, which like all signalling games has multiple equilibria. The equilibria of interest that the paper identifies are only those that survive what is known as the D1 criterion.[8] Second, there are many instances - such as the numerous experiments that document the compromise and attraction effects - in which individuals consistently exhibit these effects outside the market when they are not engaged in a non-cooperative game against some seller.

Rationalization by multiple rationales

A number of recent papers have proposed to model systematic violations of IIA as the result of a choice procedure that, in contrast to rational choice, takes as input multiple orderings ("rationales") on the set of alternatives. One set of works in this literature is

[7] To see this, recall that the compromise effect means that whenever the choice out of any pair in $\{x, y, z\}$ is the pair itself, then only a single element will be chosen from the triplet. Suppose y is chosen. If the choice correspondence satisfies "reference-dependent WARP" then either x or z act as a "potential reference point" for y, meaning that y must be chosen uniquely from $\{x, y\}$ or from $\{y, z\}$, a contradiction.

[8] In addition, the existence of these equilibria require assumptions on both the seller and the consumers.

not concerned with deriving testable implications and focuses on a different set of questions than we do. Kalai, Rubinstein and Spiegler (2002) ask what is the *minimal* number of preference relations on a set of elements, X, such that the single choice from any subset $A \subseteq X$ is the maximal element in A according to at least one of these relations. Ambrus and Rosen (2008) are concerned with the minimal number of utility functions that are needed to explain a choice function as the maximization of a cardinal SWF that aggregates these utilities. Green and Hojman (2007) propose a welfare criterion for evaluating irrational choices, by modeling these choices as reflecting a weighted aggregation of all possible strict orderings on the set of options.

Another set of papers in this literature does attempt to provide a revealed-preference characterization to decision heuristics that use multiple rationales. Manzini and Mariotti (2007) study a "shortlisting" procedure according to which a decision-maker sequentially applies two binary relations, R_1 and R_2, such that the ultimate choice from a set S is the R_2-maximal element from among the R_1-maximal elements in S (see also Manzini and Mariotti (2008)). Cherepanov, Feddersen and Sandroni (2008) propose a model in which a decision-maker is characterized by two primitives: a single, complete preference relation and a set of asymmetric, binary (possibly incomplete) relations. Given a set of options A, the decision-maker picks his most preferred option from among the elements in A that are maximal according to each of the rationales. These papers all focus on single-valued choice rules that violate IIA. In contrast, one of the key properties of our choice rule, RA, reduces to IIA when the choice rule is required to be single-valued.

Testable implications of collective decision-making

Finally, our paper is related to a small but growing literature that aims to provide testable implications for models of collective decision-making. Among those papers that employ a revealed-preference methodology, the most closely related are Sprumont (2000) and Eliaz, Richter and Rubinstein (2009). The former provides a choice theoretic characterization of Nash equilibrium and the Pareto correspondence, while the latter characterizes the choice correspondence that selects the top element(s) of two preference orderings.

A number of other papers explore similar questions but without employing a revealed-preference methodology. Chiappori (1988) characterizes the conditions under which it is possible to recover the preferences and decision process of two individuals, who consume leisure and some Hicksian composite good, from observations on their labor supply functions. Chiappori and Ekeland (2006) extend this analysis and characterize the necessary and sufficient conditions for recovering the individual preferences of a group of individuals from observations on their aggregate consumption and the common budget constraint that they face. Chiappori and Donni (2005) analyze the testable implications of the Nash bargaining solution in an environment where two individuals need to agree on the allocation of a pie among themselves and where disagreement leads each to receive some reservation payment. In a similar vein, Chambers and Echenique (2008) study the testable implications of the standard model of two-sided markets with transfers and characterize the sets of matchings that may be generated by the model.

3. DEFINITIONS

X will denote the finite set of all potential options. A *bargaining problem* is a subset of X. A *bargaining solution* C associates to each bargaining problem S a nonempty subset $C(S)$ of S. A (strict) *linear ordering* on X is a relation defined on $X \times X$ that is complete, transitive, and anti-reflexive. The set of all possible linear orderings is denoted $L(X)$. A *preference-based bargaining solution* is a function[9] C that associates a bargaining solution $C_\succ$ to each pair $\succ = (\succ_1, \succ_2)$ of linear orderings on X.

Let $\succ = (\succ_1, \succ_2) \in L(X) \times L(X)$, and let S be a bargaining problem. The *score* of x in S along dimension i ($i = 1$ or 2) is the number of feasible options that are (strictly) worse than x for $\succ_i$:

$$s_i(x, S, \succ) = |\{y \in S | x \succ_i y\}|.$$

The *fallback bargaining solution* $C^f_\succ$ associated to $\succ$ associates to each bargaining problem S the set of options in S that maximize the minimum (over $i = 1, 2$) of the scores:

$$C^f_\succ(S) = \arg\max_{x \in S} \min_{i=1,2} s_i(x, S, \succ).$$

The resulting preference-based bargaining solution will be denoted by C^f.

As pointed out in the introduction, the fallback bargaining solution has already appeared under various names in the literature on interpersonal bargaining (Sprumont (1993), Hurwicz and Sertel (1997), Brams and Kilgour (2001), Kibris and Sertel (2007)). The terminology of "fallback bargaining" is taken from Brams and Kilgour (2001), where they offer a nice reinterpretation of the solution. For each bargaining problem S, and

[9] For notational convenience, we use the same letter, C, to denote both a bargaining solution and a preference-based bargaining solution. The context will always make it clear what the right meaning is.

each integer between 1 and $|S|$, let $E_i(S,k)$ be the set of k best options in S according to i's preferences. Let k^* be the smallest k such that $E_1(S,k) \cap E_2(S,k) \neq \emptyset$. Then $\Sigma^f_{\succ}(S) = E_1(S,k^*) \cap E_2(S,k^*)$.

In other words, if both criteria agree on what the best option is, then it is the solution. Otherwise, the decision-maker looks for option(s) that would be ranked either top or second-best by both criteria. If no option satisfies this property, then the decision-maker iterates the procedure by allowing for third-best alternatives, and so on so forth. This simple algorithm for deriving the elements in the solution illustrates the appeal of the fallback solution as a descriptive model of multi-criteria decision making.

Figure 2 illustrates how the fallback solution generates the attraction and compromise effects.

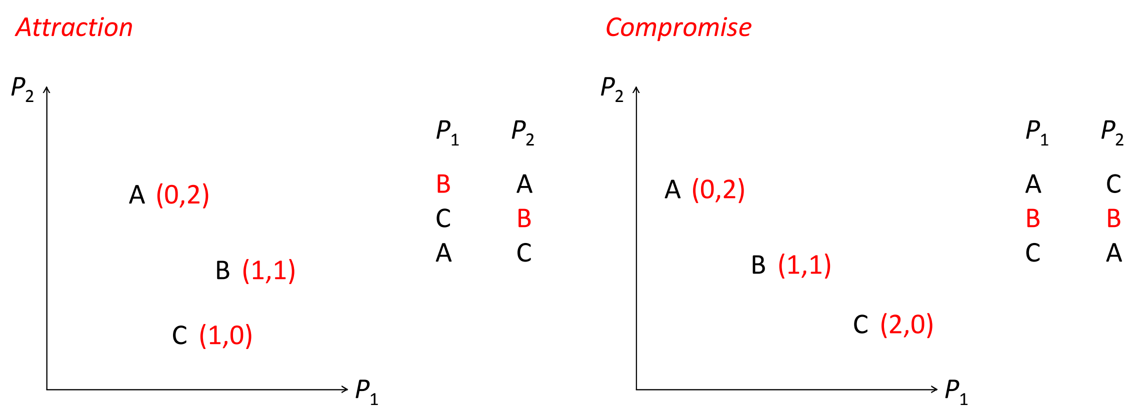


Figure 2

In both cases, both A and B would get a minimal score of 0 if C was not available. Adding C changes the scores, and B now gets the largest minimal score in both cases. It thus becomes selected uniquely by the fallback solution.

It is also interesting to note that in the spirit of the Nash program, fallback bargaining has a non-cooperative foundation. The two bargainers alternate in proposing one of the available options as a possible agreement. If the responder accepts, the game ends and the proposed option is adopted. If the responder rejects, the proposed option is removed from the set, and the responder now proposes one of the remaining options. The game continues until either an agreement is reached or there remains only a single option, which is then adopted. Anbarci (1993, 2006) shows that the unique subgame-perfect equilibrium of this game is an element in $C^f_{\succ}(S)$.[10]

[10] Interestingly, Anbarci (1993) also shows that the subgame-perfect equilibrium outcome also converges

The fallback solution applies an egalitarian criterion to a canonical representation of the ordinal preferences. It is interesting to think about applying a utilitarian criterion instead. The resulting solution would then be the analogue in our context to the rule that Borda defined in the eighteenth century to select representatives. However, this solution has two important shortcomings. First, it selects all the elements of S whenever they are all Pareto optimal for the pair of orderings $(\succ_1, \succ_2)$, and hence does not capture the compromise effect. Second, in contrast to the fallback solution, the Borda rule is not robust to common monotonic transformations of the bargainers' ordinal preferences in the sense that it is sensitive to how the score of an option changes as it moves up in the ranking (cf. "scoring rules").

4. PREFERENCE-BASED AXIOMATIC CHARACTERIZATION

The aim of this section is to establish that fallback bargaining is the unique preference-based bargaining solution that captures a number of desiderata. First, it should exhibit properties that capture a plausible interpretation of attraction and compromise. Second, we should be able to interpret the solution as a "procedurally-rational" heuristic. Thus, the solution should exhibit some form of consistency across decision/bargaining problems. Third, the solution should capture our idea that the bargaining among the selves is in some sense "cooperative". Finally, we wish to interpret all the options selected by the solution (i.e., any "agreement" reached by the two selves) as being on "equal footing" in terms of their desirability and robusteness to small changes in the bargaining prolem.

We focus our attention on a class of preference-based bargaining solutions, which satisfy some basic properties from axiomatic bargaining and social choice. This would allow us to meaningfully interpret the correspondence $C_{\succ}$ as a bargaining solution. Specifically, a preference-based bargaining solution is *regular* if

1. it is *neutral* in the sense of not having an a priori bias in favor or against some elements of X. Let $g : X \to X$ be an isomorphism. Then $C_{g(\succ)}(g(S)) = C_{\succ}(S)$, where $g(S) = \{g(x)|x \in S\}$ and $g(\succ) \in L(X) \times L(X)$ is such that $x g_i(\succ) y$ if and only if $g^{-1}(x) \succ_i g^{-1}(y)$, for all $x, y \in X$ and both $i \in \{1, 2\}$.

2. it treats both orderings with equal relevance: $C_{(\succ_2,\succ_1)}(S) = C_{(\succ_1,\succ_2)}(S)$.

to the Area Monotonic Solution if the alternatives are uniformly distributed over the bargaining set, and as the number of alternatives tends to infinity.

3. options are selected using only the parts of the two orderings that are relevant to the problem. If $\succ'$ is an alternative pair of linear orderings (defined on X) that coincide with $\succ$ on $S \times S$, then $C_{\succ'}(S) = C_{\succ}(S)$.[11]

It is certainly of interest to investigate how our theory would adapt if one eliminates some or all of these properties. Dropping neutrality would allow to accomodate some framing effects, where the label of the available options may influence the choice (e.g. options are presented in a list, or are offered by trademarks with varying impact, etc.). Dropping the second property would add the possibility of having one of the two criteria as being more relevant than the other (e.g. caring more about the size of the car than its color). Dropping the third property would allow to consider choice procedures where the decision maker is influenced by options he aspires to, but cannot afford. Yet, we believe that one must first understand the attraction and compromise effects in their purest form, in absence of all these additional features. The regularity property thus defines a benchmark which can be used to build more elaborate theories.

The following axioms are imposed on a regular preference-based bargaining solution C, and will be assumed to be valid for each $\succ \in L(X) \times L(X)$, and each $S \subseteq X$.

Attraction (ATT) - *Let* $x \in X \setminus S$ *be such that* $y \succ x$, *for some* $y \in C_{\succ}(S)$. *Then* $C_{\succ}(S \cup \{x\}) = \{y \in C_{\succ}(S) | y \succ x\}$.[12]

ATT formalizes the idea that adding a dominated alternative reinforces the appeal of an option to the decision maker. This property is best understood by decomposing it into two parts. First, whenever option x is added to a set S, it seems reasonable to expect that the set of options that were previously selected, and which dominate x, should continue to be chosen, i.e., $\{y \in C_{\succ}(S) | y \succ x\} \subseteq C_{\succ}(S \cup \{x\})$. We view the attraction effect as the converse inclusion, $C_{\succ}(S \cup \{x\}) \subseteq \{y \in C_{\succ}(S) | y \succ x\}$, i.e., when choosing from the new set, one's attention is drawn to the previously selected options that dominate x. Thus, the solution to the enlarged problem obtained by adding x as a feasible option should be the intersection of the solution to the original problem with those options that Pareto dominate x whenever that set is nonempty.[13]

[11] Similar properties have been used repeatedly in the classical theories of bargaining and social choice (first mentioned explicitly in Karni and Schmeidler (1975)).

[12] $\succ$ refers to the Pareto relation (incomplete ordering on $X \times X$) when comparing options, i.e. $x \succ y$ means $x \succ_1 y$ and $x \succ_2 y$. On the other hand, the symbol $\succ$ in $C_{\succ}$ refers to the pair $(\succ_1, \succ_2)$ of linear orderings on X. We do not introduce different symbols because the right meaning is always obvious when used in context.

[13] One could argue that the added element x may also increase the appeal of some options that were chosen in S, but do not Pareto dominate x, because of the compromise effect. Thus, one may feel

No Better Compromise (NBC) - *If both x and y belong to $C_{\succ}(S)$, then there does not exist $z \in S$ and $i \in \{1, 2\}$ such that $x \succ_i z \succ_i y$ and $y \succ_{-i} z \succ_{-i} x$.*

NBC captures the idea that the bargainers are trying to reach a compromise. If two bargainers were not able to agree on a single option - so that both x and y are identified as possible agreements - then it must be that there was no option z available that could serve as a compromise between x and y. By this we mean that there was no alternative z that falls "in between" x and y, in that it is better than x along the dimension where it is worse that y, and better than y along the dimension where it is worse than x.[14]

Removing an Alternative (RA) - *If $C_{\succ}(S) \neq \{x\}$, then $C_{\succ}(S \setminus \{x\}) \cap C_{\succ}(S) \neq \emptyset$.*

RA captures the sense in which the bargaining solution may be interpreted as a procedurally-rational heuristic. Since both ATT and NBC are typically incompatible with WARP, we propose a weaker consistency property. If an option (that is not the unique choice of the decision maker) is dropped, then at least one of the options that were chosen in the original problem belongs to the solution of the reduced problem. Observe that RA is equivalent to IIA if the bargaining solution is single-valued, as RA can be applied iteratively if one needs to eliminate multiple irrelevant alternatives. Yet, moving to correspondences, the slight difference between the two properties when eliminating a single alternative can lead to major differences in terms of choices. In addition, RA also expresses some form of continuity in our discrete setting. Indeed, making a small change in the set of available options (i.e. dropping only one alternative) should not modify too much the set of selected elements (i.e. nonempty intersection) whenever this set is not a singleton.

Efficiency (EFF) - *If $x \in C_{\succ}(S)$, then there does not exist $y \in S$ such that $y \succ x$.*

EFF captures the cooperative nature of the bargaining. It is also a standard property in axiomatic bargaining and social choice.

---

that imposing ATT is unduly restrictive, as it presumes that the attraction effect is more relevant than the compromise effects in those configurations. It turns out that both our characterization results (see Theorems 1 and 2 below) remain valid when ATT is weakened so as to apply only when those elements that were selected in S but do not Pareto dominate x do not fall in between x and another alternative of S.

[14] Note that since we are using only ordinal information, any element z such that $x \succ_i z \succ_i y$ and $y \succ_j z \succ_j x$ is interpreted a "compromise", regardless of how it is ranked relative to other elements that are ranked in between x and y. One may question this interpretation if, for example, $x \succ_i z \succ_i w \succ_i v \succ_i y$ and $y \succ_i w \succ_i v \succ_i z \succ_j x$. In this case it may seem less reasonable to consider z a compromise between x and y, since in some sense, it is "closer" to x than to y. We return to this point in the concluding section, where we discuss possible extensions.

Symmetry (SYM) - *If $x, y \in C_{\succ}(S)$ and there exists $z \in S \setminus \{x, y\}$ such that $x \notin C_{\succ}(S \setminus \{z\})$, then there exists $z' \in S \setminus \{x, y\}$ such that $y \notin C_{\succ}(S \setminus \{z'\})$.*

SYM formalizes a sense in which all the options selected by the solution are of equal "status". Suppose x and y are both in the solution. Imagine that one of the bargainers makes the following argument against the inclusion of x: "x is not selected when the option z is removed from the table; but since we did not choose z, we may consider it off the table, hence, we should not select x". Such an argument would not be convincing if the other bargainer could counter by observing that a similar claim can be made against y : if we remove z', which was not chosen, then y will no longer be selected. Observe that SYM is vacuous if the choice method is rational, but it does place a nontrivial restriction on irrational procedures. However, this property is satisfied by some well-known social welfare functions such as the Borda rule mentioned above.

