Implementation and Bounded Rationality*

Geoffroy de Clippel†

This Version: December 2011

Abstract

The paper investigates various aspects of implementation under complete information when the mechanism designer is aware that individuals suffer from biases that lead to violations of IIA, or cannot exclude the possibility of such "irrational" behavior.

Contents

---

*This paper beneficiated from insightful conversations with Kfir Eliaz, Eric Maskin, Kareen Rozen, Rene Saran, and Roberto Serrano. Financial support from the Deutsche Bank through the Institute for Advanced Study is gratefully acknowledged.

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

1 Introduction

Rationality requires individual choices to be explainable by the maximization of a complete transitive preference relation or, equivalently, to satisfy the classical axiom of independence of irrelevant alternatives (IIA). Literatures in psychology, marketing, and behavioral economics have pointed to various violations of this principle. Standard examples include status-quo biases, attraction, compromise and framing effects, temptation and self-control, consideration sets, and choice overload. There has been a recent revival of interest to better understand these choice patterns, in particular through their testable implications.[1] In light of these new models, it is natural to investigate the consequences that these biases may have in traditional problems of interactive decision making. The present paper fulfills this objective by extending implementation theory to situations where individuals may be boundedly rational.

More formally, while standard implementation theory uses types to encode complete and transitive preferences, types in this paper will encode choice functions that may possibly violate IIA. I will restrict attention to problems with complete information, meaning that individuals in the society know each others' types, while the mechanism designer does not. A social choice rule (SCR) associates a set of outcomes to each profile of types. A mechanism is a collection of sets of messages, one such set for each individual, and a function that associates an outcome to each profile of messages. A mechanism implements a SCR if the set of equilibrium outcomes coincides with the set of outcomes prescribed by the SCR, at every type profile. Implementation admits different meanings depending on what is meant by an "equilibrium." I will focus on the two main forms of implementation: in Nash equilibrium and in dominant strategies (see definitions in Section 2).

I will provide necessary, and sufficient conditions for a SCR to be implementable in either sense (see Propositions 1, 1', 8, 9 and 10), thereby extending classical results by Dasgupta et al. (1979) and Maskin (1999),[2] from the special case of rational

[1]See e.g. Kalai et al. (2002), Manzini and Mariotti (2007), Ambrus and Rozen (2009), and de Clippel and Eliaz (2009) for choice functions that result from the combination of multiple conflicting selves, see Masatlioglu and Ok (2005) for choices with a status-quo bias, see Rubinstein and Salant (2006) for choices influenced by the sequence in which options appear, see Rubinstein and Salant (2008) for the role of framing in choices, and Cherepanov et al. (2009), Manzini and Mariotti (2009), Masatlioglu et al. (2009), or Lleras et al. (2010) for choices when individuals do not pay attention to all feasible options.

[2]The paper circulated as a working paper from 1977 and 1998. Surveys on the large literature on implementation theory include Maskin (1985), Moore (1992), Palfrey (1992, 2002), Corchón (1996),

choice functions to implementation problems with any form of individual choice behavior. Identifying these properties is obviously critical to understand the limits of implementation in applications, some of which will be covered in this paper. Perhaps more insightful than their specific mathematical formulation, we will see that intuitive analogues of Maskin monotonicity - based either on Bernheim and Rangel's (2009) extended revealed preferences, or directly on choices - are inadequate for characterizing SCRs that are Nash implementable. Another interesting finding is that the notion of "augmented revelation mechanism," introduced by Mookherjee and Reichelstein's (1990) to characterize Nash-implementability on the rational domain, happens to be useful to characterize implementation in dominant strategies in our framework.

I now describe the main results derived from these abstract conditions. As a starting point, to understand the limits of implementation, I seek to extend classical "impossibility" results. When types encode rational choice functions, any SCR with a rich domain and a range that contains at least three elements is Nash implementable (or in dominant strategies) and Pareto efficient, only if it is dictatorial. I will say that a SCR has a rich domain if the mechanism designer cannot exclude some minimal form of rationality and diversity of tastes over triples of "most-preferred" options in the range of the SCR. This boils down to a simplified version of Kalai et al.'s (1979) free triple condition when considering the special case of rational choice functions. Combining this with Bernheim and Rangel's (2009) extension of the Pareto principle and a range of at least three elements, implementability then implies the existence of a dictator in a weak sense, whereby there exists an individual i with the property that the SCR systematically picks an option for which it is impossible to find an alternative in its range that is unambiguously superior for i, again in the sense of Bernheim and Rangel's (2009) extended revealed preference. Particularly, the SCR picks i's most preferred option whenever he is rational, and Propositions 2 and 11 thus indeed extend classical "impossibility" results. On the other hand, I will also show that it is possible in some cases (see Example 3) to construct implementable SCRs on rich domains for which the option being implemented varies with the type of individuals other than i when i's type is not rational. In that sense, "impossibility" results may be a bit more permissive when considering irrational types.

Even though the set of SCRs that are implementable on rich domains can be more diverse when individuals have irrational types, it remains rather small and contains

Jackson (2001), Maskin and Sjöström (2002), and Serrano (2004).

only highly non-anonymous functions. Focusing on SCRs defined for specific classes of choice biases often leads to domains that aren't rich according to the definition of the previous paragraph. Can we hope for more permissive results in these circumstances? Of particular interest are domains where individuals are subject to specific classes of biases. A first natural candidate are models of choice with psychological filters, categorization or limited attention (see Cherepanov et al. (2009), Manzini and Mariotti (2009), Masatlioglu et al. (2009), and Lleras et al. (2010)). If individuals maximize complete and transitive strict preferences over fixed subsets of actual feasible sets (e.g. restricting attention to options that are rationalizable in Cherepanov et al.'s sense, to specific cuisine when many restaurants are available, to the first three candidates on a ballot, or to the five available rentals that are closest to one's workplace when looking for an apartment to rent, etc.) and underlying preferences are unrestricted, then implementation in either Nash equilibrium or dominant strategies unfortunately turns out to be even less permissive than it was on rich domains. In particular implementable SCR whose range contains at least three elements can vary with the type of at most one individual. Also, if there is no individual who is paying attention to all options when they are all feasible, then Pareto efficiency for the true underlying preferences is incompatible with implementability in either concepts (see Propositions 3 and 12, as well as Corollaries 1 and 2).

On the other hand, more permissive results do indeed exist on other classes of biases. Example 1, for instance, illustrates how a group of individuals with a common long-term goal can better resist temptation if decisions are taken via a mechanism that wisely combines its members' limited willpower. Section 3.4 and 4.4 also provide more permissive results for some class of individual choice procedures defined over lists. Consider for instance Rubinstein and Salant's (2006) "Stop When You Start to Decline" choice procedure: given a set of options appearing in a list (think for instance of a page of google results, or a menu at the restaurant, etc.), the decision maker checks options in that order and continues to do so as long as the next option improves upon the previous according to his preference, and he then picks the last option of this improving sequence. I will show that the SCR that picks the left-most element that would be chosen by some individual, should all options be available to each one of them, is both Nash implementable and implementable in dominant strategies on this class of choice procedures (without restriction whatsoever on the class of underlying preferences used to define them). This SCR is in fact implementable over much larger classes of choice functions. For instance, it would also be implementable if individuals were to make choices by following Rubinstein and Salant's (2006) version of Simon's (1955) "Satisficing" procedure (see Propositions 4 and 13). I also establish a partial uniqueness result (see Proposition 14): the SCR I described is the only one that is anonymous, efficient in Bernheim and Rangel's (2009) sense, and implementable in dominant strategies via a "simple" mechanism.

The Pareto correspondence is probably the most classical example of SCR that is Nash-implementable.[3] I show that Bernheim and Rangel's (2009) extension of the Pareto correspondence to irrational choice functions is *not* Nash-implementable. Instead, I provide in Section 3.5 an alternative definition of efficiency that coincides with Pareto's definition when individual choices are rational, that always selects a non-empty subset of Bernheim and Rangel's concept, and that is Nash implementable on all domains (see Proposition 5). I also establish a strong negative result if the mechanism designer's objective is to implement the set of options that are Pareto efficient for the true underlying preferences in the model of choice with limited attention with a fixed consideration correspondence. Indeed, this turns out to be impossible if individuals do not systematically pay attention to all options (see Proposition 6). Similar results are derived for the core in the house allocation problem (see Proposition 7).

Choices can sometimes be influenced by external conditions. For instance, the meal picked in a cafetaria may vary with the order in which options are displayed, or the level of a person's retirement savings may depend on the level of a default rate. This idea is captured theoretically by the concepts of "frames" (see Rubinstein and Salant (2008) or "anscillary condition" (see Bernheim and Rangel (2009)). As suggested in various examples that have percolated in more popular medias (see the original papers by Camerer et al. (2003) and Thaler and Sunstein (2003), as well as Thaler and Sunstein (2008) for a popular book on the topic), the external conditions can sometimes be chosen by the mechanism designer and, if so, he must do it wisely. Since individuals must know the rules of the mechanism before playing it in the scenario considered so far, frames cannot vary with the individuals' types. All the results established in Sections 3 and 4 remain valid after the frame has been fixed by the mechanism designer. It is well-known that considering games that unfold over time expands the set of social choice functions that are implementable (in subgame-

[3]Given that I restrict attention to single-valued individual choice functions, it is impossible to implement multi-valued SCRs in dominant strategies.

perfect equilibrium). While studying subgame-perfect implementation with bounded rationality goes beyond the scope of the present paper, I conclude the paper by observing that considering dynamic games in the present model brings the added flexibility of allowing frames to vary endogenously (see Section 5).

Related Literature

As should be clear from the description above, the most relevant literatures this work build upon include a) classical papers on implementation under complete information on rational domains (see footnote 2), and b) recent papers that investigate the testable implications of specific choice biases (see footnote 1). Other aspects of bounded rationality have been discussed recently in connection to implementation and mechanism design. I conclude this Introduction by brielfy discussing these contributions, in perspective of the present paper.

The past decade has seen the development of arguments, accompanied by concrete examples, in favor of setting up institutions that help individuals make "better" choices when subject to behavioral biases, while not interfering with their preferences when rational (cf. the notions of "asymmetric paternalism" from Camerer et al. (2003) and "libertarian paternalism" from Thaler and Sustein (2003) and Thaler and Sunstein (2008)). Independently of the normative question of which social choice rules are preferable, one must first understand what are all the social choice rules that a mechanism designer can implement when individuals may be subject to behavioral biases. Bounded rationality in those papers is often understood as a dependance of one's choice function on some external condition that can be chosen by the mechanism designer, subject that will be discussed only in Section 5. Beyond this, a substancial effort is devoted in the present paper to understand which social choice rules are implementable when individuals are boundedly rational in the sense of violating IIA. Many results in what follows do not rely on a specific normative criterion. Notice that being able to implement more social choice functions is always beneficial for a mechanism designer whose objective is to maximize any function defined over final outcomes (e.g. profit). Given the role that Pareto efficiency plays in some classical implementation results over rational domains, I also present results based on two different generalizations of this normative property, either using Bernheim and Rangel's (2009) extended revealed preference, or the original Pareto criterion based on the "true" underlying preference orderings in the model of choice with limited attention.

Various recent models in IO have studied how a monopolistic firm can best deal with customers who are subject to various specific departures of rationality (see Spiegler (2011) for a brillant synthesis of many of these models). Eliaz (2002) studies full implementation in Nash equilibrium that is robust to the presence of a fixed fraction of "faulty" individuals, namely individuals that may behave in any possible way. Ray (2010) and Korpela (2011) independently derived some necessary and conditions for Nash implementability that are related to the work I present in Section 3.1 (see more detailed discussion there). Saran (2011) studies under which conditions over individuals choice correspondences over Savage acts does the revelation principle hold for weak Nash implementation with incomplete information. Glazer and Rubinstein (2011) introduce a mechanism design model in which both the content and framing of the mechanism affect the agents ability to manipulate the information he provides.

The departure of rationality considered in the present paper concerns irregularities in individual choices (specifically violations of IIA). Bounded rationality can also be modeled by a lack of rational expectations when playing the game induced by a mechanism. Papers that started to investigate this other important line of research are Crawford et al. (2009) on level-k implementation in some auction settings, and Cabrales and Serrano (2011a,b) (plus references therein) on the dynamic convergence properties of static mechanisms.

2 Definitions

Let I be the set of individuals, let Θ be the set of possible states, let X be the (finite) set of available options, and let $C_{\theta,i} : P(X) \to X$ be i's choice function when the state is θ, with $C_{\theta,i}(S) \in S$, for all $S \in P(X)$. A *social choice function*[4] is a function $f : \Theta \to X$ that picks an option for each profile of types. A *mechanism* is a profile $((M_i)_{i \in I}, \mu)$ where M_i is the (finite) set of messages available to i, and $\mu : M \to X$ is the outcome function ($M := \times_{i \in I} M_i$).

The concept of Nash equilibrium follows in the standard rational case from the combination of two different principles (see e.g. Preliminary Observation in Aumann and Brandenburger (1995)). First individuals are assumed to choose among feasible options by maximizing a preference ordering. Second each individual is assumed to

[4]I used the term social choice rule generically in the introduction to cover either social choice functions or social choice correspondences (see Section 3.5). From now on I will be more specific and systematically distinguish between single-valued and multi-valued social choice rules.

"know" the other individuals' strategies. Bounded rationality is understood in the present paper as a violation of the first principle. As a natural first step, and to best understand the consequences of such forms of bounded rationality, the assumption of mutual knowledge of the strategies played is kept. To summarize, a profile of strategies forms an (extended) Nash equilibrium if the strategy of each individual leads to an outcome that coincides with his chosen option within the set of outcomes he can generate by varying his own strategy, while others play their equilibrium strategies. Next a strategy will be said to be dominant for an individual if, whenever combined with arbitrary strategies for the others, the resulting outcome coincides with his chosen option within the set of outcomes he can generate by varying his own strategy while the others play this arbitrary strategy profile. As in the rational case, any profile of dominant strategies thus forms a Nash equilibrium. This leads to the following formal definitions in our model.

Let $\theta \in \Theta$. The profile of strategies $(m_i^*(\theta))_{i \in N}$ forms a *Nash equilibrium* of the game induced by the mechanism (M, μ) at θ if

$$\mu(m^*(\theta)) = C_{\theta,i}(\{\mu(m_i, m_{-i}^*(\theta)) | m_i \in M_i\}), \text{ for all } i \in I. \tag{1}$$

A message m_i is a *dominant strategy* for individual i at θ if

$$\mu(m_i, m_{-i}) = C_{\theta,i}(\{\mu(m_i', m_{-i}) | m_i' \in M_i\}), \text{ for all } m_{-i} \in M_{-i}.$$

Remark 1 *It is assumed in these definitions of Nash equilibrium and dominant strategy that only the set of feasible outcomes matter when individuals make choices. This rules out the possibility that the number of times an option appears in that set may matter, or more generally that the very description of the elements of M and/or how they relate to outcomes may influence an individual's choices. These are interesting topics for future research that will not be addressed in this paper.*

The mechanism (M, μ) *implements* the social choice function f *in Nash equilibrium* if it admits a Nash equilibrium at every state, and $f(\theta) = \mu(m^*(\theta))$, for all Nash equilibrium $m^*(\theta)$ at θ, and all $\theta \in \Theta$. If such a mechanism exists, then f is said to be *Nash implementable*. The mechanism (M, μ) *implements* the social choice function f *in dominant strategies* if it admits a dominant strategy profile at every state θ, and $f(\theta) = \mu(m^*(\theta))$, for all profile of strategies $m^*(\theta)$ that are dominant at θ.

