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

Geoffroy de Clippel* and Kfir Eliaz†

February 2009

PRELIMINARY AND INCOMPLETE

Abstract

The experimental literature in economics, psychology and marketing provides overwhelming evidence that individuals exhibit systematic departures from rational choice. Among the most important and robust phenomena 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. To lay out our proposal, we first characterize the unique two-self bargaining solution that - for a given pair of preference relations (one for each self) - exhibits both the attraction and the compromise effects, in addition to a number of consistency requirements. We then investigate the question of revealed preferences. First, we provide a set of properties that characterize those choice correspondences that coincide with our bargaining solution, for some pair of preference relations. Second, we characterize the extent to which these two preference relations can be uniquely identified.

Our analysis admits an alternative interpretation if one thinks of two distinct persons 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, and Kant-Rawls social compromise.

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

1. INTRODUCTION

Many of the decision problems we face consist of options with multiple attributes, which require us to trade-off different dimensions or criteria for ranking the options. In choosing an apartment or a house, one often needs to take into account the price, the size, the proximity to work, the quality of education, among other features. Buying a car also entails balancing several different dimensions: price, safety, gas efficiency, size, color, esthetics. Deciding between academic job offers requires us to consider attributes such as the department's faculty, the salary, and the location. In selecting an investment, one needs to take into account the expected annual returns, the fees, and the level of risk. Many times the available options have many different attributes, 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 attributes or dimensions, which we deem important. However, we still need to resolve the trade-off between the different dimensions.

The rationality assumption in economics assumes that individuals are able to consistently make the necessary trade-offs across the different dimensions. This stands in stark contrast to a vast literature in economics, psychology and marketing, which provides overwhelming evidence that individuals exhibit systematic departures from rational choice. In particular, numerous studies suggest that individuals often find it difficult to resolve the conflict about how much of one attribute to trade off in favor of another. Among the most studied and robust violations of rationality are the *attraction* and the *compromise* effects.

The attraction effect was first demonstrated by Huber, Payne and Puto (1989), 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. Both effects can be described by considering a choice from a pair of two-dimensional options, A and B, such that A is better than B along the first dimension while B is better than A along the second dimension (e.g., a holiday package involving 7 days in Paris and 4 days in London versus 4 days in Paris and 7 days in London). The above studies demonstrated that the choice probability of either option, say A, rises significantly in the presence

[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).

of a third alternative, which either (*i*) is inferior to A (e.g., 6 days in Paris and 3 in London), or (*ii*) is more extreme than A in the sense of being even better than A on the first dimension but being even worse on the second dimension (e.g., 9 days in Paris and 2 days in London). These studies suggest two systematic ways of violating the Weak Axiom of Revealed Preferences (WARP).

This paper proposes and characterizes a choice procedure that generates both the attraction and the compromise effects. Our choice procedure is motivated by the interpretation of these two effects 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)).[2] 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, Eliaz and Spiegler (2006), Fudenberg and Levine (2006), Benhabib and Bisin (2004)), we focus our analysis on decision problems that give rise to two selves. Note that in contrast to most of the multi-selves literature, we envision the two selves trying to reach a compromise rather than being engaged in a non-cooperative game.

We first consider an environment in which there are two clear criteria or dimensions along which options can be ranked. Moreover, it is clear how the available options may be ranked along these dimensions. Examples of decision problems of this type include 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. We model this environment as consisting 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 (observable) 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.

To be able to reasonably interpret $C_\succ$ as a bargaining solution, it should satisfy some basic set of "appealing" properties. First, we would expect $C_\succ$ to be *neutral* (NEUT)

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

with respect to the names of alternatives, *independent of the preferences over unavailable alternatives* (IPUA), and *Pareto efficient* (EFF). Second, we might also expect $C_{\succ}$ to exhibit some minimal consistency across bargaining problems. In particular, removing an alternative from a set should not cause all the previously chosen alternatives to be rejected - unless the removed alternative was the only chosen alternative (property RA). We also require the bargaining solution to be "symmetric" with respect to the bargainers and with respect to the options. By this we mean that (i) exchanging the bargainers' preferences should not affect the solution (*exchangeability,* EX), and (ii) if both x and y are chosen from a given set, but x is not chosen when we remove a third element z, then it must be that y is not chosen when we remove some other element z' (*symmetry,* SYM).