Our main result in this section relies on the following inductive characterization of the fallback bargaining solution (we relegate its proof to the Appendix).

Lemma 1 *Let $\succ \in L(X) \times L(X)$, and let S be a bargaining problem with at least four elements. Then,*

1. *$C^f_{\succ}(S) = \{x\}$ if and only if*
   - *(a) $x \in C^f_{\succ}(S \setminus \{w\})$, for each $w \in S \setminus \{x\}$, and*
   - *(b) for each $y \in S \setminus \{x\}$, there exists $w \in S \setminus \{y\}$ such that $y \notin C^f_{\succ}(S \setminus \{w\})$.*
2. *$C^f_{\succ}(S) = \{x, y\}$ if and only if*
   - *(a) $C^f_{\succ}(S \setminus \{w\}) \subseteq \{x, y\}$, for each $w \in S$, and*
   - *(b) there exists $w \in S \setminus \{x, y\}$ such that $C^f_{\succ}(S \setminus \{w\}) = \{x\}$ if and only if there exists $w' \in S \setminus \{x, y\}$ such that $C^f_{\succ}(S \setminus \{w'\}) = \{y\}$.*

In addition, if $C^f_{\succ}(S) = \{x, y\}$, and $C^f_{\succ}(S \setminus \{w\}) = \{x, y\}$, for all $w \in S \setminus \{x, y\}$, then $x \succ w$ and $y \succ w$, for all $w \in S \setminus \{x, y\}$. Also, if $C^f_{\succ}(S) = \{x\}$, $C^f_{\succ}(S \setminus \{y\}) = \{x, z\}$, and $C^f_{\succ}(S \setminus \{z\}) = \{x, y\}$, then there exists $i \in \{1, 2\}$ such that $y \succ_i x \succ_i z$ and $z \succ_{-i} x \succ_{-i} y$.

Theorem 1 *C^f is the only regular preference-based bargaining solution that satisfies EFF, ATT, NBC, RA, and SYM.*

Proof: We first check that C^f satisfies the axioms. EFF and regularity follow immediately from the definition. RA and SYM follow from Lemma 1. As for ATT, observe that $\min_{i=1,2} s_i(y, S\cup\{x\}, \succ) = \min_{i=1,2} s_i(y, S, \succ) + 1$, for each $y \in C^f_\succ(S)$ such that $y \succ x$, while the minimal score of any other option cannot increase by more than one point. Hence any such y must belong to $C^f_\succ(S\cup\{x\})$, and any option that was not selected for S does not belong to $C^f_\succ(S\cup\{x\})$. Now we only have to show that $z \notin C^f_\succ(S\cup\{x\})$ when $C^f_\succ(S) = \{y, z\}$, $y \succ x$ and $z \not\succ x$. To fix the notation, suppose that $\arg\min_{i=1,2} s_i(y, S, \succ) = 1$ and $\arg\min_{i=1,2} s_i(z, S, \succ) = 2$. Hence $z \succ_1 y$, and transitivity implies that $z \succ_1 x$. In turn, this implies that $x \succ_2 z$. The minimal score of z thus remains constant when adding x, and $z \notin C^f_\succ(S\cup\{x\})$. Finally for NBC, suppose that $C^f_\succ(S) = \{x, y\}$ and that there exists $z \in S$ such that $x \succ_i z \succ_i y$ and $y \succ_{-i} z \succ_{-i} x$. Hence it must be that the minimal score for y is reached along dimension i, and it is equal to the minimal score of x that is reached along dimension $-i$. On the other hand, z scores at least one additional point than y (resp. z) along dimension i (resp. $-i$). Hence a contradiction with the fact that x and y have the largest minimal score among all the elements of S.

We now move to the more difficult part of the proof, showing the necessary condition. Let thus C be a preference-based bargaining solution that satisfies the eight axioms. We prove that $C = C^f$ in two main steps.

Step 1 *Let C be a preference-based bargaining solution that satisfies ATT, NBC, RA, EFF, and SYM, and let $\succ \in L(X) \times L(X)$. If $C_\succ(T) = C^f_\succ(T)$, for all $T \subseteq X$ with two or three elements, then $C_\succ(S) = C^f_\succ(S)$, for all $S \subseteq X$.*

We prove that $C_\succ(S) = C^f_\succ(S)$, for all $S \subseteq X$, by induction on the number of elements in S. By assumption, the result is true when $|S| = 2$ or 3. We assume now that the result holds for any subset of X with at most $s-1$ elements, and we choose a set S with exactly s elements ($s \geq 4$). We have to prove that $C_\succ(S) = C^f_\succ(S)$.

First we observe that $C_\succ(S)$ has at most two elements. Suppose on the contrary that $x, y, z \in C_\succ(S)$. EFF implies that there is no Pareto comparison between any pair of elements in $\{x, y, z\}$. Hence one of these three options must fall "in between" the other two, leading to a contradiction with NBC.

Suppose now that $C^f_\succ(S) = \{x, y\}$, for some $x, y \in S$. Lemma 1 and the induction hypothesis imply that $C_\succ(S\setminus\{w\}) = C^f_\succ(S\setminus\{w\}) \subseteq \{x, y\}$, for each $w \in S$. Notice that $C_\succ(S)$ cannot include an element different from x and y. Indeed, $\#C(S) \leq 2$ would then imply that $C_\succ(S) = \{z\}$, $\{x, z\}$, $\{y, z\}$, or $\{z, z'\}$, for some $z, z' \in S\setminus\{x, y\}$, and RA would lead to a contradiction with $C_\succ(S\setminus\{w\})\subseteq\{x,y\}$, for all $w\in S$. So we'll be done after proving that $C_\succ(S)$ is equal to neither $\{x\}$, nor $\{y\}$. Suppose on the contrary that $C_\succ(S)=\{x\}$ (a similar reasoning applies for y). RA implies that $x\in C_\succ(S\setminus\{w\})$, for all $w\in S\setminus\{x\}$. If there exists $w\in S\setminus\{y\}$ such that $y\notin C_\succ(S\setminus\{w\})$, then $C^f_\succ(S\setminus\{w\})=C_\succ(S\setminus\{w\})=\{x\}$. Lemma 1 and the induction hypothesis imply that there exists $w'\in S\setminus\{x,y\}$ such that $C_\succ(S\setminus\{w'\})=C^f_\succ(S\setminus\{w'\})=\{y\}$, a contradiction with the fact that $x\in C_\succ(S\setminus\{w'\})$. We must conclude that $C^f_\succ(S\setminus\{w\})=C_\succ(S\setminus\{w\})=\{x,y\}$, for all $w\in S\setminus\{x,y\}$. The penultimate statement of Lemma 1 implies that $x\succ w$ and $y\succ w$, for all $w\in S\setminus\{x,y\}$, or $C_\succ(\{x,w\})=\{x\}$ and $C_\succ(\{y,w\})=\{y\}$ since $C_\succ=C^f_\succ$ on pairs. We also have $C_\succ(\{x,y\})=C^f_\succ(\{x,y\})=\{x,y\}$, and applying ATT iteratively (adding elements of $S\setminus\{x,y\}$ one at a time), we conclude that $C_\succ(S)=\{x,y\}$, contradicting the original assumption that $C_\succ(S)=\{x\}$.

To conclude the proof of Step 1, suppose that $C^f_\succ(S)=\{x\}$, for some $x\in S$. If $C_\succ(S)=\{y\}$, for some $y\neq x$, then $y\in C_\succ(S\setminus\{w\})$, for all $w\in S\setminus\{y\}$, by RA. This leads to a contradiction with Lemma 1, since there must exist $w\in S\setminus\{y\}$ such that $y\notin C^f_\succ(S\setminus\{w\})=C_\succ(S\setminus\{w\})$. It is also impossible to have $C_\succ(S)=\{y,z\}$, for some y,z different from x. Indeed, RA aplied to both $C_\succ$ and $C^f_\succ$ would then imply that $C^f_\succ(\{S\setminus\{y\})=\{x,z\}$, and $C^f_\succ(\{S\setminus\{z\})=\{x,y\}$. The last statement of Lemma 1 implies that there exists $i\in\{1,2\}$ such that $y\succ_i x\succ_i z$ and $z\succ_{-i} x\succ_{-i} y$, a contradiction with NBC. Suppose now that $C_\succ(S)=\{x,y\}$, for some y different from x. Lemma 1 implies that there exists $w\in S\setminus\{y\}$ such that $y\notin C^f_\succ(S\setminus\{w\})=C_\succ(S\setminus\{w\})$. SYM implies that there exists $w'\in S\setminus\{x\}$ such that $x\notin C_\succ(S\setminus\{w'\})=C^f_\succ(S\setminus\{w'\})$, which is impossible. Hence $C_\succ(S)=\{x\}$, as desired. This concludes the proof of Step 1.

Step 2 *Let C be a preference-based bargaining solution that satisfies ATT, NBC, RA, EFF, NEUT, EX, and IPUA. Then $C_\succ(T)=C^f_\succ(T)$ for all $T\subseteq X$ with two or three elements, and all $\succ\in L(X)\times L(X)$.*

Let $\succ\in L(X)\times L(X)$. Suppose first $T=\{x,y\}$. If $x\succ y$, then $C^f_\succ(T)=\{x\}$. Since $C_\succ(\{x\})=\{x\}$, ATT implies that $C_\succ(T)=\{x\}$ as well, as desired. A similar reasoning applies if $y\succ x$. If $x\succ_1 y$ and $y\succ_2 x$, then $C^f_\succ(T)=\{x,y\}$. Suppose, on the other hand, that $C_\succ(T)=\{x\}$. Let $g:X\to X$ be the isomorphism defined by $g(x)=y$, $g(y)=x$, and $g(z)=z$, for all $z\in X\setminus\{x,y\}$. NEUT implies that $C_{g(\succ)}(g(T))=\{y\}$. Notice though that $g(T)=T$, and $g(\succ)$ equals $(\succ_2,\succ_1)$ when restricted to T. EX and IPUA then imply that $C_\succ(T)=\{y\}$, a contradiction. Similarly, $C_\succ(T)=\{y\}$ would lead to a contradiction, and we conclude that $C_\succ(T) = \{x, y\}$, as desired. A similar reasoning applies if $y \succ_1 x$ and $x \succ_2 y$.

Let now $T = \{x, y, z\}$. If one of the elements, let's say x Pareto dominates the other two, then $C^f_\succ(T) = \{x\} = C_\succ(T)$, by EFF. If two elements, let's say x and y are not Pareto dominates, but both Pareto dominate the third one, then $C^f_\succ(T) = \{x, y\}$. The previous paragraph implies that $C_\succ(\{x, y\}) = \{x, y\}$, and ATT implies that $C_\succ(T) = \{x, y\}$, as desired. If two pairs of elements are not Pareto comparable, let's say (x, y) and (x, z), but the third one is, let's say $y \succ z$, then $C^f_\succ(T) = \{y\}$. The previous paragraph implies that $C_\succ(\{x, y\}) = \{x, y\}$, $C_\succ(\{x, z\}) = \{x, z\}$, and $C_\succ(\{y, z\}) = \{y\}$. ATT implies that $C_\succ(T) = \{y\}$ as well, as desired. Remains the last case, where there is no Pareto comparison out of any pair in T, let's say $x \succ_1 y \succ_1 z$ and $z \succ_2 y \succ_2 x$. Then $C^f_\succ(T) = \{y\}$. We already proved in Step 1 that $C_\succ(T)$ contains at most two elements. If cannot be $\{x, z\}$, because of NBC. If that $C_\succ(T) = \{x, y\}$, then consider the isomorphism $g : X \to X$ defined by $g(x) = z$, $g(z) = x$, and $g(\xi) = \xi$, for all $\xi \in X \setminus \{x, z\}$. NEUT implies that $C_{g(\succ)}(g(T)) = \{y, z\}$. Notice though that $g(T) = T$, and $g(\succ)$ equals $(\succ_2, \succ_1)$ when restricted to T. EX and IPUA then imply that $C_\succ(T) = \{y, z\}$, a contradiction. A similar argument shows that it is impossible to have $C_\succ(T) = \{y, z\}$, $\{x\}$, or $\{z\}$. Hence $C_\succ(T) = \{y\}$. This concludes the proof of Step 2, and hence the proof of the theorem. ■

Both Sprumont (1993) and Kibris and Sertel (2007) have already provided some axiomatic characterizations of the fallback solution in an *inter*-personal bargaining context. The main axioms in these previous papers restrict the behavior of the solution across problems that differ in the bargainers' *preferences*. These axioms thus become meaningless when preferences are unknown. In contrast, our axioms impose restrictions on how the composition of the solution varies across different bargaining problems. As we show in the next section, these axioms can be adapted to a setting in which preferences are not observable.

We now prove the independence of the axioms appearing in Theorem 1.

EFF: Consider a set X with four elements, let's say $X = \{a, b, c, d\}$, and let C be the preference-based bargaining solution that coincides with C^f, except as follows: $C_\succ(X) = C^f_\succ(X) \cup \{x\}$, for all $x \in X$ and all $\succ \in L(X) \times L(X)$ such that $\succ_1$ and $\succ_2$ are completely opposite on $X \setminus \{x\}$, and x is Pareto dominated by either the $\succ_1$-optimal element or the $\succ_2$-optimal element, but by no other element of X. For instance, $C_{\succ^*}(X) = \{a, b, c\}$, while $C^f_{\succ^*}(X) = \{b, c\}$, when $b \succ^*_1 a \succ^*_1 c \succ^*_1 d$ and $d \succ^*_2 c \succ^*_2 b \succ^*_2 a$. The modification thus amounts to add some options to the fallback solution in some cases, and will satisfy

RA *a fortiori.* By construction, C is regular, but violates EFF. ATT does not apply in those cases where C is different from C^f (because the Pareto dominated option falls below an option that is not chosen in the triplet obtained by deleting that Pareto dominated option), and hence C satisfies it (since C^f does). It is straightfoward to check NBC. Finally, SYM is satisfied because C^f satisfies it, and a Pareto dominated option is never selected out of any triplet.

ATT: Consider the fallback solution applied only to the set of Pareto efficient alternatives, $C_{\succ}(S) = C^f_{\succ}[EFF_{\succ}(S)]$, where

$$EFF_{\succ}(S) = \{x \in S \mid \textit{for all } y \in S,\ x \succ_i y \textit{ for some } i \in \{1, 2\}\} \tag{1}$$

Note that the fallback solution is applied here to a subset of options, whose score is unaffected by dominated elements. Hence, $C_{\succ}$ violates ATT. It is straightforward to verify that $C_{\succ}$ is regular and satisfies NBC, RA, EFF, EX and IPUA. To see that it also satisfies SYM, suppose $x, y \in C_{\succ}(S)$ but $x \notin C_{\succ}(S\backslash\{z\})$ for some $z \in S\backslash\{x, y\}$. Then $z \in EFF_{\succ}(S)$. Let $T \equiv EFF_{\succ}(S)$, then $x, y \in C^f_{\succ}(T)$ but $x \notin C^f_{\succ}(T\backslash\{z\})$ for some $z \in T\backslash\{x, y\}$. Then by SYM, $y \notin C^f_{\succ}(T\backslash\{z'\})$ for some $z' \in T\backslash\{x, y\}$, which implies that $y \notin C_{\succ}(S\backslash\{z'\})$.

NBC: Consider the analogue of the Borda rule in our setting:

$$C_{\succ}(S) = \arg\max_{x \in S}[s_1(x, S, \succ) + s_2(x, S, \succ)],$$

for each subset S of X. It is straightforward to check that this defines a regular preference-based bargaining solution that satisfies EFF and ATT. It violates NBC. For instance, it does not refine the set of Pareto efficient options when the two preferences are strict opposite to each others. It remains to show that the solution satisfies both RA and SYM. Since it satisfies EFF, the sum of the scores must decrease by at least one point for each option that is chosen, when removing x from the original problem S. Any element of $C_{\succ}(S)$ such that the sum of the scores decreases by exactly one point when removing x clearly belongs to $C_{\succ}(S \setminus \{x\})$. Hence we must consider the case where the sum of the scores decreases by two points, for each element of $C_{\succ}(S)$. This implies that x is Pareto dominated by some elements of S, and the set of Pareto efficient options remains unchanged when removing x. The sum of the scores of any element of the Pareto frontier decreases by at least one point when removing x, and hence $C_{\succ}(S) \subseteq C_{\succ}(S\backslash\{x\})$, and we are done proving RA. For SYM, suppose on the contrary that one can find $x, y \in C_{\succ}(S)$

and $z \in S \setminus \{x, y\}$ such that $x \notin C_\succ(S \setminus \{z\})$ and $y \in C_\succ(S \setminus \{z\})$. Both x and y being Pareto efficient in S, it must be that the sum of the scores decreases by at least one point for both of them when removing z. Since y remains chosen, but not x, it must be that the sum of the scores of x decreases by two points while the sum of the scores of y decreases by exactly one point. In other words, x Pareto dominates z, but y does not Pareto dominates z. It is easy to check that one would get a contradiction with $x, y \in C_\succ(S)$ if there does not exist $z' \in S$ that is Pareto dominated by y, but not by x. For any such z', we'll have $y \notin C_\succ(S \setminus \{z'\})$, and we are done proving SYM.

