Nash Implementation Under Choice Functions

Korok Ray*
Texas A&M University
korok@tamu.edu

Thursday 1st February, 2018

Abstract

This paper extends the standard implementation problem to include behavior not representable by a preference relation. In particular, the social choice correspondence has as its domain the set of profiles of choice functions, not profiles of preferences. The paper proposes a generalized Nash concept in games without preferences, and obtains necessary and sufficient conditions for Nash implementation.

JEL classification codes: C7, D0, D7.

Keywords: Game Theory, Nash Implementation, Behavioral Economics, Social Choice.

*I'd like to thank Kenneth Arrow, Lanier Benkard, Yossi Feinberg, Edward Lazear, Jonathan Levin, Eric Maskin, and Bob Wilson for comments and encouragement. Texas A&M University provided generous financial support.

1 Introduction

Economics is built on a foundation of preferences (in fact, utility functions) instead of more general choice functions. Relaxing preferences to choice functions permits behavior that cannot be represented by a preference relation; in such cases, behavior may exhibit a preference reversal or cycle. This paper asks: what can a social planner implement in a society filled with people with behavioral choice functions rather than rational preferences?

To make traction on this question, begin with primitive choice functions that do not satisfy the standard rationality axioms that give rise to a preference relation. Because the existing literature on game theory, mechanism design, and social choice is based on preference relations, it is necessary to redefine these constructs for more general choice functions. I define games with choice functions rather than games with preference orderings. I define a Nash equilibrium in games with choice functions, as well as the implementation of a social choice rule. Following Maskin [1999],[1] no veto power, and monotonicity imply Nash implementation. However, without any rationality restriction on choice functions, no veto power and a generalized form of monotonicity are no longer sufficient for Nash implementation.

I provide new axioms on the choice functions that guarantee that monotonicity and no veto power imply Nash implementation. I make full use of the fact that choice functions map sets into sets, and therefore the definition of Nash equilibrium should be similarly set-based. Under this new set-based definition, a Nash equilibrium in choices generalizes the Nash equilibrium in preferences that is now common in game theory. Under this broader definition of equilibrium, I give necessary and sufficient conditions for Nash implementation following Moore and Repullo [1990]. These conditions give a full and complete characterization of Nash implementation and these conditions impose restrictions on choice sets rather than restrictions on preferences, as in the standard model.

The claim that preferences cannot always explain human decision making has strong experimental support in the psychology and consumer marketing literatures. Kahneman and Tversky have consistently chipped away at the edifice of rational choice for the last

[1] Maskin's Theorem first appeared in a working paper in 1977, but was later published in 1999 in the Review of Economic Studies. Though the published version is recent, the result has circulated in the theory community for decades.

quarter century (see Kahneman and Tversky [2000] for an anthology of recent papers). In particular, Tversky and Simonson [1993] conduct several experiments to show how choices vary with context. For example, they find that more subjects will choose x when choosing from $\{x, y, z\}$ than from $\{x, y\}$, where z is clearly inferior to x but not to y. One explanation that they propose for such "context effects" is extremeness aversion: people tend to avoid extreme choices (for example, very high or very low quality goods) in *whichever* choice set they face. Extremeness aversion is one of many behavioral explanations of choice, and I will not advance any particular behavioral bias or heuristic here. Rather, my goal is to explore what happens in games without preferences, and to relax the rationality axioms behind preferences just enough to still get meaningful results in implementation.