The question is, what bargaining solutions, having the above properties, exhibit both attraction and compromise effects? A bargaining solution is said to exhibit an attraction effect (ATT) if whenever we add an alternative that is Pareto dominated by some chosen element in the original set, 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" 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 in between x and y.

Our first main result (Theorem 1) establishes the existence of a unique bargaining solution that satisfies all of the above properties. 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 can be implemented in subgame perfect equilibrium by a simple sequential mechanism: at each round, both bargainers simultaneously decide on an option to "remove from the table", and they continue to do so until all options have been removed. The last option or options to be removed are then selected. It is interesting to note that this bargaining solution has been previously discussed in the literature under the name of "fallback bargaining" (Brams and Kilgour (2001)), "unanimity compromise" (Kibris and Sertel (2007)) as well as the "Kant-Rawls Social Compromise" (Hurwicz and Sertel (1997)). 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).

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, serving size, etc.). 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. First, we rephrase the NBC property to account for the assumption of unobservable preferences: if both x and y are chosen from a set, then there is no z in this set that is uniquely chosen from the triplet $\{x, y, z\}$. It is straightforward to adapt EFF, ATT, RA and SYM to the new environment (note that NEUT, EX and IPUA only make sense when preferences are observable). Since the preferences of the two selves are not given to us, we need to add a condition that rules out cycles in binary choices (a property we call pairwise transitivity, PT). In addition, we need to guarantee the existence of a compromise (EC): a single element is chosen from any triplet in which no single element dominates another in pairwise choices. However, these conditions are not enough to deliver our result. We also need to impose an additional consistency requirement on revealed compromises (a property we call *Expansion*, EXP): if x is the unique choice from a set S, and this choice is a "compromise" (i.e., the choice out of any pair is the pair itself), then no new element can be chosen uniquely from the new set unless it is chosen over x in a pairwise comparison. Theorem 2 then states that when the choice data is rich enough, there exists a pair of linear orderings over X such that a choice correspondence coincides with the fallback solution if and only if it satisfies the above properties.

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 standard approach is to say that x is revealed preferred to y if x is chosen in the presence of y in a pairwise comparison. The challenge is then to show that this definition gives rise to a complete, transitive binary relation. In our framework, we may interpret a unique choice of x over y in a pairwise comparison as revealed Pareto dominance: both selves rank x above y. 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.[3] We overcome this difficulty by constructing an induction argument in which the elements of X are added in a particular order. This induction argument proves to be very useful in showing that our construction of the selves' orderings is well-defined, and also in addressing the question of "identifiability", which we discuss next.

A natural question that arises is, 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 choices. Put differently, we are able to identify a pair of rankings that are consistent with the choices, but we cannot identify which ranking belongs to which self. We argue that there is a sense in which any further multiplicity is with respect to "irrelevant alternatives". By this we mean that we can partition the set of elements into a list Π of ordered cells such that among the elements that are not in the first $k-1$ cells, the k-th cell is the minimal (non-empty) set of elements, which beat every element outside this cell in pairwise comparisons. Theorem 3 establishes that if the fallback solution generates the same choices for two distinct preference profiles, then each profile can be obtained from the other by permuting the two orderings over cells of Π that contains at least two elements. This means that for any given choice problem S, we can pin down the pair of preferences over the set of options that Pareto dominate any option outside the 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

[3]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.

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 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. Our first main result provides a new characterization of this bargaining solution, which in contrast to previous characterizations, fixes the profile of preferences and considers the consistency of the solution across bargaining problems. 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. To the best of our knowledge, we are the first to apply a revealed-preference approach to bargaining solution concepts. In future research we plan to extend this approach to cardinal bargaining solutions, of which the most well known is the Nash bargaining solution.