RA: Let CL^f be the lexicographic refinement of the fallback solution,

$$CL^f_\succ(S) = \{x \in C^f_\succ(S) \mid s_i(x, S, \succ) \geq s_i(y, S, \succ) \; \forall y \in C^f_\succ(S)\},$$

for each $S \subseteq X$, amd each $\succ \in L(X) \times L(X)$. It is easy to check that CL^f inherits the properties of regularity, EFF, ATT, and NBC from C^f. To see that it violates RA, consider $S = \{a, b, c, d\}$ and the preference pair $\succ^*$ that give rise to the following ranking on S: $b \succ_1^* a \succ_1^* c \succ_1^* d$ and $d \succ_2^* c \succ_2^* b \succ_2^* a$. Then $CL^f_{\succ^*}(S) = \{b\}$ while $CL^f_{\succ^*}(X \setminus \{a\}) = \{c\}$. All what remains is to check SYM. Suppose that $x, y \in CL^f_\succ(S)$ and that there exists $z \in S \setminus \{x, y\}$ such that $x \notin CL^f_\succ(S \setminus \{z\})$. This implies that $x, y \in C^f_\succ(S)$. If there exists $z \in S \setminus \{x, y\}$ such that $x \notin C^f_\succ(S \setminus \{z\})$, then there exists $z' \in S \setminus \{x, y\}$ such that $y \notin C^f_\succ(S \setminus \{z'\})$, by SYM. If there exists $z' \in S \setminus \{x, y\}$ such that $y \notin C^f_\succ(S \setminus \{z'\})$, then $y \notin CL^f_\succ(S \setminus \{z'\})$, as desired, since CL^f refines C^f. Hence the last case that could lead to a possible violation of SYM for CL^f is when $x, y \in C^f(S \setminus \{z\})$, for all $z \in S \setminus \{x, y\}$. But we know from Lemma 1 that this configuration of choice for C^f is possible only if $x \succ z$ and $y \succ z$, for all $z \in S \setminus \{x, y\}$. In such cases, it is impossible to have $x \notin CL^f_\succ(S \setminus \{z\})$, and we are done proving SYM.

SYM: Consider a set X with five elements, let's say $X = \{a, b, c, d, e\}$, and let C be the preference-based bargaining solution that coincides with C^f, except as follows: $C_\succ(X) = C^f_\succ(X \setminus \{x\})$, for all $x \in X$ and all $\succ \in L(X) \times L(X)$ such that $\succ_1$ and $\succ_2$ are completely opposite on $X \setminus \{x\}$, and x is Pareto dominated by either the $\succ_1$-optimal element or the $\succ_2$-optimal element, but by no other element of X. For instance, $C_{\succ^*}(X) = \{c, d\}$, while $C^f_{\succ^*}(X) = \{c\}$, when $b \succ_1^* a \succ_1^* c \succ_1^* d \succ_1^* e$ and $e \succ_2^* d \succ_2^* c \succ_2^* b \succ_2^* a$. The modification thus amounts to add some options to the fallback solution in some cases, and will satisfy RA *a fortiori*. By construction, C is regular and satisfies EFF. ATT does not apply in those cases where C is different from C^f (because the Pareto dominated option falls below an option that is not chosen in the quadruplet obtained by deleting that Pareto dominated option), and hence C satisfies it (since C^f does). Finally, SYM is violated. For instance, $C_{\succ^*}(X) = \{c, d\}$, c is selected from any quadruple that includes it, but $d \notin C_{\succ^*}(X \backslash \{e\})$.

5. REVEALED PREFERENCES

The previous two sections suggest that the fallback bargaining procedure may potentially explain systematic violations of WARP in multi-criteria decision problems. One difficulty in testing this hypothesis is that in many situations we do not directly observe the criteria used by the decision-maker, nor do we observe how the options are ranked according to each criterion. All we may hope to observe are the final choices across different decision problems. A natural question that arises is, what properties of these choices are necessary and sufficient to represent the decision-maker as if he has two criteria in his mind for ranking the options, and he resolves the conflict between these criteria by applying the fallback bargaining procedure? Suppose the observed choices do satisfy the sufficient conditions of the representation, can we identify (at least partially, and, if so, to what extent) the two underlying linear orderings? We answer both questions in this section.

These questions are also relevant for understanding the outcomes of some real life instances of bilateral bargaining between two distinct parties. One example that fits our framework is the process by which parties to a labor-management dispute choose an arbitrator from the list provided by the FMCS (see also Bloom and Cavanagh (1986)). Our main result in this section makes a first step towards providing testable implications of fallback bargaining, which may potentially explain the choices of arbitrators in these disputes.[15]

Characterization

The approach we take is to try and adapt Theorem 1 to bargaining solutions that are not preference-based. Note first that the three regularity conditions of the previous section are no longer useful as they restricted the behavior of the solution across different preference profiles. However, the main properties of Theorem 1 can be suitably adapted

[15] As we discuss in the concluding section, our characterization result relies on the existence of sufficiently rich data, which may not be available in empirical applications.

to the current environment.[16]

RA and SYM can be rephrased literally:

Removing an Alternative (RA) - *If $C(S) \neq \{x\}$, then $C(S \setminus \{x\}) \cap C(S) \neq \emptyset$.*

Symmetry (SYM) - *If $x, y \in C(S)$ and there exists $z \in S \setminus \{x, y\}$ such that $x \notin C(S \setminus \{z\})$, then there exists $z' \in S \setminus \{x, y\}$ such that $y \notin C(S \setminus \{z'\})$.*

To adapt ATT and EFF, we propose to interpret $C(\{x, y\}) = \{x\}$ as the observation that x is "revealed to be Pareto superior" to y. That is, whatever dimensions or criteria the decision-maker uses to evaluate the two options, x is better than y according to all of them. On the other hand, $C(\{x, y\}) = \{x, y\}$ means that there is a negative correlation when comparing x and y across dimensions: x is preferred to y along one, while y is preferred to x along the other. EFF and ATT can now be rephrased using only observed choices:

Efficiency (EFF) - *If $x \in C(S)$, then there does not exist $y \in S$ such that $C(\{x, y\}) = \{y\}$.*

Attraction (ATT) - *Let $x \in X \setminus S$ be such that $C(\{x, y\}) = \{y\}$, for some $y \in C(S)$. Then $C(S \cup \{x\}) = \{y \in C(S) | C(\{x, y\}) = \{y\}\}$.*

To redefine NBC, we say that z is "revealed to be a compromise between x and y" if it is chosen uniquely from $\{x, y, z\}$, but no element in this triplet is revealed to be Pareto superior to another.

No Better Compromise (NBC) - *If both x and y belong to $C(S)$, then there does not exist $z \in S$ such that the choice out of any pair in $\{x, y, z\}$ is the pair itself, and $C(\{x, y, z\}) = \{z\}$.*

The above properties, however, do not guarantee the existence of two linear orderings such that the decision maker's choices can be explained by applying the fallback solution. First, these properties (and in particular, RA, which weakens WARP) do not imply that the revealed Pareto relation is transitive. Thus, to have any hope of recovering a pair of preferences, the following condition must be met:

Pairwise Consistency (PC) - *If $C(\{x, y\}) = \{x\}$ and $C(\{y, z\}) = \{y\}$, then $C(\{x, z\}) = \{x\}$.*

Second, none of the above properties imply the compromise effect. To see this, suppose the decision-maker has a pair of orderings in his mind (which are not observeable to us)

[16] For notational simplicity, we keep the same names for the axioms than in the previous section. Of course, though their motivation is similar, their formulation is not since the models are different. We feel this would not create any confusion since the implied meaning is clear in each section given the relevant context.

such that $x \succ_1 z \succ_1 y$, while $y \succ_1 z \succ_1 x$. Then a choice rule that picks $\{x, y\}$ would satisfy NBC without exhibiting the compromise effect. We must, therefore, take into account a new testable implication: if there is no revealed Pareto comparison between any two elements of $\{x, y, z\}$, then there must be a revealed compromise.

Existence of a Compromise (EC) - *If the choice out of any pair in $\{x, y, z\}$ is the pair itself, then $C(\{x, y, z\})$ is a singleton.*

The next two examples motivate our final axiom. They demonstrate that none of our axioms thus far guarantee that revealed compromises, and their interaction with revealed Pareto dominance, are consistent with an underlying pair of preferences.

Example 1 *Let $X = \{a, b, c, d\}$ and let C be the bargaining solution that selects both elements out of any pair, and such that $C(\{a, b, c\}) = \{b\}$, $C(\{a, b, d\}) = \{d\}$, $C(\{a, c, d\}) = \{d\}$, $C(\{b, c, d\}) = \{d\}$, and $C(\{a, b, c, d\}) = \{d\}$. It is not difficult to check that C satisfies the seven axioms listed so far, but there is no pair $(\succ_1, \succ_2)$ of linear orderings such $C = C^f_{\succ}$. The inconsistency leading to this impossibility is easy to understand: $C(\{a, b, c\}) = \{b\}$ reveals that b is "in between" a and c, while $C(\{a, b, d\}) = \{d\}$ and $C(\{b, c, d\}) = \{d\}$ reveals that d is "in between" both a and b, and b and c.*

Example 2 *Let $X = \{a, b, c, d\}$ and let $(\succ_1^*, \succ_2^*)$ be the two linear orderings defined as follows: $d \succ_1^* a \succ_1^* b \succ_1^* c$ and $d \succ_1^* c \succ_1^* b \succ_1^* a$. Let C be the bargaining solution such that $C(\{b, d\}) = \{b, d\}$ and $C(S) = C^f_{\succ^*}(S)$, for all $S \subseteq X$ different from $\{b, d\}$. It is not difficult to check that C satisfies the seven axioms listed so far, but there is no pair $(\succ_1, \succ_2)$ of linear orderings such $C = C^f_{\succ}$. The inconsistency here is rooted in the way revealed Pareto comparisons should combine with revealed ompromises: b is revealed to be "in between" a and c, d is revealed to be Pareto superior to both a and c, yet b is revealed non comparable to d.*

To rule out the inconsistencies illustrated in these examples, we introduce a property that captures another sense in which compromises have a special status. Suppose y is revealed to be a compromise between x and z. One way to interpret this is that after a long process of deliberation - where one party argued in favor of x, while the other argued in favor of z - the two parties agreed to settle on y. Thus, the choice of y may be viewed as internalizing all the considerations in favor of each of the alternatives. This suggests that if a new option, w, becomes available, the parties would compare w only with y, and would not ignore the previous arguments that led to the agreement on y by opening up the discussion on all available options. Furthermore, if reaching a compromise has special status to the bargainers, then they would require a good enough reason to abandon it completely in favor of a new option. In particular, the parties may replace a compromise with a new option only when the latter Pareto dominates the former.

Overcoming a Compromise (OC) - *Suppose that* $C(\{x,y\}) = \{x,y\}$*,* $C(\{x,z\}) = \{x,z\}$*,* $C(\{y,z\}) = \{y,z\}$*, and* $C(\{x,y,z\}) = \{y\}$*. If* $C(\{w,x,y,z\}) = \{w\}$*, then* $C(\{y,w\}) = \{w\}$*.*

The fallback bargaining solution satisfies an axiom of this type for all bargaining problems, but we phrased it for bargaining problems with only four elements because this is all what is needed to establish our result, as hinted by the two previous examples.

Our second main result establishes that the testable implications we have identified are also sufficient to guarantee the existence of two linear orderings such that the decision maker's choices may be explained by the fallback solution.

Theorem 2 *A bargaining solution* C *satisfies EFF, ATT, NBC, RA, SYM, PC, EC and OC if and only if there exists* $\succ \in L(X) \times L(X)$ *such that* $C = C^f_\succ$*.*

We start by providing a sketch of the proof of the necessity part in Theorem 2. The formal proof follows. The argument unfolds in two main steps. First, we show that a choice correspondence C satisfying EFF, ATT, NBC, RA and SYM exhibits the following property: if there exists a preference profile $\succ$ such that C coincides with $C^f_\succ$ on all pairs and triplets, then this true on all subsets of X. To prove this, we adapt the arguments from the first step of the proof of Theorem 1, which established a similar claim for preference-based bargaining solutions.

In the second step, the more challenging part of the proof, we construct a preference profile $\succ$ such that C coincides with $C^f_\succ$ on all pairs and triplets. The difficulty here lies in the requirement that two preference relations defined on one pair or triplet must be consistent with relations defined on different pairs and triplets. For example, when we are given $C(\{x,y\}) = \{x,y\}$, we conclude that one bargainer prefers x to y, while the other bargainer has the opposite ranking. Suppose we are also given that $C(\{y,z\}) = \{y,z\}$. Then, again, we conclude that the two bargainers have opposite rankings of y and z. The question is, how do we determine whether or not the bargainer who ranks x to y also ranks y to z?

To answer this question, we use the choice data from triplets, and construct the two linear orderings inductively. We begin with one pair of elements and construct two preference relations over them. We then add a third element and extend the previous pair of preferences to cover all three elements. We then continue adding one element at a time and extending the relations from the previous step to cover the newly added element until we have covered all of X.

However, for this construction to succeed, the elements must be added in a particular order. First, we partition the set of elements into "revealed Pareto layers". The highest Pareto layer, denoted EFF^1, consists of all the elements in X that are not revealed to be Pareto inferior to any other element. Similarly, the second-highest Pareto layer, EFF^2, is defined as the set of elements in $X \backslash EFF^1$ that are not revealed to be Pareto inferior to any element not in EFF^1. The next revealed Pareto layers are defined in a similar manner. Each Pareto layer EFF^k is further partitioned into "inner" layers defined as follows. The most extreme layer, denoted $\mathcal{E}^{k,1}$, contains the set of elements (at most two) that are never revealed to be compromises within the Pareto layer EFF^k. The next inner layer contains those elements that are never revealed to be compromises within $EFF^k \backslash \mathcal{E}^{k,1}$. Continuing this way we end with the most interior layer. Given these partitions, the construction of the two preference relations proceeds as follows: we begin with the highest Pareto layer from which we choose the most extreme points and move inward. Once we cover the entire Pareto layer, we move to the next Pareto layer and again, begin with the extreme points and move inwards. A series of lemmas in the proof of Theorem 2 establish that the above method leads to two preference relations that are well-defined and transitive.

To better understand the role of the particular method of constructing the preferences, it is instructive to understand why a simpler inductive argument would not work. Let $X = \{v, w, x, y, z\}$ and suppose the two preference relations we wish to uncover are $y \succ_1 v \succ_1 x \succ_1 w \succ_1 z$ and $x \succ_2 z \succ_2 w \succ_2 y \succ_2 v$ (see Figure 3).

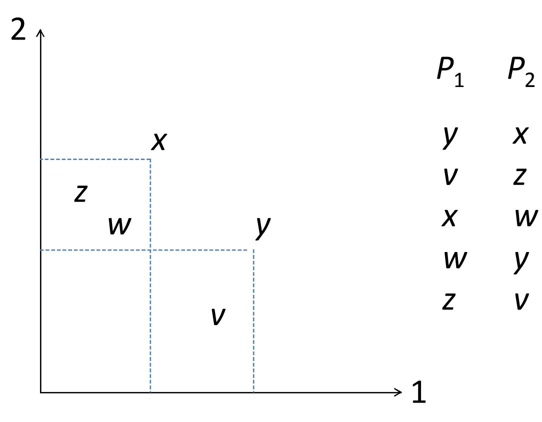


Figure 3

Assume we have data on the choices from all subsets of X such that these choices are consistent with the above two orderings. Assume also that we do not observe the above orderings but wish to infer them using a simple construction that proceeds in Pareto layers but pays no attention to whether elements within a Pareto layer are "extreme points" or not.