Once we dissolve the axioms of rational preferences, we need to define equilibrium concepts in games without preferences. There are many directions to move, since rationality naturally constrains agents choices. Part of the papers contribution is to consider alternative, more general equilibrium concepts than the standard model. Specifically, we define three kinds of equilibria: Pair-wise Nash Equilibria (PNE), Set-wise Nash Equilibria (SNE), and Nash Equilibria in Choices(NEC). A Pair-wise Nash Equilibrium is the closest to the standard model, in which an equilibrium outcome for any agent is (weakly) chosen when compared against all of his possible deviations, taking the strategies of all other agents fixed at their equilibrium levels. Set-wise Nash Equilibria introduce a concept of a deviation set, where each agents strategy is chosen within the deviation set of that agent, which can include some non-rational choices. A Generalized Nash Equilibria is the combination of both SNE and PNE. Finally, my last concept is Nash Equilibria in Choices, which is a set-based equilibrium concept exploring the special structure of choice functions that preference relations and utility functions cannot provide. It also permits some consideration of equilibrium selection. With these more general equilibrium concepts, we can then turn to the question of implementation. Hurwicz [1986] begins with the same motivation as this paper, but arrives at quite different results. Hurwicz takes choice functions (instead of preference relations) as primitives of the implementation problem, as I do here. However, Hurwicz's main motivation is to examine implementation in the context of preference cycles that arise from voting problems (the Condercet Paradox, for example). Hence, he confines attention to choice functions with a two-element domain; in other words, he extends the implementation problem not from preference relations to choice functions, but from preference relations to binary relations.[2] His main result is a generalization of Maskin [1999] theorem to a domain of (not necessarily reflexive, complete, and transitive) binary relations. This paper, on the other hand, generalizes Maskin [1999] as well as Moore and Repullo [1990] to a domain of choice functions; in doing so, I propose more general Nash concepts (defined for arbitrary, nonbinary choice functions) that Hurwicz does not consider.

There are two other papers that address a similar problem as this, de Clippel [2014] and Korpela [2012]. Both tackle the general problem, though with different approaches, definitions of equilibrium, and methods. More precisely, Korpela [2012] considers the problem of fully implementing social choice functions in Setwise Nash Equilibrium (called Behavioral Nash equilibrium in his paper) and derives necessary and sufficient conditions *á* la Moore and Repullo [1990]. An essential condition for his characterization to be valid is the axiom of independence to irrelevant alternatives (property α). There are differences between both papers. For instance, Korpela [2012] considers social choice functions, while I consider correspondences. Additionally, I do not directly assume property α, but rather a combination of property (A) about lower contour sets, property δ, and property μ. In contrast to Korpela [2012], I discuss how to extend Maskin's characterization, which used "monotonicity" and "no-veto-power." I show what additional properties on choice functions are required in order for Maskin's properties to be necessary and sufficient. In addition, Korpela, in contrast to me, also addresses the two-agent case (Theorems 3 and 4).

In contrast, de Clippel [2014] proposes a generalization of monotonicity and no-veto-power that are necessary and sufficient for implementation for all choice functions. de Clippel's paper analyzes implementation under some well-known biases and discusses what economically interesting choice rules are implementable. The Nash concepts that appear in Section 3 here also appear in Section 3 of Korpela. Section 5 extends the Moore-Repullo characterization to the choice-function setting. Theorem 4 establishes that if there are more than two agents and the choice function satisfies condition (α), then an extension of the Moore-Repullo condition is necessary and sufficient. Similar results are provided in Section 4 of Korpela; the main difference between the two papers is that I consider a slightly different solution concept, whereby a profile of actions a^* is an equilibrium if no agent can deviate to any action that is feasible given any action profile a', which is also an equilibrium. In addition, Korpela, in contrast to me, also addresses the two-agent case (Theorems 3 and 4). Other research on similar topics

[2] Recall that a preference relation is a complete, reflexive, and transitive binary relation.

include Glazer and Rubinstein [1998] on the implementation problem when agents' preferences may be influenced by the mechanism itself, Eliaz [2002] on the implementation problem when some agents may act unpredictably, Saran [2011] on implementation with menu-dependent preferences.

The paper proceeds as follows: Section 2 defines the basics of choice sets. Section 3 defines games without preferences and defines Nash Equilibria. Section 4 defines Nash implementation, monotonicity, and gives necessary and sufficient conditions for Nash implementation. Section 5 provides an expanded definition of Nash Equilibria, and provides necessary and sufficient conditions for Nash implementation in this setting. Section 6 concludes.