Example 1 *Consider the following model of choice with temptation, self-control, and* *willpower.*[5] *There are n individuals with a common long-term goal, but who have trouble to fulfill that long-term goal due to temptation and self-control issues. More formally, their choices when facing a set of options is also driven by a short-term preference. Yet each one of them has some limited ability to exercise self-control thanks to his or her willpower. Various models of willpower and self-control are conceivable. In this example, willpower will be captured by the number of tempting options an individual can overlook to better fulfill his long-term goal. Formally, if S is the set of feasible options, $\succ_L$ captures the long-term goal, $\succ_S$ captures the short-term preference, and i's willpower is captured by the positive integer k_i, then i's choice out of S is the most-preferred element for $\succ_L$ among those in S that are dominated by at most k_i elements of S according to $\succ_S$. This type of behavior does indeed often lead to choice functions that violate IIA. Indeed, one may be able to resist eating a slice of pizza for lunch when the alternative is a salad, but be unable to resist the temptation of both the burger and the pizza slice, and go for the slice if these two otions are available on the menu in addition to the salad. This choice pattern can be explained if the long-term goal ranks the salad above the pizza slice, and the pizza slice above the salad, while the short-term preference is exactly opposed, and $k_i = 1$. Consider now a situation where the state determines the common long-term goal, and (possibly different) short-term goals for the various individuals. I now construct a mechanism that Nash-implements the top-choice of the common long-term goal if $k_i \geq 1$, for all $i \in I$ and $n \geq |X|$. Let $\mathfrak{i} : X \to I$ be such that $\mathfrak{i}(x) \neq (x')$ whenever $x \neq x'$. Each individual chooses a non-negative integer. Let i be the individual in the range of $\mathfrak{i}$ that has the smallest index among all those in that set who chose the largest integer. The outcome is then the one associated to i via $\mathfrak{i}$. Let's check that the top of the long-term goal is the unique Nash equilibrium outcome at any state. Let θ be a state, let x be a Nash equilibrium outcome at that state, and let i be an individual in the range of $\mathfrak{i}$ that is associated to an option y different from x. Varying his strategy, that individual can generate the outcomes x or y. Given that $k_i \geq 1$, his choice can be x, as needed by the condition of Nash equilibrium, only if the long-term goal at θ ranks*

[5]Leading models in economics addressing temptation and self-control investigate commitment by observing choices within menus of menus, and do not explicitly deal with willpower (see Lipman and Pesendorfer (2011) for a survey). Masatlioglu et al. (2011) is a first paper in that axiomatic literature to discuss willpower, but still in perspective of commitment instead of choices in presence of temptation (see also Ozdenoren et al. (2010) for a dynamic model of optimal self-control with willpower). Willpower and its mechanics when exercising self-control for choice with temptation is the main topic of interest in the psychology literature (see e.g. Baumeister and Tierney (2011)).

x above y. Since this reasoning applies to any $y \neq x$, we conclude that x is the top element of the long-term goal at θ. Conversely, it is easy to check that letting the individual associated to that option choose 1*, while all other individuals choose* 0 *forms a Nash equilibrium of the mechanism whose outcome coincides with the social choice. However, this mechanism does not admit a dominant strategy at any state, and hence does not implement the social choice function in dominant strategies. I conjecture it is not implementable in dominant strategies (this remains an open question).*

To simplify notations, and some results (especially in regard to implementation in dominant strategies), I will restrict attention to the special case where states can be decomposed as profiles of types encoding individual choice functions. For each $i \in I$, let Θ_i denote the (finite) set of i's possible types, and $\Theta := \times_{i \in I} \Theta_i$. So I will write C_{θ_i} to describe i's choice function at state $\theta = (\theta_1, \ldots, \theta_I)$, instead of $C_{\theta,i}$.

I conclude this section by introducting Bernheim and Rangel's (2009) extended revealed preference and the associated notion of efficiency. An option a is *unambiguously preferred* to an alternative b if $C(S) \neq b$, for all S such that $a \in S$. Though necessarily incomplete when IIA is violated, this revealed ordering and the associated Pareto principle may still allow to compare some options, and will prove useful on various occasions in the remainder of the paper. An option is *BR-efficient* if there is no alternative option that is unambiguously preferred by all individuals.

3 Implementation in Nash Equilibrium

3.1 Necessary, and Sufficient Conditions

Korpela (2011) has independently derived a sufficient condition for Nash implementability that is similar to the one I will present below (see the second part of Proposition 1).[6] Interestingly, he points out to a little-known work by Hurwicz (1986), who first observed that implementation theory could be extended to domains of choice functions that need not be rationalizable via a complete transitive preference ordering. On the other hand, Hurwicz's notion of "generalized Nash equilibrium" seems less natural in our cobtext with boundedly rational individuals, as it restricts attention to choices over pairs of strategies.[7] Independently of both Korpela (2011) and

[6] I thank Rene Saran for drawing my attention to this paper.

[7] Hurwicz's primary objective was to show how Maskin's results on Nash implementability can be extended to binary relations that are possibly cyclic or incomplete. Even more specifically, Hurwicz

my work, Ray (2010) also derived a necessary condition for Nash implementability with a more resrictive notion of equilibrium that combines condition (1) above with Hurwicz's pairwise comparisons. I will include an explicit reference to those related papers in this subsection whenever appropriate.

Let's start by reminding the essence of Maskin's (1999) result. "Monotonicity" is the key property that emerged from his work.

Maskin Monotonicity *Suppose that C_{θ_i} is rational, for each $i \in I$ and each $\theta_i \in \Theta_i$. Then the social choice function $f : \Theta \to X$ is Maskin Monotonic if $f(\theta') = f(\theta)$, for each θ, θ' such that $LC_{\theta_i}(f(\theta)) \subseteq LC_{\theta'_i}(f(\theta))$ for each $i \in I$, where $LC_{\theta_i}(x) = \{y \in X | x \succ_{\theta_i} y\}$ denotes the lower contour set of x given the preference ordering $\succ_{\theta_i}$.*

Though necessary, the property is not quite sufficient, and various properties have been proposed over the years to complete it so as the guarantee Nash implementability. Perhaps the simplest and most well-known one remains Maskin's "No Veto."

No Veto *Suppose that C_{θ_i} is rational, for each $i \in I$ and each $\theta_i \in \Theta_i$. Then the social choice function $f : \Theta \to X$ satisfies no veto if $f(\theta) = x$, for each $\theta \in \Theta$ for which there exists $i \in I$ such that x is top-ranked according to $\succ_{\theta_j}$, for each $j \in I \setminus \{i\}$.*

Maskin's (1999) classical result establishes that the two following results hold on any rational domain: 1) a social choice function is Nash implementable only if it is Makin Monotonic, and 2) any social choice function that is both Maskin Monotonic and satisfies No Veto is Nash implementable if $\#I \geq 3$.

The following property is a straightforward extension of No Veto to more general individual choice functions.

Property N *Let $\theta \in \Theta$. If there exist x and $i \in I$ such that $C_{\theta_j}(X) = x$, for all $j \in I \setminus \{i\}$, then $f(\theta) = x$.*

The question of how to extend the more substancial property of Maskin Monotonicity is of greater interest. A first idea would be to use Bernheim and Rangel's (2009) extended revealed preference, as defined at the end of Section 2. For each θ, one can indeed require Maskin Monotonicity to hold at any θ, by defining the lower contour set of an option x at θ_i as the set of options that x unambiguously dominates given C_{θ_i}. Unfortunately, it is not very difficult to construct an example of implementable social choice function that does not satisfy this extension of Maskin

often thinks of an individual as a group, and the choice function as the result of some voting procedure implemented among its members, under the additional assumption that "the group never votes on more than two alternative at a time."

monotonicity (e.g. the social choice function defined in Example 3), and this property will thus not prove useful to characterize Nash-implementability.

Maskin monotonicity means that the option that is selected at θ should remain selected if it becomes more appealing in the individual preference orderings. Here is a natural choice-based extension of this idea. An option $x \in X$ is *more appealing* according to C' than to C if $C'(S) = x$, for each S such that $C(S) = x$.

Choice-Based Monotonicity *(CBM)* Let $\theta, \theta' \in \Theta$. If $f(\theta)$ is more appealing according to $C_{\theta'_i}$ than C_{θ_i}, for each $i \in I$, then $f(\theta') = f(\theta)$.

The property is in fact not new, as it first appeared in Aizerman and Aleskerov (1986) in their study of the aggregation of individual choice functions into social choice functions. It is not difficult to check that any social choice function that is Nash implementable must satisfy CBM (either directly, or as an immediate corollary to Proposition 1 below). This result was independently derived by Ray (2010) for the more general case where individual choices are encoded by correspondences.[8] While an interesting necessary condition, CBM turns out to be too weak, as it does not guarantee Nash implementability when combined with Property N. Here is a simple example to illustrate this fact.

Example 2 *Let $X = \{a, b, c\}$, let C_{θ_1} be the rational choice function associated to the ordering $a \succ b \succ c$, let C_{θ_2} be the rational choice function associated to the ordering $b \succ c \succ a$, let C_{θ_3} be the rational choice function associated to the ordering $c \succ b \succ a$, let $C_{\theta'_1}$ be the choice function that coincides with C_{θ_1} except that $C_{\theta'_1}(\{a, b\}) = b$, and let $C_{\theta''_1}$ be the choice function that coincides with C_{θ_1} except that $C_{\theta'_1}(\{a, c\}) = c$. It is then easy to see that the social choice function f, defined by $f(\theta_1, \theta_2, \theta_3) = a$, $f(\theta'_1, \theta_2, \theta_3) = b$, and $f(\theta''_1, \theta_2, \theta_3) = c$, satisfies CBM and N. Yet we will see after Proposition 1 that it is not Nash implementable.*

Ray (2010) proposes to investigate under which forms of bounded rationality do CBM and N imply implementability. Unfortunately his two results have no relevance in our context with choice functions (while his framework also encompasses the more general case of individual choice correspondences). The condition (μ) he uses in his Theorem 3 in fact implies that each individual can have only one type,[9] in which case there

[8]To be precise, the extension of the concept of Nash equilibrium I consider coincides with his notion of "setwise Nash equilibrium," while he investigates implementation in terms of the more restrictive notion of "Generalized Nash equilbrium," which requires additional pairwise conditions.

[9]Indeed, following his notations, $L_i(x, C)$, as defined in (A), is equal to $\{x\}$, since $C(\{x\}) = x$, for each choice function C and each $x \in X$.

is no reason to study implementation, while Sen's condition (α) that he uses in his Theorem 4 amounts to IIA in the case of choice functions, which thus amounts to restrict attention to rationality as already covered by Maskin's original result.

Here is a third property that coincides with Maskin Monotonicity when considering rational types, and that will be useful for our purpose in that it is necessary for Nash-implementability, and sufficient once combined with Property N.

Property M *For each i, there exists a function $X_i : \Theta \to P(X)$ such that*

1. $C_{\theta_i}(X_i(\theta)) = f(\theta)$*, for all* $\theta \in \Theta$ *and all* $i \in I$,
2. *For all* θ, θ' *such that* $f(\theta') \neq f(\theta)$*, there is* $i \in I$ *such that* $C_{\theta'_i}(X_i(\theta)) \neq f(\theta)$.

Remark 2 *If Property M is satisfied for some profile of functions $(X_i)_{i\in I}$ when type sets contain only rational choice functions, then so is it for any other profile $(X'_i)_{i\in I}$ that is larger (i.e. $X_i(\theta) \subseteq X'_i(\theta)$, for each i and θ), provided that the first condition in Property M remains satisfied. Hence Property M is equivalent to that same property applied for the largest possible such profile, i.e. $X_i(\theta) = LC_{\theta_i}(f(\theta))$, for each $i \in I$ and each $\theta \in \Theta$. Notice how the second condition in Property M then indeed boils down to Maskin monotonicity.*

One is then ready to extend Maskin's result to any domain: 1) a social choice function is Nash implementable only if it satisfies Property M, and 2) any social choice function that satisfies both M and N is Nash implementable if $\#I \geq 3$. Yet, it will be useful in applications to have slightly weaker sufficient condition (cf. e.g. Example 3 and Proposition 7). Multiple papers have been written to discuss Nash-implementability with weaker conditions than No Veto. My objective here is not to find extensions of these various conditions to my more general framework. Instead, I inspired myself from this work to derive a condition that is simple as possible, while allowing to cover all the applications discussed in the rest of the paper. Finding a single condition that would be both necessary and sufficient for Nash-implementability on any domain remains an open question of theoretical interest.[10]

Property $\mathfrak{M}$ *There exists a function $X_i : \Theta \to P(X)$, for each $i \in I$, such that the two conditions in Property M are satisfied, and if there exists $j \in I$, $\theta, \theta' \in \Theta$, and $x \in X$ such that $C_{\theta_i}(X) = x$, for all $i \in I \setminus \{j\}$, and either $C_{\theta_j}(X) = x$ or $C_{\theta_j}(X_j(\theta')) = x$, then $f(\theta) = x$.*

[10] Korpela (2011) shows that Moore and Repullo's (1990) condition (which is necessary and sufficient on rational domains) remains sufficient on any domain, but isn't systematically necessary. That condition is a bit weaker than $\mathfrak{M}$, but $\mathfrak{M}$ has so far proved strong enough in applications.

Proposition 1 *(*Necessary, and Sufficient Conditions for Nash-Implementability*)* *If a social choice function f is Nash-implementable, then it satisfies Property M. If $\#I \geq 3$, and f satisfies Property $\mathfrak{M}$, then it is Nash implementable.*

Proof: For necessity, let (M, μ) be a mechanism that implements f, let $\theta \in \Theta$, and let $m^*(\theta)$ be a Nash equilibrium of the game induced by (M, μ) at θ. Let then $X_i(\theta) = \{\mu(m_i, m^*_{-i}(\theta)) | m_i \in M_i\}$, for each $i \in I$. By definition of implementation and of Nash equilibrium, we have: $f(\theta) = \mu(m^*(\theta)) = C_{\theta_i}(X_i(\theta))$, for all i and all θ, which proves the first part of property M. Suppose now that there is a θ' such that $C_{\theta'_i}(X_i(\theta)) = f(\theta)$, for all $i \in I$. Then $m^*(\theta)$ forms a Nash equilibrium of the game induced by (M, μ) at θ', and hence $f(\theta') = \mu(m^*(\theta)) = f(\theta)$, by definition of Nash-implementability, which proves the second part of property M.

As for sufficiency, consider the set of messages $M_i = X \times \Theta \times \mathbb{Z}_+$, for each $i \in I$, and $\mu : M \to X$ defined as follows:

1. If $m_i = (f(\theta), \theta, 0)$, for each i, then $\mu(m) = f(\theta)$.
2. If there is $j \in I$ such that $m_i = (f(\theta), \theta, 0)$, for each $i \in I \setminus \{j\}$, and $m_j = (x, \theta', \alpha) \neq (f(\theta), \theta, 0)$, then $\mu(m) = x$ if $x \in X_i(\theta)$, and $\mu(m) = f(\theta)$ otherwise.
3. In all other cases, $\mu(m) = x$, where x is the first component in the report of the individual with the lowest index among those who submit the highest integer.

First notice that the strategy profile $m^*(\theta) = (f(\theta), \theta, 0)$ forms a Nash equilibrium of the game induced by (M, μ) at θ, following the first condition in Property M.

Second, consider a strategy profile m that forms an equilibrium in θ, and let x be the resulting outcome. Suppose first that $m_i = (f(\theta'), \theta', 0)$, for each i. Then $x = f(\theta')$, and the condition of Nash equilibrium imply that $C_{\theta_i}(X_i(\theta')) = f(\theta')$, for each i. The second condition in Property M implies that $f(\theta') = f(\theta)$. Transitivity thus implies that $x = f(\theta)$, as desired. Suppose next that there exists $j \in I$ such that $m_i = (f(\theta'), \theta', 0)$, for each $i \in I \setminus \{j\}$, and $m_j = (x, \theta'', \alpha) \neq (f(\theta'), \theta', 0)$. The conditions for Nash equilibrium imply that $C_{\theta_j}(X_j(\theta')) = x$ and $C_{\theta_i}(X) = x$, for all $i \in I \setminus \{j\}$. The additional condition that distinguishes $\mathfrak{M}$ from M implies that $x = f(\theta)$, as desired. Finally, in all other cases, the conditions for Nash equilibrium imply that $C_{\theta_i}(X) = x$, for each $i \in I$, and the additional condition that distinguishes $\mathfrak{M}$ from M implies that $x = f(\theta)$, as desired. ■

We can now check that the social choice function f in Example 2 is not Nash implementable, simply by showing that it violates M. Given that $f(\theta) = a$, and given the definition of C_{θ_1}, it must be that $X_1(\theta) = \{a, b, c\}$, or $\{a, b\}$, or $\{a, c\}$, or $\{a\}$. Notice that $C_{\theta'_1}(\{a, b, c\}) = a$, $C_{\theta''_1}(\{a, b\}) = a$, $C_{\theta'_1}(\{a, c\}) = a$, and $C_{\theta''_1}(\{a\}) = a$. For Property M to hold, $f(\theta'_1, \theta_2, \theta_3)$ or $f(\theta''_1, \theta_2, \theta_3)$ should have been equal to a.