According to the data, we may conclude that the highest Pareto layer includes only x and y. Hence, one party ranks x above y, while the other ranks y above x. Since we have no information as to which party prefers x and which party prefers y, let's say that one guesses correctly, i.e. $x \succ_2^* y$ and $y \succ_1^* x$, where $\succ^*$ is the pair of linear orderings that we are deriving from observing the data. We next wish to extend these preference relations to cover the remaining elements. Note that in the second Pareto layer there are two extreme points, v and z (i.e., $\mathcal{E}^{2,1} = \{z, v\}$) and a single interior point, w ($\mathcal{E}^{2,2} = \{w\}$). Suppose that instead of starting with the extreme points, z and v, we start with w. Since $C(\{x, w\}) = \{x\}$, it must be that $x \succ_1^* w$ and $x \succ_2^* w$. This means that $y \succ_1^* w$. Since $C(\{y, w\}) = \{y, w\}$, it must be that $x \succ_2^* w \succ_2^* y$. Next we add z. From $C(\{x, z\}) = \{x\}$ and $C(\{y, z\}) = \{y, z\}$, it follows that $y \succ_1^* x \succ_1^* z$ and $x \succ_2^* z \succ_2^* y$. Since $C(\{z, w\}) = \{z, w\}$, one preference ordering ranks z above w, while another ranks w above z. The information on C and the simple induction construction we are using do not provide us with any guidance on whether $z \succ_2^* w$ or $w \succ_2^* z$. We might thus (wrongly) assume that $w \succ_2^* z$ and $z \succ_1^* w$, and then obtain that $y \succ_1^* x \succ_1^* z \succ_1^* w$ and $x \succ_2^* w \succ_2^* z \succ_2^* y$. Finally, we add v. Since $C(\{y, v\}) = \{y\}$ while $C(\{x, v\}) = \{x, v\}$, we end up with $y \succ_1^* v \succ_1^* x \succ_1^* z \succ_1^* w$ and $x \succ_2^* w \succ_2^* z \succ_2^* y \succ_2^* v$. But if we now apply $C^f_{\succ^*}$ to $\{z, w, v\}$, we obtain $\{z\}$, in contradiction to our observation that $\{w\}$ is chosen from this triplet. Even though the data can be derived as the fallback solution to a pair of linear orderings, the above method did not guarantee that we will be able to construct these orderings.

The difficulty in the reasoning of the previous paragraph is that the data gave us multiple options for defining $\succ^*$ at some steps of the induction, while only some of them work. Our refined inductive argument, on the other hand, provides a construction that works if and only if the data can be generated by some pair of linear orderings. When multiple options for defining the revealed preferences occur, it is only due to a lack of identifiability (see Theorem 3). As an illustration, suppose we constructed the preference relations using our "double induction" argument. When adding the elements in $EFF^2 = \{z, w, v\}$ we begin with the extreme points, $\mathcal{E}^{k,1} = \{z, v\}$. Since $C(\{x, z\}) = \{x\}$ while $C(\{y, z\}) = \{y, z\}$, we have that $y \succ_1^* x \succ_1^* z$ and $x \succ_2^* z \succ_2^* y$. Since

$C(\{x,v\}) = \{x,v\}$ while $C(\{y,z\}) = \{y\}$, we obtain, $y \succ_1^* v \succ_1^* x$ and $x \succ_2^* y \succ_2^* v$. Therefore, $y \succ_1^* v \succ_1^* x \succ_1^* z$ and $x \succ_2^* z \succ_2^* y \succ_2^* v$. It remains to add w. Since $C(\{x,w\}) = \{x\}$ and $C(\{y,w\}) = \{y,w\}$ it must be that $y \succ_1^* x \succ_1^* w$ while $x \succ_2^* w \succ_2^* y$. Since $C(\{z,w,v\}) = \{w\}$, it must be that w is ranked in between v and z. Given our construction, this means that $v \succ_1^* w \succ_1^* z$ and $z \succ_2^* w \succ_2^* v$, and we obtain the desired pair of orderings.

We are now ready to present the proof of Theorem 2.

Proof: We have already proved in the previous section that $C_\succ^f$ satisfies RA, SYM, EFF, ATT, and NBC, for each $\succ \in L(X) \times L(X)$. PC and EC are straightforward to check, and hence only OC remains. The fallback solution generates the choice data on $\{x,y,z\}$ as in OC only if $x \succ_i y \succ_i z$ and $z \succ_{-i} y \succ_{-i} x$, for some $i \in \{1,2\}$. Hence the minimal score of y in the quadruplet is at least 1. For w to be chosen alone, it must be better than at least two alternatives for each ordering, and hence $w \succ y$, or $C_\succ^f(\{w,y\}) = \{w\}$, as desired.

Let now C be a bargaining solution that satisfies SYM, RA, PC, EFF, ATT, NBC, EC and EXP. It is not difficult to adapt the argument from the first step in the proof of Theorem 1 to show that $C = C_\succ^f$ if $\succ \in L(X) \times L(X)$ is such that $C(T) = C_\succ^f(T)$ for all $T \subseteq X$ with two or three elements. The difficult part is to show that there indeed exists a pair $(\succ_1, \succ_2)$ of linear orderings such that $C(T) = C_\succ^f(T)$ for all $T \subseteq X$ with two or three elements. We will proceed via an inductive argument. For each strictly positive integer k, let EFF^k be the following subset of X:

$$EFF^k = \{x \in X \setminus [\cup_{j=0}^{k-1} EFF^j] | \nexists y \in X \setminus [\cup_{j=0}^{k-1} EFF^j] : C(\{x,y\}) = \{y\}\}$$

(with the convention $EFF^0 = \emptyset$). EFF^1 is the set of elements that are C-Pareto efficient in X. EFF^2 is the set of alternatives that are C-Pareto efficient in $X \setminus EFF^1$. These are "second-best" options in X. Notice that EFF^k is nonempty, for each k such that $X \setminus [\cup_{j=1}^{k-1} EFF^j]$ is nonempty, since X is finite and C satisfies PC. Let K be the smallest positive integer such that $EFF^{K+1} = \emptyset$. X is thus partitioned into a collection $(EFF^k)_{k=1}^K$ of layers of options that are constrained efficient at different levels k.

Each such Pareto layer needs itself to be partitioned into subsets of one or two elements, as follows:

$$\mathcal{E}^{k,l} = \{x \in EFF^k \setminus [\cup_{j=0}^{l-1} \mathcal{E}^{k,j}] | \nexists y,z \in EFF^k \setminus [\cup_{j=0}^{l-1} \mathcal{E}^{k,j}] : C(\{x,y,z\}) = \{x\}\},$$

for each $k \in \{1, \ldots, K\}$, and each strictly positive integer l (with the convention $\mathcal{E}^{k,0} = \emptyset$, for each k). EC implies that a single element must be chosen out of any triplet in EFF^k. $\mathcal{E}^{k,1}$ is the set of elements that are never chosen out of any such triplets. These can be interpreted as extreme elements of the layer EFF^k. $\mathcal{E}^{k,2}$ is the set of elements that are extreme in the sub-layer $EFF^k \setminus \mathcal{E}^{k,1}$, and so on so forth. The next lemma, whose proof is available in the Appendix, highlights the structure of these sets.

Lemma 2 *Let $k \in \{1, \dots K\}$, and let l be a strictly positive integer. If $EFF^k \setminus [\cup_{j=1}^{l-1} \mathcal{E}^{k,j}]$ has at least two elements, then $\mathcal{E}^{k,l}$ is nonempty and contains exactly two elements.*

Let L_k be the smallest positive integer such that $\mathcal{E}^{l,L_k+1} = \emptyset$. EFF^k is thus partitioned into a collection $(\mathcal{E}^{k,l})_{l=1}^{L_k}$ of pairs of alternatives (and perhaps one singleton if $\mathcal{E}^{l,L_k}$ contains only one element). An element that belongs to a layer $\mathcal{E}^{k,l}$ for some large l can be interpreted as not too extreme, in that it is chosen as a compromise out of more triplets in EFF^k.

We are now ready to define $\succ$, and prove that $C(T) = C^f_{\succ}(T)$ for every $T \subseteq X$ with two or three elements, by induction. We start with a pair of elements in X, then add a third element, and so on so forth up to the point all the elements of X have been considered. We have to be careful, though, to follow some special order for the argument to work. It follows from our previous definition that each element of X belongs to a unique atom $\mathcal{E}^{k,l}$, for some $l \in \{1, \dots, L_k\}$ and some $k \in \{1, \dots K\}$. This fact will help us determine the right order in which elements must be added. Indeed, let $(k(x), l(x))$ be these two positive integer associated to x. We will follow the convention that x is added before x' if $(k(x), l(x))$ is lexicographically inferior to $(k(x'), l(x'))$. As we know from Lemma 2, this rule does not uniquely specify the ordering, as an atom $\mathcal{E}^{k,l}$ usually contains two elements. We do not further specify how elements are added in the inductive argument, as this is inconsequential for the construction of $\succ$, and the proof that $C = C^f_{\succ}$ on pairs and triplets.[17]

Let x and y be the two first elements of X for which $\succ$ must be defined. If $C(\{x, y\}) = \{x\}$, then we impose that $x \succ_1 y$ and $x \succ_2 y$. Similarly, if $C(\{x, y\}) = \{y\}$, then we impose that $y \succ_1 x$ and $y \succ_2 x$. Finally, if $C(\{x, y\}) = \{x, y\}$, then we impose that $x \succ_1 y$ and $y \succ_2 x$, or $y \succ_1 x$ and $x \succ_2 y$. Either way works, and one may chooses one of the two options arbitrarily. Of course, $C(\{x, y\}) = C^f_{\succ}(\{x, y\})$, by construction.

Suppose now that $\succ$ has been defined on a subset S of X, and that $C(T) = C^f_{\succ}(T)$ for each $T \subseteq S$ with two or three elements, while the next element to be added is $w \in X \setminus S$.

[17] Identifiability, i.e. the possibility of finding multiple pairs of ordering $\succ$ such that $C = C^f_{\succ}$, is the subject of the next theorem.

We now define the extension $\succ^*$ over $S \cup \{w\}$. Of course, $\succ^*$ is defined so as to coincide with $\succ$ on S, i.e. $x \succ_i^* y$ if and only if $x \succ_i y$, for each $x, y \in S$ and each $i = 1, 2$. The important question to answer is how elements of S compare with w under $\succ^*$. For this, we partition S into two subsets:

$$A_w = \{x \in S | C(\{w, x\}) = \{x\}\}$$

$$B_w = \{x \in S | C(\{w, x\}) = \{w, x\}\}.$$

Notice that $A_w \cap B_w = \emptyset$, and $S = A_w \cup B_w$, because there is no $x \in S$ such that $C(\{w, x\}) = \{w\}$ (given the way we add elements in our inductive argument). For each $x \in A_w$, we impose $x \succ_1^* w$ and $x \succ_2^* w$. As for an element $x \in B_w$, we must distinguish two cases. In the first case, we assume that there exists $z, z' \in B_w$ such that $C(\{z, w, z'\}) = \{w\}$. Then we impose $x \succ_1^* w$ and $w \succ_2^* x$ when there exists $y \in B_w$ such that $x \succ_1 y$ and $C(\{x, w, y\}) = \{w\}$, and $w \succ_1^* x$ and $x \succ_2^* w$ when there exists $y \in B_w$ such that $y \succ_1 x$ and $C(\{x, w, y\}) = \{w\}$. We need to check that this is well-defined. This follows from the next lemma, whose proof is available in the appendix.

Lemma 3 *If there exists $z, z' \in B_w$ such that $C(\{z, w, z'\}) = \{w\}$, then, for each $x \in B_w$, there exists $y \in B_w$ such that $C(\{x, w, y\}) = \{w\}$. In addition, if $y, y' \in B_w$ are such that $C(\{x, w, y\}) = C(\{x, w, y'\}) = \{w\}$, then $x \succ_i y$ if and only if $x \succ_i y'$, for both $i = 1, 2$.*

In the second case, namely when there does not exist $z, z' \in B_w$ such that $C(\{z, w, z'\}) = \{w\}$, we impose $x \succ_1^* w$ and $w \succ_2^* x$ if there exists $\xi \in A_w$ and $y \in B_w$ such that $y \succ_1 \xi$, and $w \succ_1^* x$ and $x \succ_2^* w$ if there exists $\xi \in A_w$ and $y \in B_w$ such that $y \succ_2 \xi$. If there is no $\xi \in A_w$ and no $y \in B_w$ such that either $y \succ_1^* \xi$ or $y \succ_2^* \xi$, then one is free to choose either definition, i.e. $x \succ_1^* w$ and $w \succ_2^* x$, for all $x \in S$, or $w \succ_1^* x$ and $x \succ_2^* w$, for all $x \in S$. Here too we need to check that this is well defined. This follows from the next lemma, whose proof is available in the appendix.

Lemma 4 *If there does not exist $z, z' \in B_w$ such that $C(\{z, w, z'\}) = \{w\}$, then there do not exist $\xi, \xi' \in A_w$ and $y, y' \in B_w$ such that $y \succ_2 \xi$ and $y' \succ_1 \xi'$.*

Now that the pair $\succ^*$ of linear orderings has been defined on $S \cup \{w\}$, we should check that they are transitive, i.e. for $i = 1, 2$, $x \succ_i^* w$ if $x \succ_i y$ and $y \succ_i^* w$, $x \succ_i y$ if $x \succ_i^* w$ and $w \succ_i^* y$, and the reverse rankings of both of these cases. We postpone the argument to the appendix.

We will be done with our inductive argument and the proof of Step 2 after proving that $C(T) = C^f_{\succ^*}(T)$, for all $T \subseteq S \cup \{w\}$ with two or three elements. When $w \notin T$ this follows directly from the inductive step. Consider some pair $\{x, w\}$, where $x \in S$. If $x \in A_w$, then $C(\{x, w\}) = \{x\}$ and $\succ^*$ satisfies: $x \succ^*_1 w$ and $x \succ^*_2 w$. Hence, $C^f_{\succ^*}(\{x, w\}) = \{x\}$ as well, as desired. If $x \in B_w$, then $C(\{x, w\}) = \{x, w\}$ and $\succ^*$ satisfies: $x \succ^*_i w$ and $w \succ^*_{-i} x$ for some $i \in \{1, 2\}$. Hence, $C^f_{\succ^*}(\{x, w\}) = \{x, w\}$ as well, as desired.

Consider next a triplet $\{x, y, w\}$. If $\{x, y\} \subseteq A_w$, then $x \succ^* w$ and $y \succ^* w$. The inductive step and ATT imply: $C(\{x, y, w\}) = C(\{x, y\}) = C^f_{\succ^*}(\{x, y\}) = C^f_{\succ^*}(\{x, y, w\})$, as desired.

Suppose next that only one of the alternatives in $\{x, y\}$, say x, belongs to A_w, in which case $y \in B_w$. PC implies that $C(\{x, y\}) = \{x\}$ or $\{x, y\}$. In the former case, x is the only C-efficient (resp. $\succ^*$-efficient) option in $\{x, y, w\}$, and hence $C(\{x, y, w\}) = \{x\} = C^f_{\succ^*}(\{x, y, w\})$, by EFF, as desired. If $C(\{x, y\}) = \{x, y\}$, then $C(\{x, y, w\}) = \{x\}$ by ATT. The constructed preference profile $\succ^*$ satisfies $x \succ^*_i w \succ^*_i y$ and $y \succ^*_{-i} x \succ^*_{-i} w$ (here we use the fact that $\succ^*_i$ is transitive, which is proven in the appendix), for some $i \in \{1, 2\}$. Hence $C^f_{\succ^*}(\{x, y, w\}) = \{x\}$ as well, as desired.

Finally, we consider the case in which neither x nor y belong to A_w. This means that $x, y \in B_w$. Suppose that $C(\{x, y\})$ is a singleton, say $\{x\}$. Then, $C(\{x, y, w\}) = \{x\}$, by ATT. The constructed preference profile $\succ^*$ satisfies $x \succ^*_i y \succ^*_i w$ and $w \succ^*_{-i} x \succ^*_{-i} y$ (again, remember that $\succ^*_i$ and $\succ^*_{-i}$ are transitive), for some $i \in \{1, 2\}$. Hence $C^f_{\succ^*}(\{x, y, w\}) = \{x\}$ as well, as desired.

Now comes the last, and most difficult, case where $C(\{x, y\}) = \{x, y\}$ and $x, y \in B_w$. By construction, $x \succ_i y$ and $y \succ_{-i} x$, for some $i \in \{1, 2\}$. Since the choice out of any pair in $\{x, y, w\}$ is the pair itself, EC implies that $C(\{x, y, w\})$ is a singleton. Assume w.l.o.g. that x has been added before y in the induction.

If $C(\{x, y, w\}) = \{w\}$, then by construction, $x \succ^*_i w \succ^*_i y$ and $y \succ^*_i w \succ^*_i x$. Therefore, $C^f_{\succ^*}(\{x, y, w\}) = \{w\}$ as well, as desired.

Assume $C(\{x, y, w\}) = \{x\}$. Observe that $k(x) \leq k(y) \leq k(w)$, since y is added after x, and w after y. In addition, x, y and w cannot all lie in the same C-Pareto layer, i.e. $k(x) < k(w)$. To see why, suppose on the contrary that $\{x, y, w\} \subseteq EFF^{k(x)}$. Then $l(x) \leq l(y) \leq l(w)$, since y is added after x, w is added after y. Hence, by the definition of $\mathcal{E}^{k(x),l(x)}$, $C(\{x, y, w\}) \neq \{x\}$, a contradiction. Since $k(w) > k(x)$, there must exist $w' \in S$ such that $k(w') = k(x)$ and $C(\{w, w'\}) = \{w'\}$. Lemma 9 from the Appendix implies that $C(\{x, y, w'\}) = \{x\}$. Hence $C^f_{\succ}(\{x, y, w'\}) = \{x\}$, by the induction hypothesis, and we must have: $w' \succ_i x \succ_i y$ and $y \succ_{-i} x \succ_{-i} w'$. Since $C(\{w, w'\}) = \{w'\}$, we know that $w' \succ^* w$. By transitivity, we get $x \succ^*_{-i} w$. Since $C(\{x, w\}) = \{x, w\}$, we have $w \succ^*_i x$. Hence $C^f_{\succ^*}(\{x, y, w\}) = \{x\}$, as desired.