2 Choices

Let X be the finite choice set. Let $\mathcal{X}$ be the set of all nonempty subsets of X. A *choice function* is a map $C : \mathcal{X} \to \mathcal{X}$ such that $\forall E \in \mathcal{X}, C(E) \subseteq E$. Some properties of choice functions include

(α) $\forall E, F \in \mathcal{X},\ E \subseteq F \ \Rightarrow\ C(F) \bigcap E \subseteq C(E)$

(γ) $\forall E, F \in \mathcal{X},\ C(E) \bigcap C(F) \subseteq C(E \bigcup F)$

($\alpha 2$) $\forall E \in \mathcal{X},\ x \in C(E) \ \Rightarrow\ x \in C(xy)\ \forall y \in E \backslash \{x\}$

($\gamma 2$) $\forall E \in \mathcal{X},\ x \in E$ and $x \in C(xy)\ \forall y \in E \backslash \{x\} \ \Rightarrow\ x \in C(E)$

I use the shorthand $C(xy)$ for $C(\{x, y\})$ throughout the paper. (α) and (γ) are also called Contraction and Expansion, respectively. Together they are equivalent to either Arrow's axiom, the weak axiom of revealed preference, or Houthakker's axiom; see Austen-Smith and Banks [1999], pp. 16-19) and (Kreps [1988] pp. 12-16) for definitions and proofs. ($\alpha 2$) and ($\gamma 2$) are weaker versions of (α) and (γ).

Let $\mathcal{B}$ be the set of all binary relations on a choice set X. Let $\mathcal{R} \subset \mathcal{B}$ denote the set of all preference relations (reflexive, complete, and transitive binary relations). For $R \in \mathcal{B}$ let $M(E, R) \equiv \{x \in E : xRy\ \forall y \in E\}$ be the maximal set of E under R. Moreover, for any binary relation R let P be its asymmetric part. Note that for $R \in \mathcal{R}$, R is acyclic, and hence $M(E, R)$ is nonempty for all $E \in \mathcal{X}$.

When can we consider someone making choices *as if* she is maximizing some underlying preference relation? Call a choice function *binary* if $\exists R \in \mathcal{B}: \ \forall E \in \mathcal{X}, \ C(E) = M(E, R)$. For a given choice function C, let R_C be the *base relation* associated with C: $\forall x, y \in X, \ xR_Cy \Leftrightarrow x \in C(xy)$. As usual, P_C is the asymmetric part of R_C. Standard results in single-person choice (see, for example, Austen-Smith and Banks [1999], p. 8) state that [C is binary] iff [$C(E) = M(R_C, E) \ \forall E \in \mathcal{X}$] iff [$C$ satisfies (α) and (γ)]. Let $\mathcal{C}$ be the set of all choice functions over X, and let $\mathcal{C}'$ be the subset of all binary choice functions over X. Hence imposing (α) and (γ) is equivalent to restricting the domain of choice functions to be $\mathcal{C}'$ instead of $\mathcal{C}$. In a precise sense, $\mathcal{C}'$ is equivalent to $\mathcal{B}$:

Proposition 1 *Call two spaces* equivalent *if there exists a bijection between them. Then $\mathcal{C}'$ is equivalent to $\mathcal{B}$, but $\mathcal{C}$ is not equivalent to $\mathcal{B}$.*

All proofs are in the appendix. Economic and game theoretic models usually work with utility functions. In such cases the analyst is implicitly imposing not just (α) and (γ) but additional continuity axioms (see Kreps [1988] p. 27), and if there is uncertainty, either the vNM or Savage axioms. However, not all choice behavior is binary:

Example 1. David and Simon both like chocolate, but they know it is bad for them. Their choices differ in the following way: when asked to choose between a chocolate candy bar and a granola bar, David succumbs to the chocolate while Simon exerts some self-control and takes the granola. But when asked to choose between a granola bar and thirty different candy bars (say, at a vending machine), David is nauseated by the thought of all those calories, and chooses the granola. Simon, on the other hand, loses his control (all that chocolate!) and takes the candy bar, possibly many of them.