The next few subsections apply the results and methodology developed in this subsection to various classes of individual choice functions.

3.2 Rich Domains and Dictatorial Rules

Understanding the limits of implementation starts with "impossibility" results. It is well-known in the rational case that the following properties can be met only by dictatorial rules: 1) the range of the social choice function contains at least three elements, 2) the domain of preferences over which the social choice function is defined is varied enough, 3) the social choice function selects Pareto efficient options, and 4) the social choice function is Nash implementable. The purpose of this section is to propose a definition of rich domain that need not be subsets of rational choice functions, and to investigate its implications for Nash implementability.

An option x is said to be *most preferred* according to a choice function C if $C(S) = x$, for every S that contains x. Given a most preferred option x, an option y is said to be *second most preferred* according to a choice function C if $C(S) = y$, for all S that contains y but not x. Given a most preferred option x, and a second most preferred option y, an option z is said to be *third most preferred* according to a choice function C if $C(S) = z$, for all S that contains z but not x, nor y. Rational choice functions obviously always admit a most, second most, and third most preferred option. It is plausible in some circumstances – though certainly not always – that irrational choice functions might also admit a most, second most, and third most preferred option. Inconsistency in choices might indeed occur only when these most preferred options are not feasible.

A social choice function $f : \Theta \to X$ has a *rich domain* if for each $x, y, z \in Range(f)$, there is a type profile $\theta \in \Theta$ such that x is most preferred according to C_θ, y is second most preferred, and z is third most preferred. Having a rich domain thus means that the mechanism designer cannot exclude[11] that participants have a most, a second most, and a third most preferred option, and that any three options in the range of f can play these roles (no restriction in "tastes"). This boils down to a simplified version of Kalai et al.'s (1979) free triple condition when choice functions are

[11] "Cannot exclude" means that these different properties are met by some types.

rational. The following example illustrates how the properties of BR-efficiency, Nash implementability, rich domain and large ranges can be met by social choice functions that aren't quite dictatorial when individual choice functions aren't rational.

Example 3 *Let $Z = \{a, b, c\}$, let $Z' = \{a', b', c'\}$, and let $X = Z \cup Z' \cup \{d\}$. For each $i \in I$, the set of types Θ_i is the disjoint union of two sets Δ_i and Δ'_i, each of which is isomorphic to the set of strict orderings on Z. Δ_i is interpreted as a set of rational types with the following properties:*

1. $C_{\theta_i}(S) = \arg\max_{\succ_{\theta_i}} S \cap Z$ *if* $S \cap Z \neq \emptyset$,
2. $C_{\theta_i}(S) = d$ *if and only if* $S = \{d\}$.

So options in Z are understood as dominating options outside of Z, and d (which will be interpreted below as a default option) is worse than any other option (choices over sets that do not intersect Z are irrelevant for our argument and can be defined arbitrarily). Δ'_i is interpreted as a set of irrational types, with specific biases:

1. $C_{\theta'_i}(S) = \arg\max_{\succ_{\theta'_i}} \{x \in Z | x' \in S\}$ *if* $S \cap Z \neq \emptyset$, $S \cap Z' \neq \emptyset$, *and* $S \neq X$,
2. $C_{\theta'_i}(S) = \arg\max_{\succ_{\theta'_i}} S$ *if* $S \subseteq Z$,
3. $C_{\theta'_i}(S) = d$ *if and only if* $S = \{d\}$ *or* X.

Thus individuals are rational on subsets of Z, but also subject to an "attraction effect" (first introduced by Huber et al., 1982) with $(x)'$ playing the role of a "decoy" drawing extra attention to x, for each $x \in \{a, b, c\}$. To interpret the third condition, think of the option d as a default option that is less appealing than any other alternative, but that individuals choose nevertheless when irrational and too many options are available, a phenomenon known as "choice overload" (cf. e.g. Iyengar et al. (2006)). Choices out of sets not covered by these three conditions are irrelevant for our argument and can be defined arbitrarily.

Consider the following social choice function:

$$f(\theta) = \begin{cases} C_{\theta_1}(Z) & \text{if } \theta_1 \in \Delta_1 \\ C_{\theta_2}(Z) & \text{if } \theta_1 \in \Delta'_1. \end{cases}$$

It is easy to check that f has a rich domain, is BR-efficient, and has three elements in its range. I apply Proposition 1 in the Appendix to show that it is also Nash-implementable when $\#I \geq 3$.[12] *While individual 1 is a dictator when his choice*

[12] Notice that Property N does not hold though. For instance, consider a type profile θ where $\theta_1 \in \Delta_1$, a is most preferred for C_{θ_1} and b is most preferred for C_{θ_j}, for each $j \geq 2$. Then $C_{\theta_j}(X) = b$, for each $j \neq 1$, and yet $f(\theta) = a$.

function is rational, notice that the social choice function does vary with individual 2's type in other circumstances. In that sense, the first individual qualifies as a dictator only in a weak sense.

For i to be a dictator in the rational case means that the social choice function systematically picks i's top-ranked element within the range. In the absence of rationality, and thus of revealed preferences, such statement becomes meaningless. Here is one possibility to deal with this issue based on Bernheim and Rangel's (2009) idea of extended revealed preference. Individual j is a *dictator in the weak sense* if it is impossible to find $\theta \in \Theta$ and a in the range of f such that a unambiguously dominates $f(\theta)$ given C_{θ_j}. For instance, the first individual was a dictator in the weak sense in the previous example. Indeed, f picks his most preferred option whenever he is rational, and an option in Z whenever he isn't, each of which cannot be unambiguously dominated by any alternative in the range of f. This feature is not a coincidence, and instead hints at a general result that extends classical impossibility results beyond rational domains. Dictatorial rules on the rational domain have the additional feature that the outcome depends exclusively on the type of the dictator. As we saw in Example 3, this need not be the case for rules that are dictatorial in the weak sense. Individual j is a *dictator in the strong sense* if he is a dictator in the weak sense, and the social choice function varies only with j's type.

Proposition 2 *Let $f : \Theta \to X$ be a social choice function that has a rich domain and a range with at least three elements. If it is Nash implementable and BR-efficient, then it is dictatorial in the weak sense. If, in addition, f has a full range (i.e. $Range(f) = X$), then there exists $j \in I$ such that $f(\theta) = C_{\theta_j}(X)$, for each θ, and in particular f is dictatorial in the strong sense.*

Proof: The proof is based on the following abstract result, whose proof is available in the appendix.

Lemma 1 *Let $f : \Theta \to X$ be a social choice function that has a rich domain and a range with at least three elements. If f satisfies Property M, then there exists j and $X_j : \Theta \to P(X)$ such that $f(\theta) = C_{\theta_j}(X_j(\theta))$ and $Range(f) \subseteq X_j(\theta)$, for each $\theta \in \Theta$.*

Proposition 1 and Lemma 1 imply that there exists j and $X_j : \Theta \to P(X)$ such that $f(\theta) = C_{\theta_j}(X_j(\theta))$ and $Range(f) \subseteq X_j(\theta)$, for each $\theta \in \Theta$. The definition of unambiguous domination then implies that j is a dictator in the weak sense. If $Range(f) = X$, then $X_j(\theta) = X$, for all θ, since $Range(f) \subseteq X_j(\theta)$, and the second part of the proposition thus immediately follows from Proposition 1 and Lemma 1 as well. ■

3.3 Limited Attention with Fixed Consideration

In this subsection and the next, I investigate Nash-implementability for social choice functions defined over classes of individual choice functions that display some specific choice biases that have been studied in the literature (see also Example 1). I start by focusing on situations where individuals are rational, in that they systematically maximize a complete transitive preference ordering when choosing out of a set $S \subseteq X$, but their actual choices may violate IIA because they maximize that preference over only a subset of S. Formally, for each i, let Θ_i^* be the set of types that encode all complete and transitive strict preference orderings. A *consideration correspondence* for individual i, $A_i : P(X) \to P(X)$, associates to each set $S \subseteq X$ of feasible options the subset $A_i(S) \subseteq S$ of options that he actually considers. As a consequence, when of type θ_i and when facing a set S of feasible options, individual i will pick the option that maximizes $\succ_{\theta_i}$ over $A_i(S)$:

$$C_{\theta_i}(S) = \arg\max_{\succ_{\theta_i}} A_i(S).$$

As a first example, A_i could be i's psychological filter, as in Cherepanov et al. (2009). Each individual is assumed to maximize his preference over the subset of options that he can "rationalize," meaning that it appears maximal within the feasible set, for at least one "rationale." Manzini and Mariotti (2009) offer another motivation for the framework studied in this section. They suggest that people may proceed by categorization when making choices out of large sets. For instance, when having to choose among many restaurants, one may first decide to focus only on some specific cuisine (e.g. italian), and then pick the best option within that category, while one would consider all the options (and perhaps prefer a mexican restaurant over an italian alternative) when there are only a handful of them. Similarly, one may restrict attention to apartments in a one mile-radius from one's workplace when there are numerous apartment available for rent, but would drop that restriction if only few are available. These examples are also consistent with the notion of "consideration set" in the marketing literature (see Wright and Barbour (1977), see also Masatlioglu et al. (2009) and Lleras et al. (2010) for a recent elaboration on this idea in choice theory). In some circumstances, people are not even aware of some available options. The current analysis does not quite apply to these cases, as the set of options people are aware of is then likely to depend on the description of the mechanism itself (see Remark 1). Instead, I closely follow the meaning that consideration sets take in the marketing literature, namely that the feasible set describes the set of all options the individual is aware of, while the consideration set is a subset of those options that the individual actively considers when making his decision (see e.g. Roberts and Lattin (1991)). The model of choice from lists introduced in the next subsection provide yet other examples, if individuals pay attention to the first $k(s)$ elements of any list with s elements. The key assumption for all the results in this subsection to apply is that consideration sets do not vary with types.

Types parametrize true underlying preferences for notational convenience. On the other hand, what matters for our analysis are the individual choice functions. Some types may thus be redundant, in the sense that they generate the same individual choice function given the consideration correspondence.[13] Hence, a social choice function in this subsection is a function $f : \Theta^* \to X$ such that $f(\theta) = f(\theta')$ for each $\theta, \theta' \in \Theta^*$ such that $C_{\theta_i} = C_{\theta'_i}$, for each $i \in I$. This restriction comes in addition of – and has no impact on the applicability of – the results from Section 3.1.

Lemma 2 *If f satisfies property M, then it is Maskin monotonic.*

Proof: Suppose that

$$LC_{\theta_i}(f(\theta)) \subseteq LC_{\theta'_i}(f(\theta)), \tag{2}$$

for all $i \in I$. Part 1. of property M implies that $f(\theta) \in A_i(X_i(\theta))$ and $A_i(X_i(\theta)) \subseteq LC_{\theta_i}(f(\theta))$. Combining this with condition (2), we conclude that $f(\theta)$ is the maximal element in $A_i(X_i(\theta))$ according to $\succ_{\theta'_i}$. Part 2. of Property M thus implies that $f(\theta') = f(\theta)$, as desired. ■

Proposition 3 *Let f be a social choice function whose range contains at least three options. Then f is Nash-implementable if and only if there exists $j \in I$ and $S \in P(X)$ such that $Range(f) = A_j(S)$ and $f(\theta) = C_{\theta_j}(S)$, for all $\theta \in \Theta$, in which case j is a dictator in the strong sense.*

Proof: The sufficient condition, namely that such f's are Nash-implementable, is straighforward to check. I thus focus on the necessary condition. If f is Nash-

[13] More precisely, $C_{\theta_i} \neq C_{\theta'_i}$, for each $\theta_i, \theta'_i \in \Theta_i$ if and only if for each $x, y \in X$, there exists $S \in P(X)$ such that $A_i(S) = \{x, y\}$ (which is the case, for instance, in Manzini and Mariotti (2009)). For the necessary condition, suppose that x, y are such that $A_i(S) \neq \{x, y\}$, for all $S \in P(X)$. Consider then two types θ_i and θ'_i such that x and y are ranked below any other element of X, y is below x in θ, and above it in θ'. It is easy to check then that $C_{\theta_i} = C_{\theta'_i}$. For the sufficient condition, consider $\theta_i \neq \theta'_i$. Then there exist x and y such that $x \succ_{\theta_i} y$ and $y \succ_{\theta'_i} x$. Then $C_{\theta_i}(S) \neq C_{\theta_i}(S)$, where S is a set such that $A_i(S) = \{x, y\}$.

implementable, then it satisfies property M, thanks to Proposition 1. The previous Lemma then implies that f is Maskin monotonic. Applying a standard result on the rational domain, we conclude that there exists $j \in I$ such that $f(\theta) = \arg\max_{\succ_{\theta_j}} Range(f)$, for each $\theta \in \Theta^*$. We know from property M that $f(\theta) = C_{\theta_j}(X_j(\theta)) = \arg\max_{\succ_{\theta_j}} A_j(X_j(\theta))$, for each $\theta \in \Theta^*$. I now prove that $Range(f) \subseteq A_j(X_j(\theta))$, for each θ. Suppose, to the contrary of what we want to prove, that there exist $\theta \in \Theta^*$ and $x \in Range(f) \setminus A_j(X_j(\theta))$. Consider then $\theta'_j \in \Theta^*$ obtained from θ_j by bringing x to the top, while keeping all the other comparisons unchanged. The second part of property M implies that $f(\theta'_j, \theta_{-j}) = f(\theta)$ (because j does not pay attention to x), which then contradicts the fact that $f(\theta'_j, \theta_{-j}) = \arg\max_{\succ_{\theta'_j}} Range(f)$. One must thus recognize that $Range(f) \subseteq A_j(X_j(\theta))$, for each θ, in order to avoid this contradiction. Consider now $\theta^* \in \Theta^*$ such that the elements in the range of f are ranked below any other element of X according to $\succ_{\theta^*_j}$. If there exists $x \in A_j(X_j(\theta^*)) \setminus Range(f)$, then one would reach a contradiction with the fact that $f(\theta^*) = \arg\max_{\succ_{\theta^*_j}} A_j(X_j(\theta^*))$. Hence it must be that $A_j(X_j(\theta^*)) = Range(f)$. The first part of the result then follows by choosing $S = X_j(\theta^*)$. It remains to show that j is a dictator in the weak sense, since the fact that f does not vary with the types from individuals different from j is obvious. Let $\theta \in \Theta^*$ and let x be an element of the range of f that is different from $f(\theta)$. Hence x and $f(\theta)$ belong to $A_j(S)$, and a fortiori to S. The fact that $C_{\theta_j}(S) = f(\theta)$ implies that x cannot unambiguously dominate $f(\theta)$, as desired. ■

Models of choice with limited attention impose a structure on the violations of rationality. Indeed, individuals are assumed to be endowed with a classical preference ordering, but they might fail to maximize it perfectly due to the fact that they don't pay attention to all the feasible options. Irrationality is understood as a mistake, and a paternalistic mechanism designer may care more in that case about efficiency in terms of the true underlying preferences rather than Bernheim and Rangel's (2009) notion of efficiency. A social choice function f is *Pareto efficient* according to the true underlying preferences in the model with limited attention if there does not exist θ and $x \in X$ such that $x \succ_{\theta_i} f(\theta)$, for each $i \in I$.