Assume finally that $C(\{x, y, w\}) = \{y\}$. If $k(x) = k(y) = k(w)$, then $l(x) \leq l(y) \leq l(w)$, since y is added after x, and w after y. In order to have $C(\{x, y, w\}) = \{y\}$, it must be that $l(y) > l(x)$, by definition of $\mathcal{E}^{k(x),l(x)}$. Lemma 2 implies that there exists another element x' in $\mathcal{E}^{k(x),l(x)}$. Since $l(y) > l(x)$, it must be that $C(\{x, y, x'\}) = \{y\}$. In order to satisfy the induction hypothesis and the convention $x \succ_i y$, we must have $y \succ_i x'$. Since $l(w) > l(x)$, it must be that $C(\{x, w, x'\}) = \{w\}$. The second statement from Lemma 7 in the Appendix implies that $C(\{w, y, x'\}) \neq \{y\}$, since $C(\{x, y, w\}) = \{y\}$. On the other hand, $C(\{w, x', y\})$ must be a singleton by EC, and cannot be $\{x'\}$ either, since $l(x') < l(y) \leq l(w)$. Hence $C(\{w, x', y\}) = \{w\}$, and $y \succ^*_i w$, by definition. We conclude that $x \succ_i y \succ^*_i w$ and $w \succ^*_{-i} y \succ_{-i} x$, which implies $C^f_{\succ^*}(\{w, x, y\}) = \{y\}$, as desired.

To conclude, suppose that $k(x) < k(w)$. Since $C(\{x, w\}) = \{x, w\}$ and $C(\{y, w\}) = \{y, w\}$, we have three cases to consider:

Case 1) $x \succ_i y \succ^*_i w$ and $w \succ^*_{-i} y \succ_{-i} x$,

Case 2) $x \succ^*_i w \succ^*_i y$ and $y \succ^*_{-i} w \succ^*_{-i} x$, and

Case 3) $w \succ^*_i x \succ_i y$ and $y \succ_{-i} x \succ^*_{-i} w$.

If Case 1 prevails, then $C^f_{\succ^*}(\{x, y, w\}) = \{y\}$. So we will be done after proving that Cases 2 and 3 are impossible.

In Case 2 there are elements on both sides of w according to $\succ^*$, hence, we may apply Lemma 3. Thus, there exists $x' \in B_w$ such that $C(\{x, w, x'\}) = \{w\}$. It must be that $C(\{x', y\}) = \{x', y\}$, as otherwise we get a contradiction with $C(\{w, x, y\}) = \{y\}$ via Lemma 9. Since $x \succ^*_i w$, it must be that $x \succ_i x'$. Transitivity of $\succ^*$ also implies that $x \succ_i y$. So we have two subcases to consider:

Case 2a: $x \succ_i y \succ_i x'$ and vice versa for $-i$ (because the choice out of both $\{x, y\}$ and $\{x', y\}$ is the pair itself), and

Case 2b: $x \succ_i x' \succ_i y$ and vice versa for $-i$.

Knowing that $C(\{x, w, x'\}) = \{w\}$ and $C(\{w, x, y\}) = \{y\}$, subcase 2b (leading to $C(\{x, x', y\}) = \{x'\}$, by the induction hypothesis) is incompatible with RA, given that $C(\{w, x, x', y\})$ contains at most two elements (see Lemma 5 in the Appendix). RA can be satisfied in case 2a only if $C(\{w, x, x', t\}) = \{y\}$ or $\{w, y\}$. The former leads to a contradiction with EXP. In the second case, notice that a single option must be selected out of $\{w, x', y\}$ by EC, and it must be w by RA and SYM. Recall that $y \succ_i x'$ in case 2a, and hence, $y \succ^*_i w$ by definition of $\succ^*$, in contradiction to Case 2.

As for Case 3, let $w' \in EFF^{k(x)}$ be such that $C(\{w, w'\}) = \{w'\}$. Hence $w' \succ_i^* w$, by definition, and transitivity implies that $w' \succ_i x \succ_i y$. $C(\{w', x\}) = \{w', x\}$ then implies $y \succ_{-i} x \succ_{-i} w'$. On the one hand, we could conclude that $C(\{w', y\}) = \{w', y\}$, and hence $C(\{x, y, w'\}) = \{y\}$ by Lemma 9, or $C^f_\succ(\{x, y, w'\}) = \{y\}$, by the induction hypothesis. On the other hand, if one can compute $C^f_\succ(\{x, y, w'\})$ directly from $\succ$, in which case one gets $\{x\}$, hence the contradiction. ■

We next prove the independence of the axioms appearing in Theorem 2.

EFF: Consider the choice correspondence $C_{\succ^*}$ introduced when showing that EFF does not follow from the other axioms in Theorem 1. A similar argument implies that $C_{\succ^*}$ satisfies the current versions of ATT, NBC, RA, and SYM, but violates EFF. EC and PC are satisfied since $C_{\succ^*}$ coincides with the fallback solution on pairs and triplets. OC does not apply, and is thus satisfied trivially.

ATT: Let $\succ$ be a pair of linear orderings on X satisfying that there exists at least one pair of elements, $x, y \in X$ such that $x \succ y$. Define $C(S)$ be a choice correspondence defined as the fallback solution applied only to the set of $\succ$-Pareto efficient alternatives in S, $C(S) = C^f_\succ[EFF_\succ(S)]$, where $EFF_\succ(S)$ is defined in (1). We've already shown that this choice correspondence violates ATT, while satisfying NBC, RA, EFF and SYM. $C(\{x, y\}) = \{x\}$ and $C(\{y, z\}) = \{y\}$ imply that $x \succ y$ and $y \succ z$, which in turn implies that $x \succ z$, and hence, $C(\{x, z\}) = \{x\}$. This verifies PC. If the choice out of any pair in $\{x, y, z\}$ is the pair itself, then $EFF_\succ(\{x, y, z\}) = \{x, y, z\}$. Since $C^f_\succ(\{x, y, z\})$ is a singleton, so is $C(\{x, y, z\})$, which verifies EC. Suppose the choice out of any pair in $\{x, y, z\}$ is the pair itself and $C(\{x, y, z\}) = \{y\}$. Then both individuals must rank y in between x and z. If $C(\{w, x, y, z\}) = \{w\}$, then w cannot be Pareto dominated by any of the elements. It is easy to check that, if w does not Pareto dominate y, then y also belongs to $C(\{x, y, z, w\})$, a contradiction. Hence, it must be that $w \succ y$, confirming OC.

NBC: Consider two orderings $\succ_1$ and $\succ_2$ that are opposite on X: $x \succ_1 y$ if and only if $y \succ_2 x$. Let then C be the choice correspondence defined as follows: $C(S) = C^f_\succ(S)$ if S has exactly three elements, and $C(S) = S$ otherwise. C satisfies EFF and ATT trivially since no two elements are Pareto comparable under $\succ$. NBC is clearly violated in sets with at least four elements. C is larger than the fallback solution applied to $\succ$, and hence C satisfies RA. SYM are trivially satisfied when applied to any set whose cardinality is not equal to four. For any element in a quadruplet, there exists a triplet where that element is available, yet not selected, and hence SYM is verified. PC and EC are satisfied since C coincides with the fallback on pairs and triplets. OC does not apply since C never selects a singleton in quadruplets.

RA: As in the previous example, two orderings $\succ_1$ and $\succ_2$ that are opposite on X. Inspired by Masatlioglu et al. (2009), suppose that the decision maker can pay attention to at most five options. Formally, he has an attention filter $\alpha : P(X) \to P(X)$: $\alpha(S) \subset S$, for all $S \subseteq X$ such that $|\alpha(S)| = \min\{5, |S|\}$ and $x \in \alpha(T)$ if $x \in T \subseteq S$. We will assume in addition that, if x and y belong to $\alpha(S)$, then there does not exist $z \in S \backslash \alpha(S)$ that falls in between x and y according to $\succ$ (it is easy to construct various attention filters with this property). Let then $C(S) = C^f_{\succ}(\alpha(S))$. EFF and ATT are both satisfied because the choice out of any pair is the pair itself. EC is satisfied because the choice out of any triplet is a singleton. NBC is satisfied because of the second property we imposed on the attention filter. OC is satisfied because there is no singleton choice out of quadruplets. PC is satisfied because the choice out of any pair is the pair itself. If $C(S)$ contains two elements, then it must be that S contains at most four elements, in which case C coincides with the fallback, and hence C satisfies SYM. Finally, let's check that C violates RA. Indeed, let S be a set that contains six elements. Let y be the element of S that does not belong to $\alpha(S)$. Let i be such that y is better than the element selected in S for $\succ_i$. Let then $z \in \alpha(S)$ be an option that is worse than the element selected in S for $\succ_i$. It is easy to check that the element selected in $S \setminus \{z\}$ is different from the element selected in S, thereby showing that RA is violated.

SYM: Consider the choice correspondence $C_{\succ^*}$ introduced when showing that SYM does not follow from the other axioms in Theorem 1. A similar argument implies that $C_{\succ^*}$ satisfies the current versions of EFF, ATT, NBC, and RA, but violates SYM. PC, EC and OC are satisfied since $C_{\succ^*}$ coincides with the fallback solution on pairs and triplets.

PC: Let P be a strict complete and transitive ordering on X, and let C be the choice obtained by maximizing this ordering, except that $C(\{x, y\}) = \{x, y\}$, where x is the best element in X and y is the worst element in X. It is easy to check that C satisfies all the axioms of Theorem 2, except PC.

EC: Consider the analogue of the Borda rule introduced in the previous section to show that NBC is not implied by the other axioms in Theorem 1. We already proved that it satifies EFF, ATT, NBC, RA, and SYM, for any given pair of preferences. PC is straightforward to check, and OC never applies because the choice out of any triplet is the triplet itself if no two elements are Pareto comparable. For the same reason, the choice correspondence will violate EC, as soon as there are at least three elements that are not Pareto comparable.

OC: See the two examples given before introducing OC.

Identifiability

There is no hope to identify uniquely the underlying preference relations on both dimensions. Indeed, there is no way to tell which ordering should be associated to a specific self or dimension of choice: if $C = C^f_{\succ}$, for some pair $(\succ_1, \succ_2)$ of linear orderings on X, then we also have $C = C^f_{(\succ_2, \succ_1)}$ (cf. second regularity condition in the previous section). One may wonder whether this is the only source of multiplicity. The answer is not quite, but almost, as the following example and theorem illustrate.

Example 3 *Consider $X = \{a, b, c, d\}$ and $C = C^f_{\succ}$, where $a \succ_1 b \succ_1 c \succ_1 d$ and $b \succ_2 a \succ_2 d \succ_2 c$. It is not difficult to check that C is also equal to $C^f_{\succ'}$, where $b \succ'_1 a \succ'_1 c \succ'_1 d$ and $a \succ'_2 b \succ'_2 d \succ'_2 c$. The careful reader will notice that $\succ'$ is obtained from $\succ$ by exchanging the preferences of the two selves only as far as a and b are concerned. This change is irrelevant as far as the fallback bargaining solution is concerned, because both a and b Pareto dominate both c and d according to $\succ$, implying that c and d are irrelevant when it comes to determine the solution of any subset S of X that include either a, b, or both.*

A subset S of X is *C-dominant* if it is non-empty and $C(\{x, y\}) = \{x\}$, for all $x \in S$ and all $y \in X \setminus S$.[18] Observe that if S and S' are both C-dominant, then $S \subseteq S'$ or $S' \subseteq S$. Also X is trivially C-dominant. So there exists a unique minimal C-dominant set S^*_1 in X. Similarly, a subset S of $X \setminus S^*_1$ is C-dominant in $X \setminus S^*_1$ if it is non-empty and $C(\{x, y\}) = \{x\}$, for all $x \in S^*_1$ and all $y \in X \setminus (S \cup S^*_1)$. Let S^*_2 be the minimal C-dominant set in $X \setminus S^*_1$. Iterating the procedure, one obtains a partition of X into a finite sequence $\Pi = (S^*_1, \ldots, S^*_K)$ of sets with the property that S^*_k is the minimal C-dominant set in $X \setminus \cup_{j=1}^{k-1} S^*_j$.

Theorem 3 *Let $\succ, \succ'$ be two pairs of strict linear orderings. Then $C^f_{\succ} = C^f_{\succ'}$ if and only if $\succ'$ can be obtained from $\succ$ by permuting the two orderings over atoms of Π that contains at least two elements.*

Proof: The sufficient condition is easy to check, and we focus attention only on the necessary condition. Let C be the common bargaining solution. Since it coincides with

[18] If C satisfies EFF, then S is C-dominant if and only if $C(T) \subseteq S$, for each $T \subseteq X$ such that $S \cap T \neq \emptyset$.

the fallback bargaining solution for some pair of orderings, it satisfies the axioms listed in the previous section, and the induction we followed in the proof of Theorem 2 can be reproduced here as well. Let $\succ^*$denote the preference profile that is constructed in the induction procedure, such that $C = C^f_{\succ^*}$ (note that $\succ^*$ may be different from $\succ$ or $\succ'$). For any $x, y \in X$ we write $(k(x), l(x)) \leq^L (k(y), l(y))$ to mean that $(k(x), l(x))$ is *lexicographically* lower or equal to $(k(y), l(y))$, i.e., either $k(x) < k(y)$ or $k(x) = k(y)$ and $l(x) \leq l(y)$.

Let x and y be the first and second elements in the induction, i.e., for all $z \in X \backslash \{x, y\}$,

$$(k(x), l(y)) \leq^L (k(y), l(y)) <^L (k(z), l(z))$$

If $k(x) < k(y)$, then there is only one profile of ranking that is consistent with C: both agents rank x above y. If $k(x) = k(y)$, then both $x \succ_1^* y$, $y \succ_2^* x$ and $x \succ_2^* y$, $y \succ_1^* x$ are consistent with C . Moreover, these are the only consistent profiles: i.e., either $\succ^*$ and $\succ$ agree on $\{x, y\}$, or $\succ^*$ and $\succ'$ agree. Fix one of these profiles. Let w be the first element in the induction (following x and y) with the property that there exists x' with $(k(x'), l(x')) \leq^L (k(w), l(w))$ such that either $\succ^*$ and $\succ$ or $\succ^*$ and $\succ'$ differ on $\{x', w\}$.

We first establish that the following must then be true: there is no $z, z' \in B_w$ such that $C(\{z, w, z'\}) = \{w\}$, and there is no $\xi \in A_w$ and no $y \in B_w$ such that either $y \succ_1^* \xi$ or $y \succ_2^* \xi$. To begin, note that $x' \notin A_w$, since then both agents must rank x' above w. Therefore, $x' \in B_w$. Suppose there exist $z, z' \in B_w$ such that $C(\{z, w, z'\}) = \{w\}$. Then by Lemma (3), there must exist $y' \in B_w$ such that $C(\{x', w, y'\}) = \{w\}$. Since x' and y' were added before w in the induction, the agents' preferences over them has already been determined. Moreover, having fixed the ordering of the first two elements in the induction, our definition of w implies that there is only one preference profile over $\{x', y'\}$, which is consistent with C. Assume, w.l.o.g. that this profile is $x' \succ_1^* y'$ and $y' \succ_2^* x'$. Then it must be that all three preference profiles, $\succ$, $\succ'$ and $\succ^*$, rank w in between x' and y'. Hence, these preference profiles all agree on the rankings of x' relative to w, in contradiction to the definition of w and x'.

Suppose next that there exist no $z, z' \in B_w$ such that $C(\{z, w, z'\}) = \{w\}$. Note that this means that $k(y) < k(w)$. Consider first the case in which there exists $\xi \in A_w$ and $y' \in B_w$ such that $y' \succ_i^* \xi$. Since $y' \in B_w$, PT implies that $C(\{y', \xi\}) = \{y', \xi\}$, and hence, $\xi \succ_{-i}^* y'$. By our construction of $\succ^*$, $x' \succ_i^* w$ and $w \succ_{-i}^* x'$. By Lemma (well-def2), either $x' \succ_i^* \xi$ and $\xi \succ_{-i}^* x'$, or $\xi \succ^* x'$. By our induction and the definition of w, the preference profile $\succ^*$ coincides with both $\succ$ and $\succ'$ on $\{x', \xi\}$.

Suppose $x' \succ_i^* \xi$ and $\xi \succ_{-i}^* x'$. Since $\xi \in A_w$, it must be the case that all three profiles, $\succ$, $\succ'$ and $\succ^*$, satisfy that agent 1 and 2 rank ξ above w. Hence, by transitivity, all three profiles satisfy that agent i ranks x' above ξ and ξ above w. Assume w.l.o.g. that $\succ^*$ differs from $\succ$. Then it must be that either $\xi \succ_{-i}^* x' \succ_{-i}^* w$ but $\xi \succ_{-i} w \succ_{-i} x'$, or that $\xi \succ_{-i} x' \succ_{-i} w$ but $\xi \succ_{-i}^* w \succ_{-i}^* x'$. But if agent $-i$ ranks x' above w, then we get that a contradiction to $x' \in B_w$. Therefore, both $\succ^*$ and $\succ$ must satisfy that agent $-i$ ranks ξ above w and w above x'. But this contradicts our definition of w and x'.