Let $X = \{g, x, y\}$, where g=granola bar, and x, y are chocolate bars. The choice functions for (D)avid and (S)imon are

E	$C_D(E)$	$C_S(E)$
g	g	g
x	x	x
y	y	y
gx	x	g
gy	y	g
X	g	xy

These choice functions are both nonbinary, since $x = C_D(gx) \ \Rightarrow \ xP_Dg$ but $g = C_D(X) \ \Rightarrow \ gP_Dx$. Similarly, $g = C_S(gx) \ \Rightarrow \ gP_Dx$ but $xy = C_S(X) \ \Rightarrow \ xP_Dg$. □

3 Games without Preferences

Before tackling the implementation problem, it is necessary to first define games without preferences, and an appropriate equilibrium concept. Let N be the set of players and let X be the outcome space. Let $\mathcal{X}$ be the set of all nonempty subsets of X, and let $C_i : X \to X$ be i's choice function. C is the profile of choice functions. Let S_i be the countable strategy space. Let $S = \times_{i \in N} S_i$. Let $g_i : S \to X$ map strategies into outcomes. $\langle N, S, C, g\rangle$ is a normal form game.

Definition 1 *Let i's* ***deviation set*** *be i's choice set when everyone else is playing s, so* $D_i(s) \equiv \{(s_i', s_{-i}) : s_i' \in S_i\} = S_i \times \{s_{-i}\}$.

Notice that $D_i(s)$ is invariant with respect to i's strategy, so for each $s, s' \in S$ such that $s_{-i} = s'_{-i}$, $D_i(s) = D_i(s')$. In other words, i's deviation set[3] is really only a function of s_{-i}, but writing $D_i(s)$ instead of $D_i(s_{-i})$ saves notation. Player i's choice problem is to choose from his deviation set for every $s \in S$, and so the actions of others determine i's choice problem. Note that even though i chooses from a subset of S (specifically, from $D_i(s)$), he is not choosing other players' strategies. By definition of $D_i(s)$, he can choose his strategies.

Ultimately, the definition of a game in our setting is no different than the usual definition. The main innovation is that the agents have choice functions rather than preference relations or utility functions, so we must find a way to map strategies into outcomes so that agents can evaluate them through their choice function. That is the purpose of the function g_i, which will permeate the definition of equilibria, just as the utility function permeates the definition of equilibria in the standard definition of Nash equilibria.

3.1 Equilibrium Concepts

To make traction on implementation, we must first lay out more general equilibrium concepts now that preferences need not be fully rational. This subsection lays out each concept in turn. To simplify exposition, I will consider only completely binary ($m = |S_i|$) or completely nonbinary ($m = 1$) choice. Hence at any $s \in S$, a player compares s to all other possible deviations (s_i', s_{-i}) either one–by–one, or all at once. So we have the following generalization of Nash equilibrium.

[3] de Clippel [2014] calls these deviations "opportunity sets."

Definition 2 *$s^* \in S$ is a Pairwise Nash Equilibria (PNE) if $\forall i \in N$, $g_i(s^*) \in C_i\left(g_i(\{(s_i^*, s_{-i}^*), (s_i, s_{-i}^*)\})\right)$, $\forall s_i \in S_i$.*

For comparison to the original implementation literature, the mechanism specifies an outcome for a given strategy profile. One kind of strategy, messages, is vital for research in mechanism design but not necessary for the question of Nash implementation as originally conceived. Defining mechanisms with message spaces and a corresponding revelation principle under choice functions would be a topic for future research.

Definition 3 *$s^* \in S$ is a Setwise Nash Equilibria (SNE) if $\forall i \in N$, $g_i(s^*) \in C_i\left(g_i(D_i(s^*))\right)$.*

As usual, I use the standard set notation that for any $E \in \mathcal{X}$, $g(E) = \{g(x) : x \in E\}$. Pairwise Nash Equilibria is the most direct analogue to the standard Nash equilibria in games with preference relations. Recall that the standard definition considers the equilibrium strategy against all other deviations one by one. The comparisons are made pairwise, and so is the comparison of the relevant strategies here. Because choice functions are defined over sets, each comparison set has just two elements: the outcome when the agents all play the equilibrium, and the outcome in which all other agents play the equilibrium but agent i deviates. Pairwise Nash equilibrium, however, does not take advantage of the unique properties of choice functions when all other deviations are considered together. To model this, we turn to a definition of equilibria based on sets.