Corollary 1 *Suppose that $\#X \geq 3$. Let f be a social choice function in the problem with limited attention that is Pareto efficient. Then f is Nash-implementable if and only if there exists $j \in I$ such that $A_j(X) = X$ and $f(\theta) = C_{\theta_j}(X)$, for all $\theta \in \Theta$, in which case j is a strong dictator. In particular, if no individual pays attention to* *all options when they are all feasible, then there is no social choice function that is Nash-implementable and Pareto efficient.*

Proof: The result follows at once from the previous Proposition, by observing that Pareto efficiency implies that f has full range. Indeed, $x \in X$ must be picked at any $\theta \in \Theta^*$ such that x is top-ranked according to $\succ_{\theta_i}$, for each $i \in I$. ■

We have studied so far problems where consideration sets do not vary with types. The negative conclusions derived in this subsection do not necessarily apply if types encode both preferences and consideration sets. Notice indeed that any individual choice function can be represented by the combination of an adequate preference and consideration function. In particular, both Examples 1 and 3 could be rephrased as examples with types specifying preferences and consideration sets that generate those choice functions. Another example where consideration sets vary with underlying preferences, and where more permissive results can be derived, will be proposed in the next subsection.

3.4 Choice from Lists

I now present a more permissive implementation result for specific classes of choice procedures defined on lists. Assume, for instance, that the mechanism designer has labeled the various options as "Option 1 is (...)", "Option 2 is (...)", etc. This establishes a natural *list* l, which simply amounts to an enumeration of the elements of X, $(x_k)_{k=1}^{|X|}$. It is then assumed that, for any feasible set, individuals investigate the relative appeal of its various elements in an order that is consistent with the mechanism designer's labeling system. A first example of choice procedure we will consider is given by Rubinstein and Salant's (2006, Example 6) "Stop When You Start to Decline." In this model, individuals are endowed with a complete and transitive strict preference ordering on X (which, in the present paper, may vary with their types), and check options out of any feasible set $S \subseteq X$ by following the sequence that is consistent with the list's enumeration. They continue this process as long as the next option improves upon the last according to their preference. The option picked from S is then the last in this improving sequence, i.e. the last one they paid attention to before finding a new option that is worse according to their preference. A second example of choice procedure covered in this subsection, is Simon's (1955) satisficing procedure, as formalized in Rubinstein and Salant (2006, Example 2). In this alternative model, individuals are endowed with a utility function and a real number that represents a threshold (both of which may vary in the present paper with their types). Again, they are assumed to check options out of any feasible set $S \subseteq X$ by following the sequence that is consistent with the list's enumeration. They then pick the first option whose utility happen to be larger than the threshold. If none of them meets this requirement, then they pick the last option they have looked at among those that are feasible

More generally, we will consider in this subsection any domain Θ of type profiles that satisfy the following condition: for each i and each $\theta_i \in \Theta_i$, if there exists $K \in \{1, \ldots, |X|\}$ such that $S = \{x_1, \ldots, x_K\}$, then

$$C_{\theta_i}(S) = \begin{cases} C_{\theta_i}(X) & \text{if } C_{\theta_i}(X) \in S \\ x_K & \text{if } C_{\theta_i}(X) \notin S. \end{cases} \tag{3}$$

Notice that this condition is not concerned by the way people choose out of subsets that are not of the form $\{x_1, \ldots, x_K\}$ for some $K \in \{1, \ldots, |X|\}$, and it thus imposes fewer restrictions than IIA. On the other hand, the restriction $C_{\theta_i}(S) = x_K$ if $C_{\theta_i}(X) \notin S$ imposes a restriction on choices that isn't covered by IIA. It is easy to check that the two choice procedures presented in the previous paragraph lead to individual choice functions that indeed satisfy condition (3).

Remark 3 *If individuals are rational, and have single-peaked preferences on the list, then they satisfy condition (3). Notice also that choice functions generated by applying the satisficing procedure satisfy IIA, and can thus be rationalized by the maximization of a complete transitive strict ordering, but that this underlying preference need not be single-peaked.*

Let f^l be the social choice function defined as follows: $f^l(\theta) = x_{k^l(\theta)}$, where $k^l(\theta) = \min\{k \in \{1, \ldots, |X|\} | (\exists i \in I) : x_k = C_{\theta_i}(X)\}$.

Proposition 4 *The social choice function f^l is Nash-implementable if $|I| \geq 3$.*

Proof: The result will follow by applying Proposition 1 after checking that f^l satisfies Property $\mathfrak{M}$. Let $i \in I$ and $\theta \in \Theta$. Define

$$X_i(\theta) = \begin{cases} X & \text{if } C_{\theta_i}(X) = f^l(\theta) \\ \{x_1, \ldots, f^l(\theta)\} & \text{if } C_{\theta_i}(X) \neq f^l(\theta). \end{cases}$$

We start by checking the first condition in Property M. If $C_{\theta_i}(X) = f^l(\theta)$, then $X_i(\theta) = X$ and $C_{\theta_i}(X_i(\theta)) = f^l(\theta)$. If $C_{\theta_i}(X) \neq f^l(\theta)$, then $C_{\theta_i}(X)$ comes after

$f^l(\theta)$ in the list, and hence $C_{\theta_i}(X_i(\theta))$ is the last element of $X_i(\theta)$, which is $f^l(\theta)$, as desired. Let's now check the second condition in Property M. Let θ, θ' be such that $C_{\theta'_i}(X_i(\theta)) = f^l(\theta)$, for each $i \in I$. Notice that $f^l(\theta')$ cannot come strictly after $f^l(\theta)$ in the list, since $X_i(\theta) = X$, for at least one individual i. Notice that it cannot come strictly before either, as this would require $C_{\theta'_i}(X)$ to fall strictly before $f^l(\theta)$, for at least one individual i, which would contradict $C_{\theta'_i}(X_i(\theta)) = f^l(\theta)$. Hence it must be that $f^l(\theta') = f^l(\theta)$.

We can conclude the proof by checking the additional condition that distinguishes Property $\mathfrak{M}$ from M. Let $j \in I$, $\theta, \theta' \in \Theta$, and $x \in X$ be such that $C_{\theta'_i}(X) = x$, for each $i \in I \setminus \{j\}$, and $C_{\theta'_j}(X_j(\theta)) = x$ or $C_{\theta'_j}(X) = x$. If $C_{\theta'_j}(X) = x$, then $f^l(\theta') = x$, by definition of f^l. If $C_{\theta'_j}(X_j(\theta)) = x$, then $C_{\theta'_j}(X)$ does not come before x in the list. Combining that with the other conditions on the choice for each $i \neq j$, we conclude that $f(\theta') = x_{k^l(\theta')} = x$, as desired. ■

Remark 4 *It is easy to see that f^l does not satisfy Property N, and hence Property $\mathfrak{M}$ may indeed be more useful than the combination of M and N. Notice also that f^l cannot be implemented by the direct mechanism where $M_i = \Theta_i$, for each $i \in I$, and $\mu = f^l$. For instance, reporting a θ_i for which $C_{\theta_i}(X)$ is equal to the first element in the list, for each $i \in I$, forms a Nash equilibrium at any type profile, which usually does not coincide with the option that should be implemented according to f^l at that type profile. This shows again the interest of the sufficient condition in Proposition 1, and especially of the more sophisticated mechanism proposed in its proof.*

3.5 Social Choice Correspondences

A *social choice correspondence* is a correspondence $F : \Theta \to X$ that selects a non-empty subset of options for each profile of types. The mechanism (M, μ) *implements* the social choice correspondence F in Nash equilibrium if

$$F(\theta) = \{\mu(m^*(\theta)) | m^* \text{ is a Nash equilibrium at } \theta\}, \text{ for each } \theta \in \Theta.$$

Proposition 1 immediately extends to the case of social choice correspondences.

Proposition 1' *(*Necessary, and Sufficient Conditions for Nash-Implementability of Social Choice Correspondences*) If a social choice correspondence F is Nash-implementable, then there exists a function $X_i : \{(x, \theta) \in X \times \Theta | x \in F(\theta)\} \to P(X)$, for each $i \in I$, such that:*

1. *$C_{\theta_i}(X_i(x,\theta)) = x$, for each $x \in F(\theta)$ and each $\theta \in \Theta$.*

2. *For all θ, θ', and $x \in F(\theta)$, if $C_{\theta'_i}(X_i(x,\theta)) = x$, for each $i \in I$, then $x \in F(\theta')$.*

If, in addition, $x \in F(\theta)$ for any θ and any x for which there exists i such that $C_{\theta_j}(X) = x$ for all $j \in I \setminus \{i\}$, then F is Nash-implementable.

The proof is virtually identical to that of Proposition 1, and is left to the reader.

Let F^{BRE} be the social choice correspondence that associates to each type profile θ the set of BR-efficient options at θ. As a variant, let F^{Eff} be the social choice correspondence defined as follows:

$$F^{Eff}(\theta) = \{x \in X | (\exists (Y_i)_{i\in I} \text{ subsets of } X) : C_{\theta_i}(Y_i) = x, \text{ for all } i, \text{ and } X = \cup_{i\in I} Y_i\},$$

for each $\theta \in \Theta$. Notice that F^{Eff} has non-empty values, as needed, since it includes any option x such that $x = C_{\theta_i}(X)$, for some $i \in I$.

Proposition 5 *Suppose that X contains at least three elements. F^{Eff} is Nash-implementable on all domains, while F^{BRE} isn't. $F^{Eff} \subseteq F^{BRE}$, and F^{Eff} coincides with the set of Pareto optimal options when C_{θ_i} is rational for all i.*

Proof: Let's start by showing that F^{BRE} is not Nash-implementable. Suppose that $X = \{x_1, \ldots, x_K\}$. Consider a type profile θ where, for each $i \neq 1$, $C_{\theta_i}(Y)$ is the option in Y that has the lowest index. In other words C_{θ_i} is rational with the individual prefering options with lower indices. Suppose that C_{θ_1} follows the same pattern, except that $C_{\theta_1}(\{x, x_K\}) = x_K$ instead of x, for each $x \in X \setminus \{x_K\}$. Hence $x_K \in F^{BRE}(\theta)$. By definition of θ, the first part of the necessary condition in Proposition 1' can be satisfied only if $X_i(x_K, \theta) = \{x_K\}$, for all $i \neq 1$, and $X_1(x_K, \theta) = \{x_K\}$, or $\{x_K, x\}$, for some $x \in X \setminus \{x_K\}$. If $X_1(x_K, \theta) = \{x_K\}$, then the second part of the necessary condition implies that $x_K \in F^{BRE}(\theta)$, for all θ, which is clearly not true. In the case where $X_1(x_K, \theta) = \{x_K, x\}$, for some $x \in X \setminus \{x_K\}$, consider an option y different from both x and x_K, and the type θ'_1 that differs from θ_1 only in that $C_{\theta'_1}(\{x_K, y\}) = y$ instead of x_K. The second part of the necessary condition in Proposition 1' implies that x_K must belong to F^{BRE} if it were Nash-implementable, but it doesn't, as y unambiguously dominates x_K.

I now show that F^{Eff} is Nash implementable on any Θ. By definition of F^{Eff}, one can associate to any $x \in F^{Eff}(\theta)$ a collection $(Y_i^{x,\theta})_{i\in I}$ of subsets of X such that $C_{\theta_i}(Y_i^{x,\theta}) = x$, for each $i \in I$, and $X = \cup_{i\in I} Y_i^{x,\theta}$. Let's then check the sufficient conditions in Proposition 1'. Take $X_i(x,\theta) = Y_i^{x,\theta}$. The two parts of the necessary condition follow at once by definition of F^{Eff}. As for the additional condition needed for sufficiency, it is trivially satisfied by F^{Eff} since $F^{Eff}(\theta)$ includes by definition any x such that $x = C_{\theta_i}(X)$, for some $i \in I$.

Consider a type profile θ, and $x \in F^{Eff}(\theta)$. The condition $C_{\theta_i}(Y_i) = x$ implies that none of the elements in $Y_i \setminus \{x\}$ unambiguously dominates x for i. The condition $X = \cup_{i\in I} Y_i$ implies that it is impossible to find an alternative that would be unambiguously preferred to x by all individuals. In other words, x is BR-efficient, and we have proved that $F^{Eff} \subseteq F^{BRE}$.[14]

Finally, if θ is such that C_{θ_i} is rational, for each $i \in I$, and x is Pareto efficient for the associated revealed preferences, then define Y_i as the lower contour set of x according to the revealed preference, for each i. By definition, $C_{\theta_i}(Y_i) = x$, for all i. If there is $x' \in X$ such that $x' \in X \setminus Y_i$, for each $i \in I$, then x' is revealed preferred to x for all i, thereby contradicting the efficiency of x. This shows that $X = \cup_{i\in I} Y_i$, and hence F^{Eff} contains all the Pareto efficient options. For the other inclusion, showing that $F^{Eff}(\theta)$ is included in the set of Pareto efficient, consider an option x that isn't Pareto efficient. This means that there exists $x' \in X$ such that x' is revealed preferred to x for all i. If so, and if $C_{\theta_i}(Y_i) = x$, for all i, then $x' \in X \setminus Y_i$, for all i, and hence $x \notin F^{Eff}(\theta)$, since $x' \notin \cup_{i\in I} Y_i$. ■

Let's focus now on the model of choice with limited attention from Section 3.3. For each i, Θ_i^* is the set of types that encode all complete and transitive strict preference orderings, and $A_i : P(X) \to P(X)$ is individual i's consideration correspondence, implying that $C_{\theta_i}(S) = \arg\max_{\succ_{\theta_i}} A_i(S)$, for each $\theta_i \in \Theta_i^*$, each $i \in I$, and each $S \in P(X)$. Let F^{Pareto} be the social choice correspondence that associates to any state $\theta \in \Theta^*$ the set of options that are Pareto efficient for the profile of rational orderings associated to θ. The next proposition is negative, as it shows that it is impossible to Nash-implement the Pareto correspondence for choice with limited attention with fixed consideration if at least one individual is irrational.

Proposition 6 *Consider the model of choice with limited attention. F^{Pareto} is Nash-implementable if and only if $A_i(S) = S$, for each S, i (or C_{θ_i} is rational).*

Proof: It is well-known that the Pareto correpondence is Nash-implementable on the domain of rational choice functions, and hence we will focus on the necessary condition. Suppose that there exist $S \in P(X)$ and $i \in I$ such that $A_i(S) \subsetneq S$. Let S^*

[14] Going back to the example presented in the first paragraph of the proof, notice how $x_K \in F^{BRE}(\theta) \setminus F^{Eff}(\theta)$.

be a set with that property for i, and such that no larger set has that property for i (i.e. $A_i(S) = S$ for any set S larger than S^*). Let then $x \in S^* \setminus A_i(S^*)$, and suppose that $\theta_i \in \Theta_i^*$ is such that options of $X \setminus S^*$ are ranked above all others, and x is ranked top among remaining options. Suppose also that the ranking associated to θ_j is strictly opposite to the ranking associated to θ_i, for all $j \in I \backslash \{i\}$. Observe that $x \in F^{Pareto}(\theta)$. Consider now the set $X_i(x, \theta)$ as given in Proposition 1'. $C_{\theta_i}(X_i(x, \theta)) = x$ implies that $A_i(X_i(x, \theta)) \subseteq S^*$. If $A_i(X_i(x, \theta)) = S^*$, then $X_i(x, \theta)$ is strictly larger than S^*, thereby contradicting the maximality of S^*. Hence $A_i(X_i(x, \theta)) \subsetneq S^*$. Let then $x' \in S^* \setminus A_i(X_i(x, \theta))$, and consider a type $\theta_i' \in \Theta_i^*$ that differs from θ_i only in that x' is now ranked top among all options in X. The second part of the necessary condition in Proposition 1' implies that x should also belong to $F^{Pareto}(\theta_i', \theta_{-i})$ if it were Nash-implementable, but it doesn't since x' Pareto dominates x. Hence it must be that $A_i(S) = S$, for each $i \in I$ and each $S \in P(X)$, to avoid the contradiction, and we are done proving the result. ■

The core is another important example of social choice correspondence that is Nash implementable. To present results on this topic, I now study an extension of Shapley and Scarf's (1974) house allocation problem to situations where individuals may be boundedly rational. Before starting the analysis, notice that individual choices are restricted to be single valued throughout the paper. The notions of Nash equilibrium and dominant strategies are more delicate to interpret when considering individual choice correspondences instead of functions. Indeed, while each single element in the choice set has the property of being optimal (a utility maximizer) in the case of rational choice with indifferences, it is not necessarily true with more general models of choice that each element has a property of being a reasonable choice on its own when considering correspondences (think for instance of the top cycle correspondence - Moon (1968) and Kalai and Schmeidler (1979)). There are ways of dealing with this difficulty, but I leave this topic for future research. Even so, important applications, such as the house allocation problem or, more generally, exchange economies, involve private consumption where one must face correspondences instead of functions (because individuals care only about their own consumptions, while X contains allocation vectors that specify a bundle for each individual). It is thus important to point out that all the results in this paper can be reproduced in the case of private consumption by defining individual choices over their own private consumption instead of allocation vectors. The following brief analysis of the house allocation problem provides a clear illustration of that point.