Next, suppose $\xi \succ^* x'$. Then $y' \succ_i^* \xi \succ_i^* x' \succ_i^* w$ and $\xi \succ_{-i}^* w \succ_{-i}^* x' \succ_{-i}^* y'$. Assume w.l.o.g. that $\succ^*$ differs from $\succ$. Then it must be that $w \succ_i x'$ while $x' \succ_{-i} w$. Since $\succ^*$ and $\succ$ coincide on $\{x', y'\}$, we have that $y' \succ_i^* x' \succ_i^* w$ and $w \succ_{-i}^* x' \succ_{-i}^* y'$ while $y' \succ_i w \succ_i x'$ and $x' \succ_{-i}^* w \succ_{-i}^* y'$. But this means that $C^f_{\succ^*}(\{y', x', w\}) = \{x'\}$ while $C^f_{\succ}(\{y', x', w\}) = \{w\}$, a contradiction.

It follows that there are no $z, z' \in B_w$ such that $C(\{z, w, z'\}) = \{w\}$, and there is no $\xi \in A_w$ and no $y \in B_w$ such that either $y \succ_1^* \xi$ or $y \succ_2^* \xi$. Since $x' \in B_w$, we also know that B_w is non-empty.

We now prove that A_w is C-dominant. Consider some $a \in A_w$ and $b \in X \setminus A_w$. We have to prove that $C(\{a, b\}) = \{a\}$. If b is added before w in the induction, then $b \in B_w$, and the result follows trivially from the conclusion that no element in B_w is ever chosen in a pair containing an element in A_w. Suppose now that b is added after w in the induction, i.e. $(k(b), l(b))$ lexicographically dominates $(k(w), l(w))$. Suppose first that $k(b) = k(w)$. Since there is no $z, z' \in B_w$ such that $C(\{z, w, z'\}) = \{z\}$, it must be that $l(w) = 1$. Since B_w is non-empty, it must be that there exists another element w' such that $k(w') = k(w)$ that has been added before w - this must be the other element of the atom $\mathcal{E}^{(k(w),1)}$ (remember that those atoms contain at most two elements, see Lemma ???). Hence $C(\{w, b, w'\}) = \{b\}$. Since there is no element in A_w and no element in B_w from which C picks both elements, $C(\{a, w'\}) = \{a\}$. Since $C = C_{\succ}$, we must have $a \succ w$, $a \succ w'$, and there exists $i \in \{1, 2\}$ such that $w \succ_i b \succ_i w'$ and $w' \succ_i b \succ_i w$. Hence $C(\{a, b\}) = \{a\}$, as desired. Finally, if $k(b) > k(w)$, then there exists x'' such that $k(x'') = k(w)$ and $C(\{x'', b\}) = \{x''\}$ (x'' could be w itself). By essentially the same argument as above, we may conclude that $C(\{a, x''\}) = \{a\}$, and hence, $C(\{a, b\}) = \{a\}$, by PT, as desired.

Fix now an atom S^* of the partition. We will prove that $\succ$ and $\succ^*$ must be the same or a permunation of each other on S^*. This will conclude the proof, since there is only one way of patching together the orderings obtained on the different atoms of Π ($x \succ y$ if and only if x belongs to an atom that comes before the atom to which y belongs).

For the sake of notational simplicity, we will assume that $S^* = X$, but of course the reasoning can be reproduced with any set S^* that has at least two elements, since X is an arbitrary set throughout the paper. Let us add elements as in the induction from the proof of theorem 2. Let x, y be the first two elements to be considered. Notice that $C(\{x, y\}) = \{x, y\}$, as otherwise either $\{x\}$ or $\{y\}$ would be C-dominant, a contradiction with the fact that X does not contain any strict subset that is C-dominant. Let i, j be such that $x \succ_i y$, $y \succ_{-i} x$, $x \succ'_j y$, $y \succ'_{-j} x$. The previous paragraph shows that the last case when defining the extension of $\succ^*$ when adding w in the induction cannot occur. In the three other cases, it is not difficult to check that the way we defined the extension is the only way possible to have that C coincides with the fallback bargaining solution for the extended ordering. Hence by induction, it must be that $\succ_i = \succ'_j$ and $\succ_{-i} = \succ'_{-j}$. If $i = j$, then $\succ = \succ'$ (on S^*). Otherwise $\succ'$ is simply a permutation of $\succ$, as desired. ■

6. CONCLUDING COMMENTS

Testable implications on limited data sets

The axioms listed in Section 5 remain of course necessary conditions for any data set to be consistent with the fallback solution for some pair of linear orderings. Theorem 2 establishes that they are also sufficient when working with a complete data set. While sufficiency is also guaranteed on some smaller data sets, it remains an open question to characterize all those data sets that guarantee sufficiency, and to propose new independent testable implications otherwise. In contrast, rationality is equivalent to the standard IIA axiom on any data set that includes observations on all subsets of two or three elements. The difference in results comes from the fact that our axioms of consistency across subsets (e.g., RA and OC) focus, for simplicity in the absence of rationality, on marginal changes - removing or adding a single alternative.

Even so, Theorem 2 is useful in providing an efficient method for verifying whether a limited data set is consistent with the fallback solution for some preference profile, and in deriving the corresponding preferences of the selves/bargainers. Indeed, its proof provides an algorithm for constructing a candidate profile of preferences. Given this profile, we can compute the fallback solution to predict the choice from any subset. A data set that includes observations on all subsets of two or three elements is consistent with the fallback bargaining solution for some pairs of linear orderings if and only if predictions match actual data. Using this method instead of the alternative procedure of constructing all possible preference profiles and checking for each one whether it can generate the data set with the fallback solution brings the computational complexity from exponential down to polynomial.

Allowing for indifferences

Throughout our analysis we assumed that the bargainers have strict preferences over X. Allowing for indifferences complicates the analysis, and it remains an open question how to extend our results to allow for indifference. One promising direction is to consider the case where X has a product structure, $X = X_1 \times X_2$, while criterion $\succeq_i$ on X is assumed to be strict on X_i (but note that multiple elements of X may now have the same i-component). Axioms would have to be adapted in some cases, for instance EFF would have to be strengthened to a notion of strong efficiency, but we expect that the essence of our three theorems would extend to this new framework.

More than two bargainers

A natural extension of our analysis is to situation in which the elements of choice have more than two dimensions (and the decision-maker is able to process more than two dimensions), or where there are more than two distinct individuals engaged in bargaining. While it is easy to extend the fallback solution to more than two selves, more work is needed to see how the characterization results would extend, both mathematically and in terms of the interpretation of the axioms. One difficulty, for instance, would be the definition of a compromise. Our view of a compromise was of an element that in some sense both bargainers rank "in between" other elements which one bargainer ranks in the opposite way to the other bargainer (geometrically it means that the compromise falls in the rectangle constructed on the two other alternatives). The question is, how do we extend this notion of "betweenness " to more than two bargainers, and how to interpret it in terms of choice behavior? It is interesting to note in that respect that almost all the experiments on the attraction and compromise effects were done on two-dimensional elements.

When alternatives have more than two dimensions, one may further question our assumption that all dimensions are treated equally. A natural extension would be to allow the individual to put different weights on different dimensions, and to make a choice according to, say, a "weighted" fallback solution. When the weights of the dimensions and the ranking within each dimension are not observeable, the revealed exercise would be to try and infer both from observed choices. One potential concern with this is identifiability: the additional freedom to choose the weights on the dimensions may allow the same choice correspondence to be consistent with a wide variety of preferences.

Finally, it is worth noting that the predictive power of the fallback solution diminishes with the number of bargainers since the maximal number of elements it can choose equals the number of bargainers. However, it may very well be that, indeed, if we were to replicate the experiments of the attraction and compromise effects with three-dimensional alternatives, the distribution of choices would cover all three options when offering three options that cannot be Pareto ranked.

Intensities

One has the intuition that the prevalence of the attraction and compromise effects in applications may depend on factors that cannot be captured in our ordinal model. More specifically, choices may be influenced by some trade-offs that involve a notion of distance or intensity. An individual may exhibit a compromise effect when $x = (100, 1)$, $y = (50, 50)$ and $z = (1, 100)$, but (perhaps) not when $y = (2, 2)$. Similarly, an he may be more likely to exhibit the attraction effect when $x = (60, 40)$, $y = (59, 39)$ and $z = (40, 60)$, but (perhaps) not when $y = (41, 39)$. Extending our analysis in this direction is certainly an interesting topic for future research. Note though that the analysis may now depend crucially on the context to which it is designed to be applied, as the model would have to express what it means for two options to be 'close' or 'far away' on each dimension. Perhaps a more promising avenue would be to derive these intensities, and the relevant trade-offs, from the observed choices, but the added flexibility may result in few testable implications. In addition to being detail-free, our ordinal approach has the advantage of allowing straightforward comparisons to existing standard models of choice. We hope that it will also serve as a benchmark for future extensions.

APPENDIX

Proof of Lemma 1

Necessary Condition for Subcase 1: Suppose that $C^f_\succ(S) = \{x\}$. For each $w \in S \setminus \{x\}$ and each $y \in S \setminus \{x, w\}$, we have:

$$\min_{i=1,2} s_i(x, S \setminus \{w\}, \succ) \geq \min_{i=1,2} s_i(x, S, \succ) - 1 \geq \min_{i=1,2} s_i(y, S, \succ) \geq \min_{i=1,2} s_i(y, S \setminus \{w\}, \succ),$$

and hence $x \in C^f_\succ(S \setminus \{w\})$, as desired.

Let now $y \in S \setminus \{x\}$. Suppose that $j \in \arg\min_{i=1,2} s_i(y, S, \succ)$. If there exists $w \in S$ such that $y \succ_j w$, then we have:

$$\min_{i=1,2} s_i(x, S \setminus \{w\}, \succ) \geq \min_{i=1,2} s_i(x, S, \succ) - 1 > \min_{i=1,2} s_i(y, S, \succ) - 1 = \min_{i=1,2} s_i(y, S \setminus \{w\}, \succ),$$

and hence $y \notin C^f_\succ(S \setminus \{w\})$. If there does not exist $w \in S$ such that $y \succ_j w$, then $\min_{i=1,2} s_i(y, S \setminus \{w\}, \succ) = 0$, and $y \notin C^f_\succ(S \setminus \{w\})$, for each $w \in S \setminus \{y\}$, since $|S \setminus \{w\}| \geq 3$, and the minimal score attained at the chosen element(s) is always larger or equal to the first integer below half the number of elements in the choice set.

Necessary Condition for Subcase 2: Suppose that $C^f_\succ(S) = \{x, y\}$. Let $w \in S \setminus \{x, y\}$. Let's assume that $\arg\min_{i=1,2} s_i(x, S, \succ) = 1$ and $\arg\min_{i=1,2} s_i(y, S, \succ) = 2$ (a similar reasoning applies if 1 and 2 are exchanged).

Observe that it is impossible to have $w \succ_1 x$ and $w \succ_2 y$, since the minimal score of w in S would then be larger than the minimal score of both x and y. The minimal score of x (resp. y) is the same in both S and $S \setminus \{w\}$ if $y \succ_2 w$ (resp. $x \succ_1 w$), and therefore remains strictly larger than the minimal score of any element in $S \setminus \{w, x, y\}$ (since it does not increase by deleting w). Hence $C^f_\succ(S \setminus \{w\}) \subseteq \{x, y\}$, as desired.

Suppose now that $C^f_\succ(S \setminus \{w\}) = \{x\}$. This is true if and only if $w \succ_1 x$ and $y \succ_2 w$. Hence there exists $w' \in S$ such that $x \succ_1 w'$ and $w' \succ_2 y$, as otherwise the minimal score of y is strictly larger than the minimal score of x, and $C^f_\succ(S \setminus \{w'\}) = \{y\}$, as desired.

Last Statements of the Lemma: Suppose that $C^f_\succ(S) = \{x, y\}$, and $C^f_\succ(S \setminus \{w\}) = \{x, y\}$, for all $w \in S \setminus \{x, y\}$. Continuing with the notations introduced to prove the necessary condition for subcase 2, we already observed that it is impossible to find a $w \in S$ such that $w \succ_1 x$ and $w \succ_2 y$. The previous paragraph also implies that $C^f_\succ(S \setminus \{w\}) = \{x, y\}$ if and only if we don't have $w \succ_1 x$ and $y \succ_2 w$, nor $x \succ_1 w$ and $y \succ_2 w$. Hence $x \succ w$ and $y \succ w$, as desired.

Suppose now that $C^f_\succ(S) = \{x\}$, $C^f_\succ(S \setminus \{y\}) = \{x, z\}$, and $C^f_\succ(S \setminus \{z\}) = \{x, y\}$. If $y \succ z$, then y looses one point along both dimensions when dropping z, and the minimal score of x remains strictly larger than that of y in $S \setminus \{z\}$, hence a contradiction with $C^f_\succ(S \setminus \{z\}) = \{x, y\}$. Similarly, it cannot be that $z \succ y$. There is no Pareto relation between x and z, and x and y either, since $C^f_\succ(S \setminus \{y\}) = \{x, z\}$ and $C^f_\succ(S \setminus \{z\}) = \{x, y\}$. Let $i \in \{1, 2\}$ be such that $y \succ_i z$. Three cases remain possible: 1) $x \succ_i y \succ_i z$ and $z \succ_{-i} y \succ_{-i} x$; 2) $y \succ_i x \succ_i z$ and $z \succ_{-i} x \succ_{-i} y$; or 3) $y \succ_i z \succ_i x$ and $x \succ_{-i} z \succ_{-i} y$. Consider case 1). Since y is above x along $-i$ and $C^f_\succ(S \setminus \{z\}) = \{x, y\}$, it must be that the minimal score of y in $S \setminus \{z\}$ is attained along the i-dimension, and is equal to the minimal score of x in $S \setminus \{z\}$ which is attained along the $-i$-dimension. Adding z, the minimal score of y increases by one point, while that of x remains unchanged, hence a contradiction with $C^f_{\succ}(S) = \{x\}$. Case 3) leads to a similar contradiction. Hence only case 2) remains, as desired.

Sufficient Condition for Subcase 1: Assuming that conditions 1(a) and (b) are true, we need to prove that $C^f_{\succ}(S) = \{x\}$. If $C^f_{\succ}(S) = \{y\}$ for some $y \in S \setminus \{x\}$, then the necessary condition for subcase 1 implies that $y \in C^f_{\succ}(S \setminus \{w\})$, for all $w \in S \setminus \{y\}$, thereby contradicting 1(b). If $C^f_{\succ}(S) = \{y, z\}$ for some $y, z \in S \setminus \{x\}$, then the necessary condition for subcase 2 implies that $C^f_{\succ}(S \setminus \{w\}) \subseteq \{y, z\}$, for all $w \in S$, thereby contradicting 1(a). Finally, suppose that $C^f_{\succ}(S) = \{x, y\}$ for some $y \in S \setminus \{x\}$. Condition 1(b) implies that there exists $w \in S \setminus \{y\}$ such that $y \notin C^f_{\succ}(S \setminus \{w\})$. The necessary condition for subcase 2 implies that there exists $w' \in S \setminus \{x\}$ such that $C^f_{\succ}(S \setminus \{w'\}) = \{y\}$, thereby contradicting 1(a). We must conclude that $C^f_{\succ}(S) = \{x\}$, as desired.