The rest of the paper is organized as follows. The related listerature 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 the testable implications of this model. 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 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 the necessary and sufficient conditions for deriving $(r, \mathbf{u}, u)$ from choice data such that this data will coincide with the outcome of the above bargaining model. One of these conditions, labeled "reference-dependent WARP", rules out the compromise effect.[4]

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.[5] 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.

Related to the above literature is the recent attempt to provide testable implica-

[4]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.

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

tions to decision heuristics that generate systematic departures from rational behavior. Most notably, 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)). The authors show that a choice function is equivalent to shortlisting if and only if it satisfies two properties dubbed "weak WARP" and "Expansion". Our choice procedure satisfies the first property but not the second.[6]

Rationalization by multiple rationales

Our work is also related to a few recent papers that model individual choice as the aggregation of multiple rationales. However, these papers focus on a different set of questions than we do. 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.

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 methodolgy . 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

[6]Weak WARP means that if an alternative x is chosen uniquely from the pair $\{x, y\}$ and also when y and some set of other alternatives S are available, then y is not chosen when x and a subset $T \subset S$ are available. Our choice procedure satisfies this property since it never selects alternatives that are rejected in pairwise choices. Expansion means that an alternative chosen from each of two sets is also chosen from their union. Since out bargaining solution exhibits the compromise effect it violates this property.

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 wheret 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, which 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[7] 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 (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 each integer between

[7]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.

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 bargainers agree on what the best option is, then it is the solution. Otherwise one has to look for option(s) that would be ranked either top or second-best by both bargainers. If no option satisfies this property, then one iterates the procedure by allowing for third-best alternatives, and so on so forth. It is straightforward to check that the outcome of this procedure (one or two elements of S) coincides with the fallback bargaining solution as defined in the previous paragraph. A dual interpretation of this procedure offers a natural implementation of the fallback bargaining solution in subgame-perfect Nash equilibrium, in the spirit of the Nash program. Starting with the whole set S of available options, the two bargainers alternate to propose one option (among those that remain available) to take off the table. The game ends with the single option left being enforced. If $C^f_{\succ}(S)$ is a singleton, then this is the unique subgame-perfect equilibrium of the procedure, independently of which bargainer makes the first proposal. In the other case where $C^f_{\succ}(S)$ contains two elements, the procedure admits a unique subgame-perfect equilibrium outcome independently of who makes the first proposal. There is a first-mover advantage when $C^f_{\succ}(S)$ contains two elements, and the element that is most favorable to i is the unique subgame-perfect equilibrium outcome when the game starts with i. Anecdotically, the fallback bargaining solution is also related to some classical concepts developed in social choice theory. Indeed the idea of defining scores as we do has already been proposed by Borda in the 18th century when defining a possible alternative to Condorcet's pairwise comparisons. A major difference is that Borda proposed to select those options that maximize the sum of the scores, while the fallback solution aims at maximizing the minimal score. In other words, Borda applied a utilitarian criterion to a canonical representation of the ordinal preferences, while the fallback solution applies an egalitarian/Rawlsian criterion. Notice that the Borda criterion would often be useless in our problem with only two bargainers. For instance, any element of S would be part of the solution if all the elements of S are Pareto optimal for the pair of orderings $(\succ_1, \succ_2)$ (moving from one option to the next decreases the score by one point in one dimension, and increases it by one point in the other, thereby keeping the sum constant). Particularly, the Borda solution cannot capture the compromise effect. Finally it is worth noting that the Borda rule has been criticized on the ground that its outcome depends crucially on the fact that one adds exactly one point when moving from one option up to the next in an agent's ordering. This sounds arbitrary. One can construct many other Borda-type rules (known as "scoring rules") by taking other canonical representations of

the participants' ordinal preferences, using any arbitrary function that is increasing in the Borda scores. An interesting feature of the fallback solution is that it is invariant to such alternative ways of computing scores (provided of course one uses the same method on both dimensions, which is indeed always assumed when defining scoring rules in order to preserve some form of anonymity).