Consider the classical problem where our I individuals each own one unit of an indivisible object. Objects are often thought off as houses in the classical theoretical example. One can think, for instance, of allocating on-campus housing to incoming undergrads, with an initial endowment that has been previously determined by a lottery. Let O denote the set of all objects: $O = \{o_i^*|i \in I\}$, where o_i^* is i's initial endowment. Then $X = \{z \in O^I|z_i \neq z_j, \forall i \neq j\}$. To avoid individual choice correspondences, while assuming that individuals care only about the object they consume, choices are defined over subsets of O instead of X. So I will assume in this subsection that C_{θ_i} is a function that selects a single element out of every subset of O, for each $\theta_i \in \Theta_i$. As before, a social choice function is a function that associates an element of X to each element of Θ, but elements in the range of f are now vectors with I components, given the specific structure of X. A similar point applies to the outcome function of mechanisms. Condition (1) defining a Nash equilibrium thus now becomes:

$$\mu_i(m^*(\theta)) = C_{\theta_i}(\{\mu_i(m_i, m^*_{-i}(\theta))|m_i \in M_i\}).$$

For each group S of individuals, let then $\mathcal{F}(S)$ be the set of house allocations that a coalition can achieve among its members. Formally, $\mathcal{F}(S)$ is the set of $\alpha \in O^S$ such that $\{\alpha_i|i \in S\} = \{o_i^*|i \in S\}$ and $\alpha_i \neq \alpha_j$ whenever $i \neq j$. Inspired by the definition of F^{Eff}, consider the following social choice correspondence:

$$F^C(\theta) = \{x \in \mathcal{F}(I)|(\exists (T_i)_{i\in I} \text{ subsets of } O) : C_{\theta_i}(T_i) = x_i, \forall i, \text{ and}$$
$$(\forall S \in P(I))(\forall \alpha \in \mathcal{F}(S))(\exists i \in S) : \alpha_i \in T_i\}$$

In order to make sure that F^C is well-defined as a social choice correspondence, we need to check that it is non-empty valued. For this, I propose an extension of Shapley and Scarf's (1974) *top trading cycle* procedure. Fix a state θ, and compute at that state the object that each individual i would pick if he was free to choose any element of O. Draw an arc from individual i and j if and only if $C_{\theta_i}(O)$ is j's endowment. This graph must have at least one cycle. Implement the trades induced by all these cycles, and let $O' \subsetneq O$ be the set of objects remaining, i.e. $O' = \{o_i^*|i \in S\}$ where S is the set of individuals who weren't part of a cycle. Then iterate the procedure with the remaining individuals and objects. Call $\alpha^*(\theta)$ the resulting allocation of objects. Let's check that $\alpha^*(\theta) \in F^C(\theta)$, for each $\theta \in \Theta$. For each individual i, let T_i be the set of objects remaining when he became part of a trading cycle in the procedure. The fact that $C_{\theta_i}(T_i) = \alpha_i^*(\theta)$ thus follows immediately from the definition of that cycle in the procedure. Let $S \in P(I)$, $\alpha \in \mathcal{F}(S)$, and let j be one of the individuals in S who is among the earliest to be part of a trading cycle in the induction. Then $\alpha_j \in T_j$, for each $\alpha \in \mathcal{F}(S)$, by definition of the induction, of $\mathcal{F}(S)$ and of T_j. We have thus proved indeed that $\alpha^*(\theta) \in F^C(\theta)$.

We will establish that F^C coincides with the core when individual choice functions are rational and it is well-known in that case, that the allocation resulting from the top trading cycle is the only element in the core. However, $\alpha^*(\theta)$ need not be the only element of $F^C(\theta)$ for more general type profiles, as the following example shows.

Example 4 *Suppose that $I = \{1, 2, 3\}$, and consider a type profile θ so that $C_{\theta_1}(O) = o_3^*$, $C_{\theta_2}(O) = o_2^*$, $C_{\theta_3}(O) = o_3^*$, $C_{\theta_2}(\{o_1^*, o_2^*\}) = o_1^*$, and $C_{\theta_3}(\{o_2^*, o_3^*\}) = o_2^*$. The top trading cycle leaves individuals with their initial endowments. Consider the alternative allocation $\beta = (o_3^*, o_1^*, o_2^*)$. It also belongs to $F^C(\theta)$. This can be proved using $T_1 = O$, $T_2 = \{o_1^*, o_2^*\}$ and $T_3 = \{o_2^*, o_3^*\}$. Notice indeed that $C_{\theta_i}(T_i) = \beta_i$, by definition of T_i, θ and β. To conclude the argument, one must show that, for all $S \in P(I)$ and all $\alpha \in \mathcal{F}(S)$, there exists $i \in S$ such that $\alpha_i \in T_i$. This is trivially true if $1 \in S$, since $T_1 = O$. If $1 \notin S$, then the feasibility of α implies that $\alpha_3 \in O \setminus \{o_1^*\}$. Hence the property is true by choosing $i = 3$ if he is a member of S. Finally, if neither 1 nor 3 are members of S, then it must be that $S = \{2\}$ and $\mathcal{F}(S) = \{o_2^*\} \subseteq T_2$, which establishes the property and the fact that $\beta \in F^C(\theta)$, as desired.*

Notice that Bernheim and Rangel's (2009) extended revealed preference can also be used to define an extension of the core that is applicable in particular in the house allocation problem with initial endowments. Say that coalition S *BR-blocks* an allocation α if there exists $\beta \in \mathcal{F}(S)$ such that β_i unambiguously dominates α_i, for each $i \in S$. The *BR-core*, denoted F^{BRC}, then associates to each profile θ of types the set of allocations $\alpha \in \mathcal{F}(I)$ such that there is no coalition S that BR-blocks α. We now provide a result in the spirit of Proposition 5 (see proof in the Appendix).

Proposition 7 *Suppose that O contains at least three elements. F^C is Nash-implementable on all domains, while F^{BRC} isn't. $F^C \subseteq F^{BRC}$, and F^C coincides with the core when C_{θ_i} is rational for all i.*

Remark 5 *Bade (2008) studies the house allocation problem for general choice functions. She defines two notions of efficiency and of cores, one based on Bernheim* *and Rangel's (2009) revealed preference, cf.* F^{BRE} *and* F^{BRC}*, and one on an alternative revealed ordering whereby* $a \succ b$ *if there exists* S *that contains* b *and such that* $C(S) = a$*. The latter ordering is complete but often intransitive, leading in many cases to the emptiness of both concepts (efficiency and core). She does not study the implementability of the various concepts she defines, but instead defines the top-trading cycle (same as above) and serial dictatorship procedures to show that an equivalence result by Abdulkadiroglu and Sonmez (1998) does not extend to the case where individuals are irrational.*

4 Implementation in Dominant Strategies

4.1 Necessary, and Sufficient Conditions

I now present an extension of the revelation principle to our more general domain in order to characterize the set of social choice functions that are implementable in dominant strategies.

Proposition 8 *(*Weak Revelation Principle*) The social choice function* f *is implementable in dominant strategies if and only if there exists a set* $\mathcal{M}_i$ *of messages, for each* $i \in I$*, such that truth-telling is a dominant strategy for every type profile in a mechanism where individual* i*'s message space is* $\Theta_i \cup \mathcal{M}_i$*, for each* $i \in I$*, and the outcome function coincides with* f *when all reports fall within the type sets.*

Proof: The sufficient condition is obvious. For the necessary condition, suppose that f is implementable in dominant strategies via the mechanism $((M_i)_{i \in I}, \mu)$. Let $m^* : \Theta \to M$ be a profile of strategies such that $m_i^*(\theta)$ is a dominant strategy at θ, and $m_i^*(\cdot)$ varies only with θ_i (such a profile can always be constructed since a dominant action at a state θ remains dominant for individual i at any θ' such that $\theta'_i = \theta_i$). Let then $\mathcal{M}_i$ be the set of messages that do not belong to the image of the function $m_i^*(\cdot)$. For each $i \in I$, let $g_i : \Theta_i \cup \mathcal{M}_i \to M_i$ be the function defined as follows: $g_i(\mathfrak{m}_i) = m_i^*(\mathfrak{m}_i)$ if $\mathfrak{m}_i \in \Theta_i$ and $g_i(\mathfrak{m}_i) = \mathfrak{m}_i$ if $\mathfrak{m}_i \in \mathcal{M}_i$. Let then $\mu^* : \times_{i \in I}(\Theta_i \cup \mathcal{M}_i) \to X$ be the function defined as follows: $\mu^*(\mathfrak{m}) = \mu((g_i(\mathfrak{m}_i))_{i \in I})$. It is easy to check that truth-telling is a dominant strategy in every type profile for this new mechanism. ■

This weak revelation principle highlights that, as in the rational case, there is no loss of generality for the mechanism designer to make his decision on which option to implement based solely on the individuals' direct reports about their types. On the other hand, the fact that individuals' choices might violate IIA implies that the mechanism designer might want to add some 'decoy messages' to 'nudge' individuals to see truth-telling as a dominant strategy. Obviously, Proposition 8 boils down to the classical revelation principle when choice functions are rational, as the messages in $\mathcal{M}_i$ can be deleted without changing the individuals' behavior in the mechanism. Note that the exact same class of mechanisms was introduced by Mookherjee and Reichelstein (1990) under the name of "augmented revelation mechanism" to state their "augmented revelation principle." They show that any social choice rule which is Nash-implementable (in the standard case of rational choice functions, but allowing also for incomplete information) can be implemented by an augmented revelation mechanism for which truthtelling is one equilibrium, and then state the "selective elimination condition" on the social choice function to guarantee that truth is the only equilibrium, thereby providing a sufficient condition for Nash implementability. The non-type auxiliary messages are used there to destroy undesired equilibria without introducing new ones, while here they are used to influence individual choice behaviors.

The fact that $\mathcal{M}_i$ may have to be non-empty makes this weak revelation principle less helpful for characterizing the whole class of social choice functions that are implementable in dominant strategies in specific examples. Phrasing a tractable condition that is both necessary and sufficient seems out of reach. Yet the weak revelation principle allows to derive a necessary condition for implementability in dominant strategies that share some similarities with Property M.

Property P *For each $i \in I$, there exists a function $X_i : \Theta_{-i} \to P(X)$ such that $C_{\theta_i}(X_i(\theta_{-i})) = f(\theta)$, for all $\theta \in \Theta$.*

Proposition 9 *(*Necessary Condition for Implementability in Dominant Strategies*) If a social choice function f is implementable in dominant strategies, then it satisfies Property P.*

Proof: Consider a 'direct' mechanism as derived in Proposition 8, and let μ denote its outcome function. Define $X_i(\theta_{-i}) = \{\mu(m_i, \theta_{-i}) | m_i \in \Theta_i \cup \mathcal{M}_i\}$, for each i, θ. Property P then follows at once from the fact that truth-telling is a dominant strategy. ■

Remark 6 *Properties M and P have a similar flavor, but differ in some important aspects. First, the X_i functions do not depend on θ_i. This makes Property P more difficult to satisfy than Property M. Property P trivially implies a special case of the* *second condition in Property M: if* $C_{\theta'_i}(X_i(\theta_{-i})) = f(\theta)$*, then* $f(\theta'_i, \theta_{-i}) = f(\theta)$*. On the other hand, Property M is stronger in that it also restricts the value of* f *when types of multiple individuals change in a way that their choices out of the relevant* X_i*'s remain unchanged. As a result, there is no logical implication between M and P (and more generally no general logical implication between Nash-implementation and implementation in dominant strategies). The fact that M does not always implies P is already true on some domains containing only rational choice functions (see Theorem 7.2.3 and the discussion that follows it in Dasgupta et al. (1979)). On the other hand, it is always true that P implies M when types in the domain encode only rational choice functions (see Theorem 3.2.1 in Dasgupta et al. (1979)). The next example shows that departing from rationality breaks down that logical implication.*

Example 5 *Let* $X = \{a, b, c, d\}$*,* $I = \{1, 2\}$*, and* $\bar{\Theta}_i$ *be the set of all strict rational preferences on* X *such that* d *is first or second-best. Suppose that* $C_{\theta_i} = \arg\max_{\succ_{\theta_i}} A_i(S)$*, where* $A_i(S)$ *contains the first two elements of* S *according to the alphabetical order.*[15] *These choice procedures are consistent with the model described in Section 3.3 and, more specifically, are rationalizable in the sense of Cherepanov et al. (2009), Manzini and Mariotti (2009), and Lleras et al. (2010), as well as in the sense of Masatlioglu et al. (2009). Consider the social choice function* $f : \bar{\Theta} \to X$ *defined as follows:*

$$f(\theta) = \begin{cases} a & \text{if } a \succ_{\theta_1} x \text{ and } a \succ_{\theta_2} x, \forall x \in \{b, c, d\} \\ b & \text{if } b \succ_{\theta_1} x \text{ and } b \succ_{\theta_2} x, \forall x \in \{a, c, d\} \\ c & \text{if } c \succ_{\theta_1} x \text{ and } c \succ_{\theta_2} x, \forall x \in \{a, b, c\} \\ d & \text{otherwise.} \end{cases}$$

In other words, f *picks d except if an alternative is top best for both individuals. It is easy to check that* f *satisfies Property P (and is implementable in dominant strategies, as a consequence of Proposition 10 below) with*

$$X_i(\theta_{-i}) = \begin{cases} \{a, d\} & \text{if } a \succ_{\theta_{-i}} x, \forall x \in \{b, c, d\} \\ \{b, d\} & \text{if } b \succ_{\theta_{-i}} x, \forall x \in \{a, c, d\} \\ \{c, d\} & \text{if } c \succ_{\theta_{-i}} x, \forall x \in \{a, b, c\} \\ \{d\} & \text{otherwise.} \end{cases}$$

[15] Notice that $C_{\theta_i} \neq C_{\theta'_i}$, for all $\theta_i, \theta'_i \in \bar{\Theta}_i$ such that $\theta_i \neq \theta^*_i$ - see footnote 13.

On the other hand, f violates Property M (and hence is not Nash-implementable). Let θ be such that d is ranked top for both individuals. The first condition in Property M implies that $C_{\theta_i}(X_i(\theta)) = d$, for both $i = 1, 2$. Hence i pays attention to d, and $X_i^*(\theta)$ contains at most two elements: $X_1(\theta) = \{x, d\}$ and $X_2(\theta) = \{y, d\}$, for some $x, y \in \{a, b, c, d\}$. If θ' is such that $z \in X \setminus \{d, x, y\}$ is top-ranked and d is second-best, for both individuals, then $C_{\theta'_i}(X_i(\theta)) = d$, for both $i = 1, 2$, which would contradict the second condition in Property M, given that $f(\theta') = z$. Hence f violates Property M, and is not Nash-implementable.