Sufficient Condition for Subcase 2: Assuming that conditions 2(a) and (b) are true, we need to prove that $C^f_{\succ}(S) = \{x, y\}$. If $z \in C^f_{\succ}(S)$, for some $z \in S \setminus \{x, y\}$, then the necessary condition for subcases 1 and 2 implies that $z \in C^f_{\succ}(S \setminus \{w\})$, for some $w \in S$, thereby contradicting 2(a). If $C^f_{\succ}(S) = \{x\}$, then 1(b) and 2(a) imply that $C^f_{\succ}(S \setminus \{w\}) = \{x\}$, for some $w \in S \setminus \{x\}$. On the other hand, 1(a) implies that $x \in C^f_{\succ}(S \setminus \{w'\})$, for all $w' \in S \setminus \{x\}$, and this leads to a contradiction with condition 2(b). A similar reasoning shows that $C^f_{\succ}(S) \neq \{y\}$, and hence $C^f_{\succ}(S) = \{x, y\}$, as desired. ■

Proof of Lemma 2

Lemma 5 *Let C be a bargaining solution that satisfies EFF, NBC, and EC. Then $|C(S)| \leq 2$, for all $S \subseteq X$.*

Proof: Suppose that one can find three elements x, y, z in $C(S)$, for some $S \subseteq X$. EFF implies that the choice out of any pair in $\{x, y, z\}$ is the pait itself, and EC implies that a single element must be chosen out of the triplet. This contradicts NBC. ■

Lemma 6 *Let C be a bargaining solution that satisfies SYM, RA, PC, EFF, ATT, NBC, and EC. Let w, x, y, z be four distinct elements of X. If $C(\{w, x, y, z\}) = \{x, y\}$, then $C(\{w, x, z\}) = \{x\}$.*

Proof: RA implies that $x \in C(\{w,x,z\})$. Lemma 5 implies that we will be done after proving that $C(\{w,x,z\})$ is not equal to $\{w,x\}$, nor $\{x,z\}$. Since the argument is similar in both cases, we will only show how to rule out the first one. Suppose on the contrary that $C(\{w,x,z\}) = \{w,x\}$. EFF implies that $C(\{w,x\}) = \{w,x\}$, $C(\{w,z\}) \neq \{z\}$, and $C(\{x,z\}) \neq \{z\}$. EC implies that it is impossible to have $C(\{w,z\}) = \{w,z\}$ and $C(\{x,z\}) = \{x,z\}$. ATT also implies that it is impossible to have $C(\{w,z\}) = \{w\}$ and $C(\{x,z\}) = \{x,z\}$, or $C(\{w,z\}) = \{w,z\}$ and $C(\{x,z\}) = \{x\}$. Hence $C(\{x,z\}) = \{x\}$ and $C(\{w,z\}) = \{w\}$. Also, $C(\{w,x,y,z\}) = \{x,y\}$ implies that $C(\{x,y\}) = \{x,y\}$, $C(\{y,z\}) \neq \{z\}$, and $C(\{w,y\}) \neq \{w\}$, by EFF. Notice that $C(\{w,x,y\})$ must be a singleton, because of EC if $C(\{w,y\}) = \{w,y\}$, and because of ATT if $C(\{w,y\}) = \{y\}$. RA implies that $C(\{w,x,y\}) = \{y\}$. If $C(\{y,z\}) = \{y\}$, then $C(\{x,y,z\}) = \{x,y\}$, by ATT, and we get a contradiction with SYM. If $C(\{y,z\}) = \{y,z\}$, then it must be that $C(\{w,y\}) = \{w,y\}$ to avoid a contradiction with PC. ATT thus implies that $C(\{w,y,z\}) = \{w\}$, which contradicts RA. Hence the original hypothesis that $C(\{w,x,z\}) = \{w,x\}$ is false, and we are done with the proof. ■

Lemma 7 *Let C be a bargaining solution that satisfies SYM, RA, PC, EFF, ATT, NBC, EC, and EXP, and let w,x,y,z be four distinct elements of X such that the choice out of any pair is the pair itself. Then the three following statements are true:*

1. *If $C(\{w,x,y\}) = \{x\}$ and $C(\{x,y,z\}) = \{y\}$, then $C(\{w,x,z\}) = \{x\}$.*

2. *It is impossible to have $C(\{x,y,z\}) = \{y\}$, $C(\{x,w,y\}) = \{w\}$, and $C(\{y,w,z\}) = \{w\}$.*

3. *If $C(\{w,x,z\}) = \{x\}$ and $C(\{x,y,z\}) = \{y\}$, then $C(\{w,x,y\}) = \{x\}$.*

Proof: For the first statement, assume that $C(\{w,x,y\}) = \{x\}$ and $C(\{x,y,z\}) = \{y\}$. RA implies that $C(\{w,x,y,z\})$ cannot be w nor y since $w,y \in \{w,x,y\}$ and $C(\{w,x,y\}) = \{x\}$, and cannot be x, nor z, since $x,z \in \{x,y,z\}$ and $C(\{x,y,z\}) = \{y\}$. Lemma 5 implies that $C(\{w,x,y,z\})$ must contain two elements. RA rules out $\{w,y\}$, $\{x,z\}$, $\{w,x\}$, $\{y,z\}$, and $\{w,z\}$. Hence it must be $\{x,y\}$. Applying Lemma 6, we conclude that $C(\{w,x,z\}) = \{x\}$, as desired.

For the second statement, assume that $C(\{x,y,z\}) = \{y\}$, $C(\{x,w,y\}) = \{w\}$, and $C(\{y,w,z\}) = \{w\}$. It is not difficult to check that RA and Lemma 5 imply that $C(\{w,x,y,z\})$ must equal $\{w\}$ or $\{w,y\}$. The former case leads to a contradiction with EXP, while the other leads to a contradiction with SYM.

For the third statement, assume that $C(\{w,x,z\}) = \{x\}$ and $C(\{x,y,z\}) = \{y\}$. EC implies that $C(\{w,x,y\})$ must be a singleton. Suppose that $C(\{w,x,y\}) = \{w\}$. Thanks to the first statement, we can combine this with $C(\{w,x,z\}) = \{x\}$, to conclude that $C(\{w,y,z\}) = \{w\}$. Hence a contradiction with the second statement (w is "in between" both x and y, and y and z, while y is "in between" x and z). If $C(\{w,x,y\}) = \{y\}$, then one gets again a contradiction with the second statement (y is "in between" both w and x, and x and z, while x is "in between" w and z). ■

<u>Proof of Lemma 2</u>: We want to prove that, for each set $Y \subseteq X$ with at least two elements and such that the choice out of any pair in Y is the pair itself, there exist exactly two elements in Y that are not chosen out of any triplet in Y. This is done by induction on the number of elements in Y. The result is trivial if $\#Y = 2$ or 3. Let α be a positive integer larger or equal to 3, and suppose that the result holds for all set with no more than α elements. Consider now a set Y with $\alpha + 1$ elements.

First notice that there cannot be more than two elements in Y that are not chosen out of any triplet, since the choice out of any triplet in Y is a singleton, by EC. Since Y has more than three elements, we can choose $y, x, x' \in Y$ such that $C(\{x,y,x'\}) = \{y\}$. Let ξ, ξ' be the two elements in $Y \setminus \{y\}$ that are not chosen out of any triplet in $Y \setminus \{y\}$ (using the induction hypothesis). We will be done with the proof after showing that these two elements are not chosen out of any triplet in Y. This amounts to show that $C(\{\xi,y,z\}) \neq \{\xi\}$, for all $z \in Y \setminus \{\xi, y\}$, and $C(\{\xi',y,z\}) \neq \{\xi'\}$, for all $z \in Y \setminus \{\xi', y\}$ (since we already know that ξ and ξ' are not chosen out of any triplet in $Y \setminus \{y\}$). We prove the first statement only, the argument with ξ' instead of ξ being similar. We proceed by considering three cases.

Case 1: $\{x, x'\} = \{\xi, \xi'\}$. In that case, we know that $C(\{\xi, y, \xi'\}) = \{y\}$. Suppose to the contrary of what we want to prove that $C(\{\xi, y, z\}) = \{\xi\}$, for some $z \in Y \setminus \{\xi, y\}$. It must be that $z \neq \xi'$, and hence $C(\{\xi, z, \xi'\}) = \{z\}$, by definition of ξ, ξ'. On the other hand, the first statement of Lemma 7 implies that $C(\{\xi, z, \xi'\}) = \{\xi\}$, hence the desired contradiction.

Case 2: $\{x, x'\} \cap \{\xi, \xi'\} \neq \emptyset$, but $\{x, x'\} \neq \{\xi, \xi'\}$. Suppose for instance that $x = \xi$ (the argument for the three other cases $x = \xi'$, $x' = \xi'$, and $x' = \xi$ is similar). We know that $C(\{\xi, y, x'\}) = \{y\}$ and $C(\{\xi, x', \xi'\}) = \{x'\}$ (by definition of ξ, ξ'). Suppose to the contrary of what we want to prove that $C(\{\xi, y, z\}) = \{\xi\}$, for some $z \in Y \setminus \{\xi, y\}$. Observe that $C(\{y, x', \xi'\})$ cannot be $\{y\}$ because of the second statement of Lemma 7, and it cannot be $\{\xi'\}$ to avoid a contradiction with the first statement of Lemma 7.

EC implies that $C(\{y, x', \xi'\}) = \{y\}$. The first statement of Lemma 7 now implies that $C(\{\xi, y, \xi'\})$ cannot be $\{\xi\}$ nor $\{\xi'\}$, i.e. $C(\{\xi, y, \xi'\}) = \{y\}$. Hence we can assume that z is different from ξ', and we know that $C(\{\xi, z, \xi'\}) = \{z\}$, by definition of ξ, ξ'. This leads to a contradiction with the first statement of Lemma 7, since $C(\{\xi, y, z\}) = \{\xi\}$.

Case 3: $\{x, x'\} \cap \{\xi, \xi'\} = \emptyset$. Suppose to the contrary of what we want to prove that $C(\{\xi, y, z\}) = \{\xi\}$, for some $z \in Y \setminus \{\xi, y\}$. If $C(\{x, x', \xi\}) = \{\xi\}$, then we reach a contradiction with $C(\{\xi, x, \xi'\}) = \{x\}$ and $C(\{\xi, x', \xi'\}) = \{x'\}$, via the first statement of Lemma 7. Hence $C(\{x, x', \xi\}) = \{x\}$ or $\{x'\}$. We consider only the first case, the argument for the second case being similar. The third statement of Lemma 7 implies $C(\{x, y, \xi\}) = \{x\}$, since $C(\{x, y, x'\}) = \{y\}$. Hence $C(\{\xi, y, \xi'\}) \neq \{\xi\}$, as otherwise one would get a contradiction with the second statement of Lemma 7 (with x being "in between" both y and ξ, and ξ and ξ', while ξ is "in between" y and ξ'). So $z = \xi'$ is impossible. If $z \neq \xi'$, then $C(\{\xi, z, \xi'\}) = \{z\}$. Once combined with $C(\{\xi, y, z\}) = \{\xi\}$, the first statement of Lemma 7 implies that $C(\{\xi, y, \xi'\}) = \{\xi\}$, a contradiction again. ■

Proof of Lemma 3

Let $z, z' \in B_w$ such that $C(\{z, w, z'\}) = \{w\}$, and let $x \in B_w$. We will be done with the first part of the statement after proving that either $C(\{x, w, z\}) = \{w\}$ or $C(\{x, w, z'\}) = \{w\}$ (meaning that one can actually choose y in $\{z, z'\}$). Notice first that $C(\{x, w, z\})$ must be a singleton, by EC if $C(\{x, z\}) = \{x, z\}$, or by ATT if $C(\{x, z\})$ is a singleton. A similar argument implies that $C(\{x, w, z'\})$ is a singleton as well. Suppose now, on the contrary to what we want to prove, that $C(\{x, w, z\}) \in \{x, z\}$ and $C(\{x, w, z'\}) \in \{x, z'\}$. Notice that we must have $C(\{x, w, z\}) = C(\{x, w, z'\})$, as otherwise we would have a contradiction with Lemma 5 and RA (there is no way to select at most two elements out of $\{w, x, z, z'\}$, that lead to a nonempty intersection with three different singleton choices in three subsets of cardinality 3). Hence it must be that both $C(\{x, w, z\})$ and $C(\{x, w, z'\})$ equal $\{x\}$. It is not difficult to check that this, combined $C(\{z, w, z'\}) = \{w\}$, implies that $C(\{w, x, z, z'\}) = \{x\}$ or $\{w, x\}$, again as a consequence of Lemma 5 and RA. SYM makes the second case impossible. Indeed, w does not belong to neither $C(\{x, w, z\})$, nor $C(\{x, w, z'\})$. So we are forced to conclude that $C(\{w, x, z, z'\}) = \{x\}$, but then we get a contradiction with EXP since $x, z, z' \in B_w$. We are thus done with the proof of the first part of the statement.

As for the second part, let $y, y' \in B_w$ be such that $C(\{x, w, y\}) = \{w\}$ and $C(\{x, w, y'\}) = \{w\}$. Suppose, to the contrary of what we want to prove, that $x \succ_1 y$ and $y' \succ_1 x$. Notice that $C(\{x, y\}) = \{x, y\}$, as otherwise $C(\{x, w, y\}) = \{x\}$ or $\{y\}$, by ATT. Similarly, $C(\{x,y'\}) = \{x,y'\}$. Hence $y \succ_2 x$ and $x \succ_2 y'$. By the induction hypothesis, $C(\{x,y,y'\}) = C^f_{\succ}(\{x,y,y'\})$. Hence $C(\{x,y,y'\}) = \{x\}$. Combining this with $C(\{x,w,y\}) = \{w\}$ and $C(\{x,w,y'\}) = \{w\}$, Lemma 5 and RA imply that $C(\{w,x,y,y'\}) = \{w\}$ or $\{w,x\}$. The second case would lead to a contradiction with SYM, and hence $C(\{w,x,y,y'\}) = \{w\}$, but this leads to a contradiction with EXP, since $C(\{w,x\}) = \{w,x\}$, $C(\{x,y\}) = \{x,y\}$, $C(\{x,y'\}) = \{x,y'\}$, and $C(\{y,y'\}) = \{y,y'\}$. We are thus done with the proof of the second and last part of the statement. ■

Proof of Lemma 4

Lemma 8 *Let C be a bargaining solution that satisfies SYM, RA, EFF, NBC, EC, and EXP. Suppose that the choice out of any pair in $\{x,y,y'\}$ is the pair itself, and that $C(\{x,y,y'\}) = \{x\}$. If $x \in A_w$ and $y,y' \in B_w$, then $C(\{y,y',w\}) = \{w\}$.*

Proof: ATT implies that $C(\{x,y,w\}) = C(\{x,y',w\}) = \{x\}$. Since $C(\{x,y,y'\}) = \{x\}$, it follows from Lemma 5, RA and SYM that $C(\{x,y,y',w\}) = \{x\}$. EC implies that $C(\{y,y',w\})$ is a singleton. If $C(\{y,y',w\}) = \{y\}$, then we get a contradiction with EXP, since $C(\{x,y\}) = \{x,y\}$. By a similar argument, $C(\{y,y',w\}) \neq \{y'\}$, and hence $C(\{y,y',w\}) = \{w\}$. ■

Proof of Lemma 4: Assume, by contradiction, that there exist $\xi,\xi' \in A_w$ and $y,y' \in B_w$ such that $y \succ_2 \xi$ and $y' \succ_1 \xi'$. Hence $C(\{\xi,y\}) \neq \{\xi\}$, by definition of $\succ$ on S. Also, $C(\{\xi,y\}) \neq \{y\}$, as otherwise we would get a contradiction with $y \in B_w$ via PC, since $\xi \in A_w$. Hence $C(\{\xi,y\}) = \{\xi,y\}$. A similar argument implies that $C(\{\xi',y'\}) = \{\xi',y'\}$. By definition of $\succ$ on S, we have:

$$\begin{array}{cc} \xi \succ_1 y & y \succ_2 \xi \\ y' \succ_1 \xi' & \xi' \succ_2 y' \end{array} \tag{2}$$

The proof proceeds by considering two cases.

Case 1 $C(\{\xi,y'\}) = \{\xi\}$ *and* $C(\{\xi',y\}) = \{\xi'\}$

By definition of $\succ$, we have: $\xi \succ y'$ and $\xi' \succ y$. Combining this with (2), it follows that $\xi \succ_1 y' \succ_1 \xi' \succ_1 y$ and $\xi' \succ_2 y \succ_2 \xi \succ_2 y'$. Since $C = C^f_{\succ}$ on triplets in S, we conclude that $C(\{\xi,y,y'\}) = \{\xi\}$ and $C(\{\xi',y,y'\}) = \{\xi'\}$. ATT implies that $C(\{w,\xi,y\}) = C(\{w,\xi,y'\}) = \xi$, and $C(\{w,\xi',y\}) = C(\{w,\xi',y'\}) = \xi'$. SYM, Lemma 5, and RA imply that $C(\{w,\xi,y,y'\}) = \{\xi\}$ and $C(\{w,\xi',y,y'\}) = \{\xi'\}$. This leads to a contradiction with EXP if $C(\{w,y,y'\}) = \{y\}$ or $\{y'\}$, since $y, y' \in B_w$, $C(\{y,y'\}) = \{y,y'\}$, $C(\{\xi,y\}) = \{\xi,y\}$, and $C(\{\xi',y\}) = \{\xi',y\}$. EC implies that $C(\{w,y,y'\})$ is a singleton, and hence $C(\{w,y,y'\}) = \{w\}$, but this contradicts the assumption of Lemma 4. Hence this first case is impossible, and we have to look into the second case.

Case 2 $C(\{\xi,y'\}) \neq \{\xi\}$ *and/or* $C(\{\xi',y\}) \neq \{\xi'\}$.