4. PREFERENCE-BASED AXIOMATIC CHARACTERIZATION

The following axioms are imposed on a preference-based bargaining solution C, and will be assumed to be valid for each $\succ \in L(X) \times L(X)$, and each $S \subseteq X$. We start with the axioms that are most original to our analysis.

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

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 $y \succ_{-i} z \succ_{-i} x$.*

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

RA and NBC are directly inspired from, and generalize, some "irrational" choice patterns identified in simple experiments (cf. Introduction). ATT formalizes the idea that adding a dominated alternative reinforces the appeal of an option to the decision maker. Thus, the solution to the enlarged problem obtained by adding x as an option should be the intersection of the solution to the original problem with those options that Pareto dominate x whenever that set is nonempty. In NBC, the alternative z 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. In such cases, decision makers that are subject to the compromise effect would not be enclined to select both x and y as good choices without including z as well. Both ATT and NBC are typically incompatible with the property of independence of irrelevant alternatives that characterizes rational choice methods. We propose to restrict attention to bargaining solutions that satisfy a weaker consistency property, RA. 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 thinks of eliminating 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 to being a simple consistency propert, 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 move the set of selected elements too far away (i.e. nonempty intersection). In addition to these three properties that capture the essence of our analysis, we add a few regularity axioms. It is certainly worthwile to characterize the classes of solutions that would satisfy ATT, NBC, and RA, together with weaker alternatives of the next five axioms. Yet we feel that these five properties define a natural benchmark that needs to be investigated first.

Efficiency (EFF) - *If* $x \in C_{\succ}(S)$*, then there does not exist* $y \in S$ *such that* $y \succ x$*.*[8]

Neutrality (NEUT) - *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* $i \in \{1, 2\}$*.*

Exchangeability (EX) - $C_{(\succ_2,\succ_1)}(S) = C_{(\succ_1,\succ_2)}(S)$*.*

Independence with respect to Preferences over Unavailable Alternatives (IPUA) - *Let* $\succ'$ *be an alternative pair of linear orderings on* X *that coincide with* $\succ$ *on* $S \times S$*. Then* $C_{\succ'}(S) = C_{\succ}(S)$*.*

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* $y \notin C_{\succ}(S \setminus \{z'\})$*.*

EFF assumes that the decision maker will not choose an option that is dominated by an alternative along both dimensions. NEUT means that the label of the available options does not influence the choice. Thus we rule out at the outset the potential effect of some forms of framing. EX means that the bargaining solution is invariant to renaming of the underlying characteristics (the first characteristic becoming the second, and vice versa). This amounts to assume that both characteristics are equally relevant. IPUA is a rather weak property that has been used repeatedly in bargaining and social choice (first mentioned explicitly in[9] Karni and Schmeidler (1975)). It requires that the solution to a problem S does not depend on the bargainers' preferences over alternatives that are not available (i.e. in $X \setminus S$). As for SYM, it seems reasonable to assume that the options that are selected by the decision maker are equally appealing to him. We feel that this

[8] $\succ$ 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.

[9] Karni and Schmeidler themselves refer to a 1969 *mimeo* written by A. Gibbard.

general property would be violated if x, y where both selected in S, while the decision maker looses interest in x when an option z is dropped, but never looses interest in y when dropping any single alternative from S.

In order to prove the main result, we need the following inductive characterization of the fallback bargaining solution, whose proof is available in 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 *A preference-based bargaining solution C satisfies ATT, NBC, RA, EFF, NEUT, EX, IPUA, and SYM if and only if $C = C^f$.*

Proof: We first check that C^f satisfies the nine axioms. *[text to be inserted here - details available from the authors upon request]*

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. ■

Independence of the axioms. *[text to be inserted here - details available from the authors upon request]* Comparison with Kibris and Sertel's axiomatic result (2007), and perhaps reference to axiomatic characterizations of scoring rules (Young, Myerson + ?). *[text to be inserted here]*

5. REVEALED PREFERENCES

Suppose that a decision maker has to select options that are characterized by two dimensions as in the previous section, but that we do not know what these preferences are (e.g. color, shape, fabric, etc.). We want to address two related questions: 1) Which characteristic property will the decision maker's choices display if he chooses according to the fallback bargaining solutions give his preferences over the two dimensions?, and 2) Assuming that the decision maker behaves according to the fallback bargaining solution, can we identify (at least partially, and, if so, to what extent) his preferences over the two dimensions. More generally, the options do not need to be characterized by two observable dimensions. One may wonder under which conditions on the observable choices can we conclude that the decision maker behaves as if she is compromising between two conflicting orderings.