When the domain contains only rational choice functions, the weak revelation principle also holds true with $\mathcal{M}_i = \emptyset$ for all $i \in I$, by direct application of IIA, and thus a social choice function is implementable in dominant strategies in that case if and only if

$$\forall i, \theta_i, \theta'_i, \theta_{-i} : f(\theta) \succ_{\theta_i} f(\theta'_i, \theta_{-i}) \text{ or } f(\theta) = f(\theta'_i, \theta_{-i}). \tag{4}$$

This standard condition is in turn equivalent to Property P, simply by taking $X_i(\theta_{-i}) = \{f(\theta_i, \theta_{-i}) | \theta_i \in \Theta_i\}$. Property P is not equivalent anymore with implementability in our more general problem where individual choice functions may violate IIA. Yet such construction hints at a sufficient condition that proves useful in some problems (see Example 5 above and Proposition 13 below).

Proposition 10 *If a social choice function f satisfies property P for the specific sequence of functions $(X_i^*)_{i \in I}$, where $X_i^*(\theta_{-i}) = \{f(\theta_i, \theta_{-i}) | \theta_i \in \Theta_i\}$, for each θ_{-i} and each i, then f is implementable in dominant strategies.*

Proof: The fact that f satisfies Property P for $(X_i^*)_{i \in I}$ implies that truthtelling is a dominant strategy in the direct mechanism where individuals report their types and the outcome function coincides with f. ■

It should be clear from the weak revelation principle that the condition from the previous proposition is not sufficient. It may help indeed to consider non-empty sets $\mathcal{M}_i$. Adding such messages can be useful to make truthtelling a dominant strategy when others report their true types, but at the same time introduce new constraints for truthtelling to be a dominant strategy, as other individuals have now more strategies at their disposal. We will encounter some examples of social choice functions that are implementable in dominant strategies without satisfying the assumption of Proposition 10, but capturing these conditions in a tractable property that would further characterize implementation in dominant strategies seems out of reach.

4.2 Rich Domains and Dictatorial Rules

I follow the same model and notations as those introduced in Section 3.2. As explained in Remark 6, there is no logical implication between Properties M and P. It turns out, though, that an analogue of Lemma 1 holds when M is replaced by P. In turn, that result allows to prove an analogue Proposition 2, thanks to Proposition 9. Details are available upon request.

Proposition 11 *Let $f : \Theta \to X$ be a social choice function that has a rich domain and a range with at least three elements. If it is implementable in dominant strategies and BR-efficient, then it is dictatorial in the weak sense. If, in addition, f has a full range (i.e. Range$(f) = X$), then there exists $j \in I$ such that $f(\theta) = C_{\theta_j}(X)$, for each θ, and in particular f is dictatorial in the strong sense.*

The following example, that is closely related to Example 3, shows that the properties of BR-efficiency, Nash implementability, rich domain and large ranges can be met by social choice functions that are dictatorial only in a weak sense.

Example 6 *Let's consider the same problem as in Example 3, except for one additional restriction:*

$$C_{\theta_2}(Z') = (C_{\theta_2}(Z))',$$

for all $\theta_2 \in \Theta_2$. Let's check now that the social choice function f defined there is also implementable in dominant strategies. Consider the mechanism (M, μ), where $M_1 = \Theta_1 \cup \{m\}$, $M_i = \Theta_i$, for each $i \in I \setminus \{1\}$, $\mu(\theta) = f(\theta)$, and $\mu(m, \theta_{-i}) = (C_{\theta_2}(Z))'$. We have to prove that $m_i^*(\theta_i) = \theta_i$ is a dominant strategy at θ, for each $i \in I$ and each $\theta \in \Theta$. Let's start with the first individual. Notice that the set of options that he can generate by choosing his message, while others report θ'_{-1} is $\{f(\theta'_1, \theta'_{-1}) | \theta'_1 \in \Theta_1\} \cup \{(C_{\theta'_2}(Z))'\} = Z \cup \{(C_{\theta'_2}(Z))'\}$. His choice out of that set is $C_{\theta_1}(Z)$ if his actual type θ_1 belongs to Δ_1, and $C_{\theta'_2}(Z)$ if it belongs to Δ'_1, which indeed coincides with $\mu(\theta_1, \theta'_{-1})$. Let's now consider the second individual. Notice that the set of options that he can generate by choosing his message, while others report θ'_{-2} is $\{f(\theta'_2, \theta'_{-2}) | \theta'_2 \in \Theta_2\} = \{C_{\theta'_1}(Z)\}$ if $\theta'_1 \in \Delta_1$, and Z if $\theta'_1 \in \Delta'_1$. The second individual has no freedom to choose in the former case, and in particular reporting his true type is a consistent with a dominant strategy. In the latter case, his choice out of Z when of type θ_2 coincides with $\mu(\theta'_{-2}, \theta_2)$, and hence reporting the true type is also consistent with a dominant strategy, as desired. The set of options that he can generate by *choosing his message, while others report* (m, θ'_{-12}) *is* $\{f(m, \theta'_2, \theta'_{-12}) | \theta'_2 \in \Theta_2\} = Z'$. *His choice out of that set is* $C_{\theta_2}(Z') = (C_{\theta_2}(Z))'$ *if his actual type is* θ_2, *which indeed coincides with* $\mu(m, \theta_2, \theta'_{-1})$. *Other individuals cannot influence the outcome of the mechanism, and hence reporting their true types is a dominant strategy. We have thus proved that* f *is indeed implementable in dominant strategies. Notice that the sufficient condition in Proposition 10 does not apply, and hence is not necessary, since* $X_1^*(\theta_{-1}) = Z$ *and* $C_{\theta_1}(Z) \neq f(\theta)$ *when* $\theta_1 \in \Delta'_1$ *and* $C_{\theta_2}(Z) \neq C_{\theta_1}(Z)$.

Remark 7 *The mechanisms used to implement* f *in Examples 3 and 6 are not the same. Particularly, the mechanism defined in Example 3 does not implement* f *in dominant strategies, and the mechanism proposed in Example 6 does not Nash-implement* f. *This is not a coincidence - one can show that there is no single mechanism that implements* f *both in Nash equilibrium and in dominant strategies (cf. Saijo et al.'s (2007) notion of secure implementation).*

4.3 Limited Attention with Fixed Consideration

This subsection is devoted to the study of implementation in dominant strategy for the class of models introduced in Section 3.3. Notations will be borrowed from there.

Lemma 3 *If* f *satisfies property P, then it satisfies condition (4).*

Proof: Let $i \in I$, $\theta_i, \theta'_i \in \Theta_i$, and $\theta_{-i} \in \Theta_{-i}$. By Property P, $f(\theta'_i, \theta_{-i}) = C_{\theta'_i}(X_i(\theta_{-i}))$, and hence $f(\theta'_i, \theta - i) \in A_i(X_i(\theta_{-i}))$. Property P also implies that $f(\theta) = C_{\theta_i}(X_i(\theta_{-i}))$, and hence it must be that $f(\theta) \succ_{\theta_i} f(\theta'_i, \theta_{-i})$ or $f(\theta) = f(\theta'_i, \theta_{-i})$. ■

Proposition 12 *Let* f *be a social choice function whose range contains at least three options. Then* f *is implementable in dominant strategies if and only if there exists* $j \in I$ *and* $S \in P(X)$ *such that* $Range(f) = A_j(S)$ *and* $f(\theta) = C_{\theta_j}(S)$, *for all* $\theta \in \Theta$, *in which case* j *is a dictator in the strong sense.*

Proof: The sufficient condition, namely that such f's are implementable in dominant strategies, is easy to check. As for the necessary condition, f must satisfy condition (4) if it is implementable in dominant strategies. A fortiori it must be Maskin monotonic (see Theorem 3.2.1 in Dasgupta et al. (1979)), and the proof of Proposition 3 can then be replicated almost word for word, using Property P instead of M. ■

We also have an analogue to Corollary 1

Corollary 2 *Let f be a social choice function in the problem with limited attention that is Pareto efficient. Then f is implementable in dominant strategies if and only if there exists $j \in I$ such that $A_j(X) = X$ and $f(\theta) = C_{\theta_j}(X)$, for all $\theta \in \Theta$, in which case j is a strong dictator. In particular, if no individual pays attention to all options when they are all feasible, then there is no social choice function that is implementable in dominant strategies and Pareto efficient.*

Proof: The result follows at once from the previous Proposition, in the same way Corollary 1 followed from Proposition 3. ■

4.4 Choice from Lists

I now follow the notations from Section 3.4. I start by establishing that f^l is also implementable in dominant strategies.

Proposition 13 *f^l is implementable in dominant strategies on any domain that satisfies condition (3).*

Proof: For each i and each θ, let $k^l_{-i}(\theta) = \min\{k \in \{1, \ldots, |X|\} | (\exists j \in I \setminus \{i\}) : x_k = C_{\theta_j}(X)\}$. Observe that $k^l_{-i}(\theta)$ is independent of θ_i. Hence $\{x_k \in X | k \leq k^l_{-i}(\theta)\}$ is also independent of θ_i, and will thus be denoted $X_i(\theta_{-i})$. Let's check that $f^l(\theta) = C_{\theta_i}(X_i(\theta_{-i}))$, for all θ. Suppose first that $k^l(\theta) = k^l_{-i}(\theta)$. Then $C_{\theta_i}(X)$ does not come before $x_{k^l_{-i}(\theta)}$ in the list, and $C_{\theta_i}(X_i(\theta_{-i})) = x_{k^l_{-i}(\theta)}$, given the assumption on choice functions. Hence $C_{\theta_i}(X_i(\theta_{-i})) = x_{k^l(\theta)} = f^l(\theta)$, as desired. Suppose next that $k^l(\theta) < k^l_{-i}(\theta)$. Hence $C_{\theta_i}(X)$ comes before $x_{k^l_{-i}(\theta)}$ in the list, and $C_{\theta_i}(X_i(\theta_{-i})) = C_{\theta_i}(X) = x_{k^l(\theta)}$, given the assumption on choice functions. Again, $C_{\theta_i}(X_i(\theta_{-i})) = f^l(\theta)$, as desired.

The result from the previous paragraph also implies that $\{f^l(\theta_i, \theta_{-i}) | \theta_i \in \Theta_i\} = \{C_{\theta_i}(X_i(\theta_{-i})) | \theta_i \in \Theta_i\} = X_i(\theta_{-i})$. Hence Proposition 10 applies, and we have established that f^l is implementable in dominant strategies. ■

I now establish a weak converse. For each $i \in I$, let $\bar{\Theta}_i$ be the domain of all choice functions that satisfy condition (3). Of particular interest is the case where choice functions can be derived by the 'stop when you start to decline' procedure. Let Θ_D be the subsets of choice functions that can be generated by this procedure (varying the underlying preference to which it is applied). We will focus on (the more restrictive concept of) implementation in dominant strategies via simple mechanisms. Simple mechanisms allow each individual to send a message which is an option in X, i.e. $M_i = X$. Otherwise, the concept of implementation is the same as in Section 2.

Proposition 14 *Let f be a social choice function defined over a domain $\Theta \subseteq \bar{\Theta}$ that includes Θ_D. If f is BR-efficient, anonymous, and implementable in dominant strategies via a simple mechanism, then $f = f^l$.*

Proof: Let Θ_i^* be a set encoding all the choice functions associated to the maximization of a preference ordering that is single-peaked on the list. Notice that applying the "stop when you start to decline" procedure to a single-peaked preference delivers the same choice function as by maximizing it. Hence we may think of Θ_i^* as a subset of Θ_D and thus also of Θ. The simple mechanism will implement f in dominant strategies on Θ^*. By Moulin (1980), we know that the must exist $I - 1$ elements of X (interpreted as choices of "phantom" voters) such that f coincides on Θ^* with the median of these points and $\{C_{\theta_i}(X)|i \in I\}$. The unique dominant strategy for type $\theta_i \in \Theta_i^*$ is to send the message $x = C_{\theta_i}(X)$. All the elements of X are thus already used by types in Θ_i^*, and hence, for any set of messages, f must coincide with the median of these messages and the phantom voters' choices. I now conclude the proof by showing that all the phantom voter's choices must fall on the first element of the list. Suppose not. Then it is not difficult to check that one can select messages for $I - 1$ individuals such that, combined with the phantom voters' choices, the set of medians that the missing voter can generate by changing his message constitutes a connected subset Y of the list that does not contain the first element of the list. Consider now an individual i, and a type θ_i^* for that individual such that the first element of the list is best for him, the second one is worst, and each element thereafter in the list is better than the previous. If the individuals other than i send those messages associated to Y, then i's best response is to send a message that coincides or follows the last element of Y. On the other hand, if all the other individuals send the message that coincides with the first element of the list, then i's best response is to send that same message. Hence the simple mechanism has no dominant strategy if phantom voters' choices do not all fall on the first element of the list - a contradiction. Hence all these choices do fall on the first element, and the mechanism must select the left-most message among those that have been sent. Since $\Theta \subseteq \bar{\Theta}$, sending $C_{\theta_i}(X)$ is a best response for each type $\theta_i \in \Theta_i$ and each i, and hence $f = f^l$. ■

Remark 8 *Another interesting subclass of $\bar{\Theta}$ is the class of (rational) choice procedures associated to the satisficing procedure. It is not difficult to check that there are* *some choice functions that can be derived by the maximization of a single-peaked preference, but not via the satisficing procedure. For instance, suppose that the peak falls on the second element of the list, and the first element of the list happens to be second-best. The choice out of the two first elements of the list being the peak, it must be that the first element falls below the threshold of satisfaction. Satisficing then implies that the choice out of the first and third elements of the list is the third, which does not coincide with the maximization of the original preference ordering. Hence the previous proof does not apply to the case where* Θ_D *is replaced by the set of types that encode choice functions that can be rationalized via the satisficing procedure, and it remains an open question whether the previous proposition remains valid in that case. Notice also that other social choice functions that are implementable in dominant strategies on the class of single-peaked preferences (e.g. the median with no phantom voter) are not implementable when individuals follow the satisficing procedure instead.*

Remark 9 f^l *is BR-efficient, since it coincides with the choice out of* X *for at least one individual. BR-efficiency was used in the previous proof only to apply Moulin (1980), and so the previous result also holds when BR-efficiency is replaced by the weaker requirement that* f *is Pareto efficient at each type profile for which all individuals are rational, or by any other condition that guarantees this.*

5 Concluding Observation: Endogenous Frames and Backward Induction

Choices can sometimes be influenced by external conditions. For instance, the meal picked in a cafetaria may vary with the order in which options are displayed, or the level of a person's retirement savings may depend on the level of a default rate. This idea is captured theoretically by the concepts of "frames" (see Rubinstein and Salant (2008)) or "anscillary condition" (see Bernheim and Rangel (2009)). As suggested in various examples that have percolated in more popular medias (see the original papers by Camerer et al. (2003) and Thaler and Sunstein (2003), as well as Thaler and Sunstein for a popular book on the topic), the external conditions can sometimes be chosen by the mechanism designer and, if so, he must do it wisely.

Formally, the model introduced in Section 2 can be amended as follows to encompass this added flexibility. Let Φ be a finite set of frames that the mechanism designer must choose from (e.g. the default saving rate for retirement, the list in which to present items, an initial allocation of houses in the house allocation problem, etc.). Choice functions are now indexed by frames: $C^{\phi}_{\theta_i} : P(X) \to X$, for each i and each θ_i, with $C^{\phi}_{\theta_i}(S) \in S$, for all $S \in P(X)$.[16] A social choice function $f : \Theta \to X$ is implementable in Nash equilibrium (resp. dominant strategies) if there exists a frame $\phi \in \Phi$ such that f is implementable in Nash equilibrium (resp. dominant strategies) in the problem (Θ, C^{ϕ}). It should be clear that all the results I have described so far extend to the problem after the frame has been fixed. Adding frames thus expands the set of social choice functions that the mechanism designer can implements, but does not require additional theoretical work.