We consider the case where $C(\{\xi,y'\}) \neq \{\xi\}$. A similar reasoning applies if $C(\{\xi',y\}) \neq \{\xi'\}$. $C(\{\xi,y'\}) = \{y'\}$ would lead to a contradiction with $y' \in B_w$ via PC, since $\xi \in A_w$. Hence $C(\{\xi,y'\}) = \{\xi,y'\}$. If $\xi \succ_1 y'$, then $\xi \succ_1 y' \succ_1 \xi'$ and $\xi' \succ_2 y' \succ_2 \xi$, by (2) and the fact that $C = C^f_\succ$ on pairs in S. Also, $C = C^f_\succ$ on triplets in S, and hence $C(\{\xi,\xi',y'\}) = \{y'\}$. On the other hand, ATT implies that $C(\{w,\xi,y'\}) = \{\xi\}$ and $C(\{w,\xi',y'\}) = \{\xi'\}$. There is no way of defining $C(\{w,\xi,\xi',y'\})$ so as to satisfy Lemma 5 and RA. Hence it must be that $y' \succ_1 \xi$. In turn, this implies that $y' \succ_1 \xi \succ_1 y$ and $y \succ_2 \xi \succ_2 y'$, by (2) and the fact that $C = C^f_\succ$ on pairs in S. Also, $C = C^f_\succ$ on triplets in S, and hence $C(\{\xi,y,y'\}) = \{\xi\}$. Lemma 8 implies $C(\{y,y',w\}) = \{w\}$, a contradiction with the assumption of Lemma 4. Case 2 is thus impossible as well. ■

$\succ_1^*$ and $\succ_2^*$ are transitive

Transitivity is the subject of Lemmas 10 and 11. Before stating and proving them, we need to establish a useful property.

Lemma 9 *Let C be a bargaining solution that satisfies SYM, RA, EFF, ATT, NBC, EC, and EXP. Let x, y, z, z' be four elements of X such that the solution out of any pair in $\{x,y,z\}$ is the pair itself, $C(\{y,z'\}) = \{y,z'\}$, and $C(\{z,z'\}) = \{z'\}$. Then $C(\{x,y,z\}) = \{y\}$ if and only if $C(\{x,y,z'\}) = \{y\}$.*

Proof: Notice that $C(\{x,z'\}) \neq \{x\}$, as otherwise we would get a contradiction with $C(\{x,z\}) = \{x,z\}$ via PC, since $C(\{z,z'\}) = \{z'\}$. Independently of whether $C(\{x,z'\}) = \{z'\}$ or $\{x,z'\}$, ATT implies that $C(\{x,z,z'\}) = C(\{y,z,z'\}) = \{z'\}$.

If $C(\{x,y,z\}) = \{y\}$, then Lemma 5 and RA imply that $C(\{x,y,z,z'\}) = \{z'\}$ or $\{y,z'\}$. The former case leads to a contradiction with EXP. In the latter case, SYM implies that $z' \notin C(\{x,y,z'\})$, since $C(\{y,z,z'\}) = \{z'\}$. $C(\{x,z'\}) = \{z'\}$ would imply $C(\{x,y,z'\}) = \{z'\}$, by ATT, a contradiction. Hence $C(\{x,z'\}) = \{x,z'\}$, and EC implies that $C(\{x,y,z'\})$ must be a singleton, or $C(\{x,y,z'\}) = \{y\}$ given RA, as desired.

If $C(\{x, y, z'\}) = \{y\}$, then Lemma 5 and RA imply that $C(\{x, y, z, z'\}) = \{y, z'\}$. Lemma 6 implies in turn that $C(\{x, y, z\}) = \{y\}$, as desired. ■

Lemma 10 *Let $(\succ_1, \succ_2)$ be two complete, transitive and anti-reflexive orderings defined over $S \subseteq X$ such that $C = C^f_\succ$ on pairs and triplets in S, let $w \in X \setminus S$, let $(\succ^*_1, \succ^*_2)$ be the extensions of $(\succ_1, \succ_2)$, as defined in the main text, let x, y be two elements of S, and let $i \in \{1, 2\}$. If $x \succ_i y$ and $y \succ^*_i w$, then $x \succ^*_i w$. Similarly, if $w \succ^*_i y$ and $y \succ_i x$, then $w \succ^*_i x$.*

Proof: The second statement being symmetric to the first, its proof is very similar and is therefore omitted. We are thus assuming that $x \succ_i y$ and $y \succ^*_i w$, and we want to prove that $x \succ^*_i w$. If $x \in A_w$, then we are done. So we'll assume $x \in B_w$.

Suppose that there is no $z, z' \in B_w$ such that $C(\{z, w, z'\}) = \{w\}$. If $y \in A_w$, then $x \succ^*_i w$, by definition of $\succ^*_i$. Suppose now that $y \in B_w$. Our construction of $\succ^*$ is such that either $z \succ^*_i w$ for all $z \in B_w$, or $w \succ^*_i z$ for all $z \in B_w$. Hence $x \succ^*_i w$, as desired. So from now on we assume that there exist $z, z' \in B_w$ such that $C(\{z, w, z'\}) = \{w\}$.

By Lemma 3, there exists $x' \in B_w$ such that $C(\{x, w, x'\}) = \{w\}$. If $x \succ_i x'$, then $x \succ^*_i w$, by construction, and we are done. So we prove in the remainder that $x' \succ_i x$ is impossible. So we will assume, on the contrary, that $x' \succ_i x$ and $x \succ_{-i} x'$.

Suppose first that $y \in A_w$. In that case, $C(\{x, y\})$ is different from $\{x\}$, as otherwise we get a contradiction with $x \in B_w$ via PC. $C(\{x, y\})$ is also different from $\{y\}$, since $x \succ_i y$, and $C = C^f_\succ$ on all pairs in S. Hence $C(\{x, y\}) = \{x, y\}$. Since $C = C^f_\succ$ on all pairs in S, we conclude $y \succ_{-i} x$. Given that $x' \succ_i x$ and $x \succ_{-i} x'$, the transitivity of $\succ$ implies that $x' \succ_i y$ and $y \succ_{-i} x'$. Since $C = C^f_\succ$ on all pairs in S, $C(\{x', y\}) = \{x', y\}$. Given that $y \in A_w$, ATT now implies that $C(\{x, y, w\}) = C(\{x', y, w\}) = \{y\}$. Since $C = C^f_\succ$ on all triplets in S, it follows that $C(\{x', x, y\}) = \{x\}$. But because $C(\{x, w, x'\}) = \{w\}$, there is no way to define $C(\{x, x', y, w\})$ so as to satisfy RA, given Lemma 5, and we get the desired contradiction.

Suppose next that $y \in B_w$. Then, it follows from $y \succ^*_i w$ that $w \succ^*_{-i} y$, by construction. If $C(\{x', y\}) = \{x'\}$, then $x' \succ y$, by construction, and hence $x \succ y$ (by assumption for i and by transitivity for $-i$). Since $C = C^f_\succ$ on pairs in S, we conclude that $C(\{x, y\}) = \{x\}$. ATT implies that $C(\{x, y, w\}) = \{x\}$ and $C(\{x', y, w\}) = \{x'\}$. It becomes impossible to define $C(\{x, x', y, w\})$ so as to satisfy RA and Lemma 5, given that $C(\{x, x', w\}) = \{w\}$. So we must conclude that $C(\{x', y\}) \neq \{x'\}$, and hence $C(\{x', y\}) = \{x', y\}$ since $x' \succ_i y$ (this follows from our assumptions that $x \succ_i y$ and $x' \succ_i x$, and from the transitivity of $\succ$). If $C(\{x, y\}) = \{x\}$, then Lemma 9 implies that

$C(\{x', y, w\}) = \{w\}$, and we get a contradiction with $y \succ_i^* w$, since $x' \succ_i y$ (see Lemma 3). As in the previous paragraph, we cannot have $C(\{x, y\}) = \{y\}$ either, because $x \succ_i y$. Hence $C(\{x, y\}) = \{x, y\}$. So $x' \succ_i x \succ_i y$ and $y \succ_{-i} x \succ_{-i} x'$, and $C(\{x, x', y\}) = \{x\}$ since $C = C_{\succ}^f$ on triplets in S. In addition, we also know that $C(\{x', w, x\}) = \{w\}$. Since $x', y \in B_w$ and $C(\{x', y\}) = \{x', y\}$, then $C(\{x', w, y\})$ must be a singleton, by EC. If $C(\{x', w, y\}) \in \{x', y\}$, then there is no way of defining $C(\{x, x', y, w\})$ so as to satisfy Lemma 5 and RA. Hence, $C(\{x', w, y\}) = \{w\}$, and we get a contradiction with $y \succ_i^* w$, since $x' \succ_i y$ (see Lemma 3). ■

Lemma 11 *Let $(\succ_1, \succ_2)$ be two complete, transitive and anti-reflexive orderings defined over $S \subseteq X$ such that $C = C_{\succ}^f$ on pairs and triplets in S, let $w \in X \setminus S$, let $(\succ_1^*, \succ_2^*)$ be the extensions of $(\succ_1, \succ_2)$, as defined in the main text, let x, y be two elements of S, and let $i \in \{1, 2\}$. If $x \succ_i^* w$ and $w \succ_i^* y$, then $x \succ_i y$.*

Proof: We wish to show that $x \succ_i y$. If $C(\{x, y\}) = \{x\}$, then we are done. Assume $C(\{x, y\}) \neq \{x\}$.

We first consider the case where $x \in A_w$. Hence $C(\{x, y\}) \neq \{y\}$, or $C(\{x, y\}) = \{x, y\}$, since otherwise we get a contradiction with $w \succ_i^* y$ via PC. Now assume that the conclusion of the lemma is wrong, i.e. $y \succ_i x$. Notice that there must exist $y' \in B_w$ such that $C(\{y, w, y'\}) = \{w\}$, as otherwise $y \succ_i^* w$, by definition of $\succ^*$, a contradiction. Since $w \succ_i^* y$, it must be that $y' \succ_i y$ and $y \succ_{-i} y'$, again by definition of $\succ^*$. Since $y \succ_i x$, $x \succ_{-i} y$, and $C = C_{\succ}^f$ on triplets in S, it follows that $C(\{x, y, y'\}) = \{y\}$. Given that w is added after y in our induction, it cannot be that $C(\{w, y\}) = \{w\}$. Since $w \succ_i^* y$, it cannot be that $C(\{w, y\}) = \{y\}$ either. Hence $y \in B_w$. ATT implies that $C(\{x, w, y\}) = \{x\}$, but then there is no way of defining $C(\{x, y, y', w\})$ so as to satisfy Lemma 5 and RA. We, therefore, conclude that $x \succ_i y$, as desired.

Consider next the case where $x \in B_w$. As in the previous paragraph, $y \in B_w$. By our construction of $\succ^*$, there must exist $x', y' \in B_w$ such that $C(\{x, w, x'\}) = \{w\}$ and $C(\{y, w, y'\}) = \{w\}$. If this was not true, then w would be ranked above or below both x and y according to $\succ_i^*$, thereby contradicting our assumption that $x \succ_i^* w$ and $w \succ_i^* y$.

Suppose that $C(\{x, y\}) = \{y\}$. Lemma 9 implies that $C(\{x', y, w\}) = \{w\}$. Since $w \succ_i^* y$, we must have $x' \succ_i y$. We must also have $x \succ_i x'$, since $C(\{x, x', w\}) = \{w\}$ and $x \succ_i^* w$. Transitivity of $\succ_i$ implies that $x \succ_i y$, as desired.

Suppose now that $C(\{x, y\}) = \{x, y\}$, and that $y \succ_i x$, contrarily to what we want to prove. Then $y' \succ_i y \succ_i x \succ_i x'$ and $x' \succ_{-i} x \succ_{-i} y \succ_{-i} y'$ in order to have $x \succ_i^* w$ and $w \succ_i^* y$. The solution out of any pair in $\{x, y, w\}$ is the pair itself. So $C(\{x, y, w\})$

is a singleton, by EC. It cannot be w, as this would imply $w \succ_i^* x$. Suppose that $C(\{x, y, w\}) = \{y\}$. Since $C(\{y, w, y'\} = \{w\}$, the first statement of Lemma 7 implies that $C(x, w, y') = \{w\}$, hence a contradiction with $x \succ_i^* w$, since $y' \succ_i x$. Suppose now that $C(\{x, y, w\}) = \{x\}$. Since $C(\{x, w, x'\} = \{w\}$, the first statement of Lemma 7 implies that $C(\{x', y, w\}) = \{w\}$, hence a contradiction with $w \succ_i^* y$, since $y \succ_i x'$. ■

References

Ariely, Dan. 2008. *Predictably Irrational: The Hidden Forces That Shape Our Decisions.* Harper Collins, NY.

Anbarci, Nejat. 1993. "Noncooperative Foundation of the Area Monotonic Solution." *Quarterly Journal of Economics*, 108: 245-258.

Anbarci, Nejat. 2006. "Finite Alternating-Move Arbitration Schemes and the Equal Area Solution." *Theory and Decision*, 61: 21-50.

Ambrus, Attila and Kareen Rosen. "Revealed Conflicting Preferences: Rationalizing Choice with Multi-Self Models." Mimeo.

Benhabib, Jess and Alberto Bisin. 2004. "Modeling Internal Commitment Mechanisms and Self-Control: A Neuroeconomics Approach to Consumption-Saving Decisions." *Games and Economic Behavior*, 52(2): 460-92.

Bloom, David E. and Christopher L. Cavanagh. 1986. "An Analysis of the Selection of Arbitrators." *American Economic Review*, 76(3): 408-422.

Brams, Steven J. and D. Marc Kilgour. 2001. "Fallback Bargaining." *Group Decision and Negotiation*, 10(4): 287-316.

Chambers, Christopher P. and Federico Echenique. 2008. "The Core Matchings of Markets with Transfers." Mimeo.

Cerepanov, Vadim, Timothy Feddersen and Alvaro Sandroni. 2008. "Rationalization." Mimeo.

Chiappori, Pierre-André. 1988. "Rational Household Labor Supply." *Econometrica,* 56: 63-89.

Chiappori, Pierre-André and Olivier Donni. 2005. "Learning From a Piece of Pie : The Empirical Content of Nash Bargaining." Mimeo.

Chiappori, Pierre-André and Ivar Ekeland. Forthcoming. "The Micro Economics of Efficient Group Behavior: Identification." *Econometrica.*

Eliaz, Kfir, Michael Richter and Ariel Rubinstein. 2009. "An Étude in Choice Theory: Choosing the Two Finalists." Mimeo.

Eliaz, Kfir and Ran Spiegler. 2006. "Contracting with Diversely Naïve Agents." *Review of Economic Studies*, 73(3): 689-714.

Fudenberg, Drew and David K. Levine. 2006. "A Dual-Self Model of Impulse Control," *American Economic Review,* 96: 1449-76.

Green, Jerry R. and Daniel A. Hojman. 2007. "Choice, Rationality and Welfare Measurement." Mimeo.

Huber, Joel, John W. Payne, and Christopher Puto. 1982. "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis." *Journal of Consumer Research,* 9(1): 90-98.

Hurwicz, Leonid and Murat R. Sertel. 1997. "Designing Mechanisms, in particular for Electoral Systems: The Majoritarian Compromise." Mimeo.

Kamenica, Emir. 2008. "Contextual Inference in Markets: On the Informational Content of Product Lines." *American Economic Review*, 98(5): 2127-2149.

Kalai, Gil, Ariel Rubinstein and Ran Spiegler. 2002. "Rationalizing Choice Functions By Multiple Rationales." *Econometrica,* 70(6): 2481-2488.

Karni, Edi and David Schmeidler. 1975. "Independence of Nonfeasible Alternatives, and Independence of Nonoptimal Alternatives." *Journal of Economic Theory*, 12: 488-493.

Kibris, Özgur and Murat R. Sertel. 2007. "Bargaining Over a Finite Set of Alternatives." *Social Choice and Welfare*, 28: 421-437.

Kivetz, Ran, Oded Netzer, and V. Srinivasan. 2004. "Alternative Models for Capturing the Compromise Effect." *Journal of Marketing Research,* 41(3): 237-57.

Manzini, Paula and Marco Mariotti. 2007. "Sequentially Rationalizable Choice." *American Economic Review,* 95(5): 1824-39.

Manzini, Paula and Marco Mariotti. 2008. "Categorize Then Choose: Boundedly Rational Choice and Welfare." Mimeo.

Mariotti, Marco. 1998. "Nash Bargaining Theory Theory When the Number of Alternatives Can Be Finite." *Social Choice and Welfare,* 15: 413-421.

Masatlioglu, Yusufcan, Daisuke Nakajima, and Erkut Y. Ozbay. 2009. "Revealed Attention." Mimeo.

Ok, Efe A., Pietro Ortoleva, and Gil Riella. 2008. "Rational Choice with Endogenous Reference Points." Mimeo.

Shafir, Eldar, Itamar Simonson, and Amos Tversky. 1993. "Reason-based Choice." *Cognition,* 49: 11-36.

Simonson, Itamar. 1989. "Choice Based on Reasons: The Case for the Attraction and Compromise Effects." *Journal of Consumer Research,* 16(2): 158-74.

Sprumont, Yves. 1993. "Intermediate Preferences and Rawlsian Arbitration Rules." *Social Choice and Welfare,* 10(2): 1-15.

Sprumont, Yves. 2000. "On the Testable Implications of Collective Choice Theories." *Journal of Economic Theory,* 93(2): 205-32.

Tversky, Amos and Eldar Shafir. 1992. "Choice Under conflict: The Dynamics of Deferred Decision." *Psychological Science,* 3(6): 358-61.