Looking back at the previous section, we see that the properties of IPUA and EX are meaningless when studying bargaining solutions instead of their preference-based variants. NEUT was useful in combination with these two properties, and will therefore not serve any purpose in the present section either. SYM and RA can be rephrased literally:

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* $y \notin C(S \setminus \{z'\})$*.*

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

Now comes the interesting exercise of thinking about what could be identified by observing choices. A first relation that comes to mind is the Pareto relation, that is, $C(\{x, y\} = \{x\}$ is interpreted as the revelation that x is better than y in both dimensions. 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. Of course, the revealed Pareto relation should not display cycles - a usual condition when discussing revealed preferences:

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

PC was automatically satisfied in the preference-based model of the previous section, since $\succ$ was assumed to be transitive. We are now ready to rephrase EFF and ATT:

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 $C(S \cup \{x\}) = \{y \in C(S) | C(\{x,y\}) = \{y\}\}$.*

Of course, the fallback bargaining solution also captures another feature that is more original: if the pair itself is chosen out of any pair in $\{x,y,z\}$, and x is chosen out of $\{x,y,z\}$, then one should conclude that x is a good compromise, which means that x falls"in between" y and z. So we can rephrase NBR:

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 $C(\{x,y,z\}) = \{z\}$.*

For NBC to be effective, we also need to require that the compromise effect is indeed present:

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.*

Observe that EC was always satisfied in the preference-based approach, since one of the three elements always had to fall in between the other two in the absence of Pareto comparisons.

We have argued so far that SYM, RA, PC, EFF, ATT, NBC, and EC are natural analogue to the axioms used from Section 3 in our new setting. One may wonder whether these axioms would be sufficient to guarantee the existence of a pair of preferences $\succ \in L(X) \times L(X)$ such that $C = C^f_\succ$. It turns out that this is not the case, because the fallback bargaining solution satisfies some natural consistency properties in the way compromises are made that are not captured by the axioms listed so far,[10] as the following two examples show.

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*

[10] PC expresses a consistency property for Pareto comparisons that are revealed when a single option is chosen out of pairs, but remains silent as far as compromises are concerned, those being revealed only when considering triplets as well.

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.*

It is possible to prove our next characterization result by adding axioms that rule out these kinds of inconsistencies. However, to avoid the multiplication of axioms, we will add a single axiom that will prevent all of them at once (as will follow from various lemmas in the Appendix).

Expansion (EXP) - *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.

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

Proof: We first check that $C^f_{\succ}$ satisfies the eight axioms, for each $\succ \in L(X) \times L(X)$. *[text to be inserted here - details available from the authors upon request]*

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 $C = C^f_{\succ}$ if $\succ \in L(X) \times L(X)$ is such that $C(T) = C^f_{\succ}(T)$ 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 $C(T) = C^f_{\succ}(T)$ 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 $k \in \{1, \ldots K\}$, 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 $l \in \{1, \ldots, L_k\}$ and some $k \in \{1, \ldots K\}$. 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.[11]

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 $w \in X \setminus S$. 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

[11] 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.

$x \in S$. 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. ■

Explain why a simpler induction argument would not work *[text to be inserted here]*

Independence of the axioms. *[text to be inserted here - details available from the authors upon request]*

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 $C = C^f_{(\succ_2, \succ_1)}$ (as already observed with the axiom EX). 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 $y \in X \setminus S$.[12] 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 $C_{\succ}^f = C_{\succ'}^f$ 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 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 $C = C_{\succ^*}^f$ (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.