Since individuals must know the rules of the mechanism before playing it, frames cannot vary with the individuals' types. It is well-known that considering games that unfold over time expands the set of social choice functions that are implementable (in subgame-perfect equilibrium). While studying subgame-perfect implementation with bounded rationality goes beyond the scope of the present paper, I wish to conclude it by observing that considering dynamic games in the present model brings the added possibility of allowing frames to vary endogenously with underlying types. To make this point as clear and simple as possible, I will consider only problems involving an individual who is "fully" rational (see footnote 16) in the sense of satisfying IIA and being unaffected by frames (e.g. a profit maximizing firm or the head cook of a school cafetaria), and another who need not be. I will also restrict attention to two-stage mechanisms of perfect information where the mechanism designer systematically delegates the choice of the frame to the first individual (while obviously one could also consider mechanisms where the designer sets the frame). Formally, a *two-stage mechanism* is composed by a finite set M_1 of the first individual, a finite set $M_2(m_1)$ of messages for the second individual, for each $m_1 \in M_1$, and a function μ that associates an outcome in X to each pair of messages. The extensive-form associated with this mechanism starts with the first individual choosing an action $(m_1, \phi) \in M_1 \times \Phi$, and ends with the second individual choosing afterwards an action in $m_2 \in M_2(m_1)$. The outcome is then $\mu(m_1, m_2)$. The *backward induction equilibrium* at θ of such a

[16] Bounded rationality now acquires a new dimension, namely the dependance on the frames. As an illustration, notice that an individual's choice may satisfy IIA, and thus be rationalizable by a complete transitive preference relation for each frame, but be boundedly rational because these choices (and the underlying preferences) may change with the frame. Many examples of endowment effects fall in this category. Applying Rubinstein and Salant's (2006) version of Simon's (1955) satisficing procedure also falls in this category.

mechanism is a couple $(m_1^*(\theta), m_2^*(\theta))$ of messages and a frame $\phi^*(\theta)$ such that

$$\mu(m_1^*(\theta), m_2^*(\theta)) = C_{\theta_2}^{\phi^*(\theta)}(\{\mu(m_1^*(\theta), m_2)|m_2 \in M_2(m_1^*(\theta))\})$$

$$\mu(m_1^*(\theta), m_2^*(\theta)) = C_{\theta_1}(C_{\theta_2}^{\phi}(\{\mu(m_1, m_2)|m_2 \in M_2(m_1)\})|m_1 \in M_1, \phi \in \Phi\}).$$

The social choice function $f : \Theta_1 \times \Theta_2 \to X$ is *implementable by backward induction* if there exists a mechanism $(M_1, M_2(\cdot), \mu)$ such that, for each θ, the backward induction equilibrium $(m_1^*(\theta), m_2^*(\theta), \phi^*(\theta))$ is such that $f(\theta) = \mu(m_1^*(\theta), m_2^*(\theta))$. It is easy to characterize the set of social choice functions that are implementable in this sense.

Proposition 15 *The social choice function f is implementable by backward induction if and only if there exists a sequence $(S_k)_{k=1}^K$ of non-empty subsets of X such that $f(\theta) = C_{\theta_1}(X_1(\theta))$, where $X_1(\theta) = \{C_{\theta_2}^{\phi}(S_k)|k = 1, \ldots, K, \phi \in \Phi\}$, for each $\theta \in \Theta$.*

Proof: To establish the sufficient condition, let $M_1 = \{m_1, \ldots, m_K\}$. Let $M_2(m_k)$ be a set of $|S_k|$ messages, and μ be such that $\{\mu(m_k, m_2)|m_2 \in M_2(m_k)\} = S_k$, for each k. It is easy to check that the resulting two-stage extensive-form game implements f by backward induction.

As for the necessary condition, let f be a social choice function that is implementable by backward induction via some mechanism $(M_1, M_2(\cdot), \mu)$. Let K be the total number of messages in M_1, and enumerate them so that $M_1 = \{m_1, \ldots, m_K\}$. Let then $S_k = \{\mu(m_k, m')|m' \in M_2(m_k)\}$, for each k. The two conditions for $(M_1, M_2(\cdot), \mu)$ to implement f then imply that $f(\theta) = C_{\theta_1}(X_1(\theta))$, where $X_1(\theta) = \{C_{\theta_2}(S_k)|k = 1, \ldots, K, \phi \in \Phi\}$, for each $\theta \in \Theta$, as desired. ■

The following example illustrates how implementation by backward induction allows for frames to vary with types, which was impossible when considering implementation in Nash equilibrium or in dominant strategies.

Example 7 *Consider the following variant of the arbitrator selection problem studied by de Clippel et al. (2011) for the selection of arbitrators. There are five candidates to be appointed as arbitrator for a case: $X = \{a, b, c, d, e\}$. Types encode strict preference orderings for both parties, but while the first chooses by maximizing that ordering, candidates are always presented in a list, and the second individual makes his choice by applying Rubinstein and Salant's (2006) "Stop When You Start To Decline" procedure introduced in Section 3.4. Here is a variant with frames of de Clippel et al.'s (2011) shortlisting mechanism. The first party can choose an order for the candidates,* *and any one subset of X that contains three candidates. Candidates in the subset selected by the first individual are then presented in a list that is consistent with the order chosen by the first individual, and the second individual is free to appoint any candidate from that list to rule the case. The mechanism implements the social choice function f, where $f(\theta) = C_{\theta_1}(X_1(\theta))$, with $X_1(\theta) = \{C^{\phi}_{\theta_2}(S)|S \in \Sigma, \phi \in \Phi\}$, for each $\theta \in \Theta$, where Σ is the set of subsets of X that contains three candidates, and Φ is the set of lists over X. It is easy to check that this social choice function picks the first individual's top candidate whenever that candidate is not the worse option for the second party, and the first individual's second-best candidate otherwise. Frames at equilibrium change with the second individual's type, with the first individual's top candidate appearing first on the sortlist and the worst candidate for the second individual appearing second, when the first individual's top candidate is not the worst for the second party, and the first individual's second-best candidate appearing first on the shortlist and his most preferred candidate appearing second, when the first party's top candidate is the worst for the other individual. It is also easy to check that there is no frame that the mechanism designer can pick in order to implement this social choice function either in Nash equilibrium (because it violates property M for any list), in dominant strategies (because it violates property P for any list), or by backward induction (there is no way to satisfy the condition of Proposition 15 when ϕ is fixed for all θ in the definition of $X_1(\theta)$ instead of being a variable).*

References

Abdulkadiroglu, A. and T. Sonmez, 1998. Random Serial Dictatorship and the Core from Random Endowments in House Allocation Problems. *Econometrica* **66**, 689-701.

Aizerman, M. A., and F.T. Aleskerov, 1986. Voting operators in the space of choice functions. *Mathematical Social Sciences* **11**, 201-242.

Ambrus, A., and K. Rozen, 2009. Rationalizing Choice with Multi-Self Models. Cowles Foundation Discussion Paper.

Aumann, R., and A. Brandenburger, 1995. Epistemic Conditions for Nash Equilibrium. *Econometrica* **63**, 1161-1180.

Bade, S., 2008. Housing Problems with Non-Rationalizable Behavior. *Mimeo.*

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

Baumeister, R. F., and J. Tierney, 2011. Willpower: Rediscovering the Greatest

Human Strength. Penguin Press.

Cabrales, A., and R. Serrano, 2011a. Implementation in Adaptative Better-Response Dynamics: Towards a General Theory of Bounded Rationality in Mechanisms. *Games and Economic Behavior* **73**, 360-374.

Cabrales, A., and R. Serrano, 2011b. Stochastically Stable Implementation. *Journal of the Spanish Economic Association*, forthcoming.

Camerer, C., S. Issacharoff, G. Loewenstein, T. O'Donoghue, and M. Rabin, 2003. Regulation for Conservatives: Behavioral Economics and the Case for "Asymmetric Paternalism." University of Pennsylvania Law Review **151**, 1211-1254.

Cherepanov, V., T. Feddersen, and A. Sandroni, 2009. Rationalization. Working paper, Northwestern University.

Corchón, L., 1996. The Theory of Implementation of Socially Optimal Decisions in Economics. St. Martin's Press, New York.

Crawford, V. P., T. Kugler, Z. Neeman, and A. Pauzner, 2009. Behaviorally Optimal Auction Design: Examples and Observations. *Journal of the European Economic Association* **7**, 377-387.

Dasgupta, P. S., P. J. Hammond, and E. Maskin, 1979. The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility. *Review of Economics Studies* **46**, 185-216.

de Clippel, G., and K. Eliaz, 2009. Reason-Based Choice: A Bargaining Rationale for the Attraction and Compromise Effects. *Theoretical Economics*, forthcoming.

de Clippel, G., K. Eliaz, and B. Knight, 2011. On the Selection of Arbitrators. *Mimeo.*

Eliaz, K., 2002. Fault-Tolerant Implementation. *Review of Economic Studies* **69**, 589-610.

Glazer, J., and A. Rubinstein, 2011. A Model of Persuasion with a Boundedly Rational Agent. *Mimeo.*

Gul, F., and W. Pesendorfer, 2001. Temptation and Self-Control. *Econometrica* **69**, 1403-1435.

Huber, J., J. W. Payne, and C. Puto, 1982. Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Slimilarity Hypothesis. *Journal of Consumer Research* **9**, 90-98.

Hurwicz, L., 1986. On the Implementation of Social Choice Rules in Irrational Societies, in Heller et al., editors, Essays in Honor of Kenneth J. Arrow, Vol. I.

Iyengar, S., W. Jiang, and E. Kamenica, 2006. The Psychological Cost of Ever Increasing Choice: A Fallback to the Sure Bet. Working Paper, Graduate School of Business, Columbia University.

Jackson, M. O., 2001. A Crash Course in Implementation Theory. *Social Choice and Welfare* **18**, 655-708.

Kalai, E., E. Muller, and M. A. Satterthwaite, 1979. Social welfare functions when preferences are convex, strictly monotonic, and continuous. *Public Choice* **34**, 87-97.

Kalai, E., and D. Schmeidler, 1977. An admissible set occurring in various bargaining situations. *Journal of Economic Theory* **14**, 402-411.

Kalai, G., A. Rubinstein, and R. Spiegler, 2002. Rationalizing Choice Functions by Multiple Rationales. *Econometrica* **70**, 2481-2488.

Korpela, V., 2011. Implementation without Rationality Assumptions, *Theory and Decision*, forthcoming.

Lipman, B., 2011. Temptation. *Mimeo.*

Lleras, J. S., Y. Masatlioglu, D. Nakajima, and E. Ozbay, 2010. When More is Less: Choice by Limited Consideration. *Mimeo.*

Manzini, P., and M. Mariotti, 2007. Sequentially Rationalizable Choice. *American Economic Review* **97**, 1824-1839.

Manzini, P., and M. Mariotti, 2009. Categorize Then Choose: Boundedly Rational Choice and Welfare. *Journal of the European Economic Association*, forthcoming.

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

Masatlioglu, Y., D. Nakajima, and E. Ozbay, 2009. Revealed Attention. *American Economic Review*, forthcoming.

Masatlioglu, Y., D. Nakajima, and E. Ozdenoren, 2011. Revealed Willpower. *Mimeo.*

Maskin, E., 1985. The Theory of Implementation in Nash Equilibrium: A Survey. In Social Goals and Social Organization, L. Hurwicz, D. Schmeidler, and H. Sonnenschein, eds., Cambridge University Press, Cambridge, UK, 173-204.

Maskin, E., 1999. Nash Equilibrium and Welfare Optimality. *Review of Economic Studies* **66**, 23-38.

Maskin, E., and T. Sjöström, 2002. Implementation Theory. In Handbook of Social Choice and Welfare, Vol. I, K. J. Arrow, A. Sen, and K. Suzumura, eds., Elsevier Science, New York, 237-288.

Mookherjee, D., and S. Reichelstein, 1990. Implementation via Augmented Revelation Mechanisms, *Review of Economic Studies* **57**, 453-475.

Moon, J., 1968. Topics on Tournaments. Holt, Rinehart and Winston, New York.

Moore, J., 1992. Implementation, Contracts and Renegotiation in Environments with Complete Information. In Advances in Economic Theory, 4th World Congress of the Econometric Society, Vol. I, J. J. Laffont, ed., Cambridge University Press, Cambridge, UK, 182-282.

Moore, J., and R. Repullo, 1990. Nash Implementation: A Full Characterization. *Econometrica* **56**, 1191-1220.

Moulin, H., 1980. On Strategy-Proofness and Single-Peakedness. *Public Choice* **35**, 437-455.

Ozdenoren, E., S. Salant, and D. Silverman, 2010. Willpower and the Optimal Control of Visceral Urges. *Journal of the European Economic Association*, forthcoming.

Palfrey, T. R., 1992. Implementation in Bayesian Equilibrium: The Multiple Equilibrium Problem in Mechanism Design. In Advances in Economic Theory, 4th World Congress of the Econometric Society, Vol. I, J. J. Laffont, ed., Cambridge University Press, Cambridge, UK, 283-323.

Palfrey, T. R., 2002. Implementation Theory. In Handbook of Game Theory with Economic Applications, Vol. III, R. J. Aumann and S. Hart, eds., Elsevier Science, New York, 2271-2326.

Ray, K. T., 2010. Nash Implementation under Irrational Preferences. *Mimeo*, McDonough School of Business, Georgetown University.

Roberts, J. H., and J. M. Lattin, 1991. Development and Testing of a Model of Consideration Set Composition. *Journal of Marketing Research* **28**, 429-440.

Rubinstein, A., and Y. Salant, 2006. A Model of Choice from Lists. *Theoretical Economics* **1**, 3-17.

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

Saijo, T., T. Sjöström, and T. Yamato, 2007. Secure Implementation. *Theoretical Economics* **3**, 203-229.

Saran, R., 2011. Menu-Dependent Preferences and Revelation Principle. *Journal of Economic Theory*, forthcoming

Serrano, R., 2003. The Theory of Implementation of Social Choice Rules. *SIAM Review* **46**, 377-414.

Shapley, L., and H. Scarf, 1974. On Cores and Indivisibility. *Journal of Mathematical Economics* **1**, 23-37.

Simon, H. A., 1955. A behavioral model of rational choice. *Quarterly Journal of Economics* **69**, 99-118.

Spiegler, R., 2011. Bounded Rationality and Industrial Organization. *Oxford University Press.*

Thaler R. H., and C. R. Sunstein, 2003. Libertarian Paternalism. *American Economic Review* **93**, 175-179

Thaler R. H., and C. R. Sunstein, 2008. Nudge: Improving Decisions About Health, Wealth, and Happiness. *Yale University Press.*

Wright, P., and F. Barbour, 1977. Phased Decision Strategies: Sequels to an Initial Screening. In Studies in Management Sciences, Multiple Criteria Decision Making, ed. Martin K. Starr, and Milan Zeleny, 91-109, Amsterdam: North-Holland.

Appendix

Implementability of f in Example 3

It is sufficient to prove that f satisfies Property $\mathfrak{M}$ (or almost), and apply Proposition 1 (or a minor variant of it – see the very end of this proof). Define $X_1 : \Theta \to P(X)$ as follows:

$$X_1(\theta) = \begin{cases} Z \cup \{(\sigma(C_{\theta_1}(Z)))'\} & \text{if } \theta_1 \in \Delta_1 \\ Z \cup \{(C_{\theta_2}(Z))'\} & \text{if } \theta_1 \in \Delta_1', \end{cases}$$

where $\sigma : Z \to Z$ is the function that associates to each letter in Z the next letter in the alphabet, with the convention that a follows c. Let $X_2(\theta) = Z$ if $\theta_1 \in \Delta_1'$, $X_2(\theta) = \{f(\theta)\}$ if $\theta_1 \in \Delta_1$, and let $X_i(\theta) = \{f(\theta)\}$ for all θ and all $i > 2$.

The first property in M is easy to check. Let's focus on the second property. For this, let θ, θ' be such that $C_{\theta_i'}(X_i(\theta)) = f(\theta)$, for each $i \in I$. If $\theta_1' \in \Delta_1$, then

$$f(\theta') = C_{\theta_1'}(Z) = C_{\theta_1'}(X_1(\theta)) = f(\theta),$$

where the first equality follows from the definition of f, the second equality follows from the fact that $X_1(\theta)$ is equal to Z plus an element of Z', out of which a rational individual picks the same element as when choosing in Z, and the third equality follows by assumption. Hence $f(\theta') = f(\theta)$, as desired. Suppose now that $\theta_1' \in \Delta_1'$. If $\theta_1 \in \Delta_1$, then $C_{\theta_1'}(X_1(\theta)) = \sigma(f(\theta)) \neq f(\theta)$, and hence there is nothing to prove in this case. If $\theta_1 \in \Delta_1'$, then $X_2(\theta) = Z$, and hence $C_{\theta_2'}(X_2(\theta)) = f(\theta)$ means $C_{\theta_2'}(Z) = f(\theta)$. Since $\theta_1' \in \Delta_1'$, $f(\theta') = C_{\theta_2'}(Z)$, and transitivity implies that $f(\theta') = f(\theta)$, as desired. This concludes the proof of the second property in M. We now check the additional property that distinguishes $\mathfrak{M}$ from M. Let thus $j \in I$, $x \in X$, and $\theta, \theta' \in \Theta$ be such that $C_{\theta_i}(X) = x$, for each $i \in I \setminus \{j\}$, and either $C_{\theta_j}(X) = x$ or $C_{\theta_j}(X_j(\theta')) = x$. Suppose first that $j = 1$ and $C_{\theta_1}(X_1(\theta')) = x$. If $\theta_1 \in \Delta_1$, then $f(\theta) = C_{\theta_1}(Z) = C_{\theta_1}(X_1(\theta)) = x$, and the property thus holds. Consider now the case where $\theta_1 \in \Delta_1'$. The condition $C_{\theta_1}(X_1(\theta')) = C_{\theta_2}(X) = x$ implies that $\theta_2 \in \Delta_2$ (as otherwise $C_{\theta_2}(X) = d$ while $C_{\theta_1}(X_1(\theta')) \in Z$), and hence $C_{\theta_2}(Z) = x$. Since $f(\theta) = C_{\theta_2}(Z)$, by definition of f, we conclude that $f(\theta) = x$, as desired. Suppose now that $j \geq 2$. If $\theta_1 \in \Delta_1$, then $C_{\theta_1}(X) = x$ implies that $C_{\theta_1}(Z) = x$, by definition of Δ_1. Since $f(\theta) = C_{\theta_1}(Z)$, by definition of f, we conclude that $f(\theta) = x$, as desired. Suppose now that $\theta_1 \in \Delta_1'$. In that case $C_{\theta_1}(X) = d$, while $C_{\theta_j}(X_j(\theta')) \in Z$.

Hence the property is trivially satisfied in that case, since no x satisfying its premises can be found. The last case to be considered is when $C_{\theta_i}(X) = x$, for all $i \in I$. Here Property $\mathfrak{M}$ is not systematically satisfied, as $C_{\theta_i}(X) = d$ when $\theta_i \in \Delta_i'$, for all $i \in I$, and yet $f(\theta) \neq d$. It is easy to adapt the mechanism in the proof of Proposition 1 to show that f is nevertheless implementable. Specifically, simply change the mechanism in the third case of its definition by having the outcome equal to a if the third individual is the one quoting the largest integer (instead of picking the first component of his report). Notice that this slight modification indeed prevents d from being an equilibrium outcome when $\theta_i \in \Delta_i'$, for all i (as the third individual never wants to pick d given the limited set of alternatives he can generate), while not changing any of the other properties of the mechanism. We can thus conclude that f is Nash implementable. ■

Proof of Lemma 1

Let x, y, z be any three elements in the range of f, let Θ_i^{xyz} be the set of all six rational strict orderings on $\{x, y, z\}$, and let $\alpha_i^{xyz} : \Theta_i^{xyz} \to \Theta_i$ be a function that associates a type $\theta_i \in \Theta_i$ to each $\succ_i \in \Theta_i^{xyz}$ such that x is most preferred according to C_{θ_i}, y is second most preferred, and z is third most preferred. Next consider the social choice function $\phi^{xyz} : \Theta^{xyz} \to X$ defined as follows:

$$\phi^{xyz}(\succ) = f(\alpha^{xyz}(\succ)),$$

for each $\succ \in \Theta^{xyz}$ (with the convention $\alpha^{xyz}(\succ) := (\alpha_i^{xyz}(\succ_i))_{i \in I}$). The proof of this lemma now proceeds in various steps.

Step 1 *Range*$(\phi^{xyz}) = \{x, y, z\}$.

Proof: For each $a \in \{x, y, z\}$, consider the profile of strict orderings $\succ$ on $\{x, y, z\}$ such that a is top-ranked by all individuals. Hence a is most-preferred by all individuals at $\alpha^{xyz}(\succ)$, and BR-efficiency implies that $\phi^{xyz}(\succ) = f(\alpha^{xyz}(\succ)) = a$. Hence $\{x, y, z\} \subseteq$ *Range*(ϕ^{xyz}). We prove the opposite inclusion by contradiction. By definition of most preferred, second most preferred, and third most preferred elements, any element in $\{x, y, z\}$ unambiguously dominates any element in $X \setminus \{x, y, z\}$ at type profiles in the image of α^{xyz}. The definition of ϕ^{xyz} and the fact that f is BR-efficient then imply that *Range*$(\phi^{xyz}) \subseteq \{x, y, z\}$ as well, and hence *Range*$(\phi^{xyz}) = \{x, y, z\}$, as desired. □

Step 2 ϕ^{xyz} *is Maskin monotonic.*

Proof: Suppose that $\phi^{xyz}(\succ) = a \in \{x, y, z\}$, and that a's rank increases for each individual when moving from $\succ$ to $\succ'$. $a = \phi^{xyz}(\succ) = f(\alpha^{xyz}(\succ))$ and the first part of Property M imply that $C_{\alpha_i^{xyz}(\succ_i)}(X_i(\alpha^{xyz}(\succ))) = a$, for each $i \in I$. The definition of $\succ'$ implies that $C_{\alpha_i^{xyz}(\succ'_i)}(X_i(\alpha^{xyz}(\succ))) = a$, for each $i \in I$. The second part of property M implies that $f(\alpha^{xyz}(\succ')) = a$, and hence $\phi^{xyz}(\succ') = a$, as desired. $\square$

Step 3 *There exists a unique j such that, for all a, b, c in A and all $\succ \in \Theta^{abc}$, $\phi^{abc}(\succ)$ is top ranked for $\succ_j$.*

Proof: By the previous step and the usual result on the rational domain, for each x, y, z, there exists a unique j such that $\phi^{xyz}(\succ) = top(\succ_j)$, for each $\succ \in \Theta^{xyz}$. We just need to prove that j does not vary with x, y, z. It is obviously sufficient to show that x, y, z and x, y, z' lead to the same j. Suppose, on the contrary, that the first triplet leads to j, while the second leads to j', with $j \neq j'$. Consider then $\succ \in \Theta^{xyz}$ such that x is top ranked according to $\succ_j$, y is top ranked according to $\succ_i$, and x is second best according to $\succ_i$, for each $i \in I \setminus \{j\}$. By definition of j, we must have $\phi^{xyz}(\succ) = x$. The first part of property M implies that $x \in X_i(\alpha^{xyz}(\succ))$ and $y \notin X_i(\alpha^{xyz}(\succ))$, for each $i \in I \setminus \{j\}$. Let $\succ' \in \Theta^{xyz'}$ be derived from $\succ$ by replacing z by z'. We have that $C_{\alpha_i^{xyz'}(\succ'_i)}(X_i(\alpha^{xyz}(\succ))) = x$, for all $i \in I$, and hence $\phi^{xyz'}(\succ') = x$, by the second part of Property M and the definition of $\phi^{xyz'}$. This implies a contradiction with the definition of j', as desired. $\square$

Step 4 *There exists $j \in I$ such that $f(\theta) = a$, for each $a \in Range(f)$ and each $\theta \in \Theta$ such that a is most preferred according to C_{θ_j}.*

Proof: Let j be the individual as identified in the previous step, and let a and θ be as in the statement. Suppose, to the contrary of what we want to prove, that $f(\theta) = b \neq a$. The first part of Property M implies that $a \notin X_j(\theta)$. Let c be a third element in the range of f, and let $\succ \in \Theta^{abc}$ be such that a is top ranked according to $\succ_j$, b is second best according to $\succ_j$, and b is top ranked according to $\succ_i$, for each $i \in I \setminus \{j\}$. The second part of Property M implies that $\phi(\succ) = f(\alpha(\succ)) = b$, thereby contradicting the definition of j. $\square$

Step 5 *There exists $j \in I$ and $X_j : \Theta \to P(X)$ such that $Range(f) \subseteq X_j(\theta)$ and $f(\theta) = C_{\theta_j}(X_j(\theta))$, for each $\theta \in \Theta$.*

Proof: Pick j as identified in the previous step. The existence of the X_j follows from Property M. All what needs to be proved is that $Range(f) \subseteq X_j(\theta)$, for all θ. Suppose, on the contrary, that $x \in Range(f) \setminus X_j(\theta)$, for some $\theta \in \Theta$. So $f(\theta) \neq x$.

Part 2 of Property M then implies that $f(\theta'_j, \theta_{-j}) = f(\theta)$, for any θ'_j such that x is most preferred and $f(\theta))$ is second most preferred according to C_{θ_j}, a contradiction with the previous step. ■

Proof of Proposition 7

As discussed in the second half of Section 3.5, in order to have individual choice functions instead of correspondences, one must consider choices as defined over subsets of O instead of subsets of allocation vectors. Proposition 1 can easily be rewritten as follows.

Lemma 4 *(*Necessary, and Sufficient Conditions for Nash-Implementability of Social Choice Correspondences in the House Allocation Problem*)* *If a social choice correspondence F is Nash-implementable, then there exists a function $X_i : \{(x, \theta) \in X \times \Theta | x \in F(\theta)\} \to P(O)$, for each $i \in I$, such that:*

1. *$C_{\theta_i}(X_i(o, \theta)) = o_i$, for each $o \in F(\theta)$ and each $\theta \in \Theta$.*
2. *For all θ, θ', and all $o \in F(\theta)$, if $C_{\theta'_i}(X_i(o, \theta)) = o_i$, for each $i \in I$, then $o \in F(\theta')$.*

If, in addition, $o \in F(\theta)$ for any θ and any o for which there exists i and o' such that $C_{\theta_j}(O) = o_j$, for all $j \in I \setminus \{i\}$, and either $C_{\theta_i}(O) = o_i$ or $C_{\theta_i}(X_i(o', \theta)) = o_i$, then F is Nash-implementable.

We can now proceed with the proof of Proposition 7. Let's start by showing that F^{BRC} is not Nash-implementable. Consider a type profile θ where each individual $i \neq 1$'s choice function is derived by maximizing the ordering where o_1^* is most preferred, o_{i+1}^* is second most preferred, o_{i+2}^* is third most preferred, etc., with the convention that $I+1 = 2$. The first individual's choice function at that state is derived by maximizing the ordering where o_2^* is most preferred, o_3^* is second most preferred, etc., and o_1^* is least preferred, except that $C_{\theta_1}(\{o_1^*, o_i^*\}) = o_1^*$, for each $i > 1$. The allocation $\bar{o}$ where the first individual keeps his endowment, and each subsequent individual gets the endowment of his follower (with I getting 2's endowment) belongs to the $F^{BRC}(\theta)$. Indeed, no object unambiguously dominates the endowment for the first individual, and hence a blocking coalition cannot contain 1. Notice then that $\bar{o}$ is such that each individual other than the first receives his most preferred object in $O \setminus \{o_1\}$, and hence it is impossible to find a blocking coalition that does not contain 1 either, which shows indeed that $\bar{o} \in F^{BRC}(\theta)$. By definition of θ, the first part of the necessary condition in the previous Lemma can be satisfied only if $\bar{o}_i \in X_i(\bar{o}, \theta)$ and $o_1^* \notin X_i(\bar{o}, \theta)$, for all

$i \neq 1$, and $X_1(\bar{o}_1, \theta) = \{\bar{o}_1\}$, or $\{\bar{o}_1, o_i^*\}$, for some $i \geq 2$. If $X_1(\bar{o}_1, \theta) = \{\bar{o}_1\}$, then the second part of the necessary condition implies that $\bar{o} \in F^{BRC}(\theta_1', \theta_{-1})$, where θ_1' is the rational choice function derived from θ_1 by changing the first individual's choice over pairs that contain o_1^*. To see that is is impossible, observe that $F^{BRC}(\theta_1', \theta_{-1})$ coincides with the regular core since all individual choice functions are rational at that type profile, and hence the core coincides with the outcome of the top-trading cycle procedure, which will have 1 and 2 achieve their most-preferred choice by trading their objects. In the case where $X_1(\bar{o}_1, \theta) = \{\bar{o}_1, o_i\}$, for some $i \geq 2$, consider $j \in I \setminus \{1, i\}$, and the type θ_1' that differs from θ_1 only in that $C_{\theta_1'}(\{o_1^*, o_j^*\}) = o_j^*$ instead of o_1^*. The second part of the necessary condition in Proposition 1' implies that $\bar{o} \in F^{BRC}(\theta_1', \theta_{-1})$ if it were Nash-implementable, but it doesn't, as $\{1, j\}$ can unambiguously improve upon $\bar{o}$ by trading with each others.

I now show that F^C is Nash implementable on any Θ. By definition of F^C, one can associate to any $x \in F^C(\theta)$ a collection $(T_i^{x,\theta})_{i \in I}$ of subsets of O such that $C_{\theta_i}(T_i^{x,\theta}) = x_i$, for each $i \in I$, and $X = \cup_{i \in I} T_i^{x,\theta}$. It is easy to check that the conditions 1. and 2. given in the previous Lemma are satisfied for $X_i(x, \theta) = T_i^{x,\theta}$. Suppose now that θ and $i \in I$ are such that $C_{\theta_j}(O) = o_j$, for each $j \in I \setminus \{i\}$. If $C_{\theta_i}(O) = o_i$ as well, then one concludes that $o \in F^C(\theta)$ by taking $T_j = O$, for all $j \in I$. If $C_{\theta_i}(X_i(\theta)) = o_i$, then one concludes that $o \in F^C(\theta)$ by taking $T_j = O$, for all $j \in I \setminus \{i\}$ and $T_i = X_i(\theta)$ (notice that $o_i^* \in T_i$, by definition of $X_i(\theta)$).

Consider a type profile θ, $x \in F^C(\theta)$, a group S, and an allocation $o' \in \mathcal{F}(S)$. Hence there exists $i \in S$ such that $o_i' \in T_i$ and $C_{\theta_i}(T_i) = x_i$. It is thus impossible to find $o' \in \mathcal{F}(S)$ such that o_i' unambiguously dominates x_i, for all $i \in S$, and $x \in F^{BRC}(\theta)$.

Finally, if θ is such that C_{θ_i} is rational, for each $i \in I$, and x belongs to the core for the associated revealed preferences, then define T_i as the lower contour set of x according to the revealed preference, for each i. By definition, $C_{\theta_i}(T_i) = x$, for all i. If there is S and $x' \in \mathcal{F}(S)$ such that $x_i' \in O \setminus T_i$, for each $i \in S$, then x_i' is revealed preferred to x_i for all $i \in S$, thereby contradicting the fact that x belongs to the core. This shows that F^C contains the core. For the other inclusion, showing that $F^C(\theta)$ is included in the core at θ, consider an option x that does not belong to the core. This means that there exist S and $x' \in \mathcal{F}(S)$ such that x_i' is revealed preferred to x_i for all i. If so, and if $C_{\theta_i}(T_i) = x_i$, for all $i \in S$, then $x_i' \in O \setminus T_i$, for all $i \in S$, and hence $x \notin F^C(\theta)$. ■