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

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

Often times the available data is limited and does not contain the choices from all possible subsets. The axioms of Theorem 2 have "bite" (i.e., they can potentially be violated) if the data is rich enough. Note that to falsify the standard IIA axiom, it is enough to have choice observations on all subsets of two or three elements. However, in contrast to IIA, our axioms of consistency across subsets (e.g., RA and EXP) focus, for simplicity in the absence of rationality, on marginal changes - removing or adding a single alternative, hence, to guarantee the sufficiency of our axioms we need observations on subsets that differ in only a single alternative. Our current proof of sufficiency relies on knowldge of the full data set (choices from all subsets), and it still an open question what is the minimal data set that guarantees sufficiency.

Theorem 2 is still useful in providing an efficient method for verifying whether a limited data set is consistent with the fallback solution for some preference profile. Moreover, for consistent data sets, Theorem 2 derives the corresponding preferences of the selves/bargainers. Our construction of the preferences relies only on observed choices from triplets. Hence, for any data set that contains all such choices we provide a method of constructing a candidate profile of preferences. Given this profile, we can use the fallback solution to predict the choice from any subset. If all the predictions match the data, we can conclude that the data is consistent with the fallback bargaining solution and the preferences we constructed. Compare this method with 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.

Allowing for indifferences

Throughout our analysis we assumed that the bargainers have strict preferences over X. Allowing for indifferences complicates the analysis. In particular, indifferences may lead to non-singleton compromises. It remains an open question how to extend our analysis to allow for indifference without imposing any structure on X. We conjecture that Indifferences may be easier to handle if X has the product structure, i.e., if $X = X_1 \times X_2$, and if we also assume that individual i has strict ranking of the elements in X_i. That is, suppose each individual i is characterized by a pair of preference relations, a weak preference $\succeq_i$ on X and a strict preference $\succ_i^*$ on X_i such that for any $x, y \in X$, $x \succ_i y$ if $x_i \succ y_i$ and $x \sim y$, otherwise (which means that $x_i = y_i$). For our analysis to go through we need to strengthen the efficiency axiom to strong efficiency and check whether the fallback solution still selects at most two elements.

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 bargining. One conceptual difficulty with this extension is 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. The question is, how do we extend this notion of "betweenness " to more than two bargainers? It is interesting to note 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 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 concerns 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 if we were to replicate the experiments of the attraction and compromise effects with three-dimensional alternatives, the distribution of choices would be roughly $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ on pairs of elements that cannot be Pareto ranked.

"Measurable" alternatives

Our focus on an ordinal bargaining solution captures situations in which individuals are simply unable to trade-off different dimensions or features of alternatives. However, in some situations individuals may be able to compute trade-offs in a limited way, and hence, would take into account preference intensities. For example, when the elements of choice can be mapped onto a two-dimensional grid (recall the holiday packages example from the introduction), individuals may be sensitive to the distance between alternatives. Thus, they 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 individual may be more likely to exhibit the attraction effect when $x = (60, 40)$, $y = (59, 39)$ and $z = (40, 60)$, but (perhaps) not when $y = (41, 39)$. It is therefore interesting to search for a choice procedure that takes into account the sensitivity of individuals to the distance between elements, and to ask whether this sensitivity can be inferred from their choices.

[text to be inserted here - details available from the authors upon request]

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). ■

Proof of Lemma 2: 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 $C(\{y, y', w\}) = \{w\}$.*

<u>Proof</u>: 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 $C(\{y, y', w\}) = \{w\}$. ■

<u>Proof of Lemma 4</u>: 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:

$$\begin{array}{cc} \xi \succ_1 y & y \succ_2 \xi \\ y' \succ_1 \xi' & \xi' \succ_2 y' \end{array} \tag{1}$$

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 (1), 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 (1) 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 (1) 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 $C = C^f_\succ$ 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 $C = C^f_\succ$ 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 $x \succ^*_i w$ 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 $C = C^f_\succ$ 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

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

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.

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.

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 Naive 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.

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.

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. 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.