To be precise, suppose the agent considers the full deviation set, namely the set of all possible deviations that he might play given that all other agents play their equilibrium strategies. Because we do not assume the rationality axioms, this setwise comparison may differ from the pairwise comparison above. Note that if (α) holds, every SNE is a PNE, and if (γ) holds, every PNE is a SNE. If both hold, we are back to preferences, and the two concepts are the same. Because choices functions place fewer restrictions on behavior than preferences, there are multiple definitions of Nash equilibria. The Pairwise Nash equilibria (PNE) is the most faithful representation of Nash equilibria under preference relations, whereas Setwise Nash Equilibria (SNE) takes advantage of the fact that agents make choices among sets of alternative strategies.

Definition 4 *$s^* \in S$ a Generalized Nash Equilibria (GNE) if it is a SNE and PNE.*

Generalized Nash Equilibria (GNE) is more restrictive than either because it must satisfy both criteria. Ideally, we would like to look within the class of GNE, but implementation is not always possible within this restrictive set. Since choice functions emphasize choices made from sets, the rest of this paper will (mostly) consider Setwise Nash Equilibria. The games under choice sets are identical to games under preferences, except for how the players evaluate those outcomes. Therefore, the game itself will not be any different, only the evaluation of the payoffs.

The following examples show how SNE and PNE differ.

Example 2. Consider a voter deciding between three candidates in the 2000 U.S. Presidential election: Bush (b), Gore (g), and Nader (n). Also, the voter may choose to abstain (a). When choosing between Bush and Gore, he cannot discern any significant difference between them, and hence is indifferent.[4] If he were to compare Bush against Nader or Gore against Nader, he clearly sees Nader's faults and chooses either Bush or Gore. But choosing among all three, he picks Nader — Bush and Gore squabble so much that they plummet in the voter's perception, relative to Nader. Throughout this example (and the following), voters strictly prefer any candidate to abstaining.

E	$C(E)$
bg	bg
gn	g
bn	b
bgn	n
ba	b
ga	g
na	n

	l	c	r
u	(a,n)	(b,a)	(n,a)
m	(a,b)	(n,n)	(a,g)
d	(n,a)	(g,a)	(a,n)

Formally, let $N = \{1,2\}$, $X = \{b,g,n,a\}$, $S_1 = \{u,m,d\}$, $S_2 = \{l,c,r\}$. Let $g : S \to X^2$ map strategies into outcomes, and let i's choice be given by $C_i(E) = C(g_i(E))$, $\forall E \subseteq X$. The figure shows the game form on the right and the choices on the left; for example, $g_2(m,l) = b$. In the notation of the figure, if Player 1 plays u and Player 2 plays l, then the joint payoff (a,n) means that Player 1 gets a and Player 2 gets n. This is the standard treatment of normal form games and game theory. The choice sets on

[4] Strictly speaking, indifference is a property of preference relations, not choice functions. I use indifference in a loose sense here.

the left side of the figure play the role of preferences. Specifically, they illustrate how each agent chooses $C(E)$ from a set E. Then (m, c) is a SNE but not a PNE. □

Example 3. Now consider the 1992 U.S. Election. The candidates are Bush (b), Clinton (c), and Perot (p). The voter is indifferent between Bush and Clinton, chooses Perot against Bush and Clinton separately, but really sees Clinton's virtues when stacked up against Perot and Bush simultaneously.

E	$C(E)$
pc	p
pb	p
bc	bc
pbc	c
ba	b
ca	c
pa	p

	l	c	r
u	(a, c)	(c, a)	(c, a)
m	(a, c)	(p, p)	(a, b)
d	(c, a)	(b, a)	(a, c)

Here, (m, c) is a PNE but not a SNE. □

These examples show that the concepts of PNE and SNE vary. Just as with preference relations, pure Nash equilibria may not exist when defined with choice sets. As such, the existence of equilibria requires mixed strategies. This existence is guaranteed with a similar proof to the existence of Nash equilibria in normal form games under preference relations.[5] Of course, the non-existence of pure strategies is not important for implementation because the whole purpose of implementation is to find a game that implements a social choice function. Nonetheless, I leave the deeper problem of defining games without preferences for future research.

4 Nash Implementation

Now turn to the implementation problem. A Social Choice Correspondence (SCC) specifies a nonempty choice set $f(C) \subseteq X$ for each $C \in \mathcal{C}$, i.e. $f : \mathcal{C} \to \mathcal{X}$. This is the natural generalization of SCC's over $\mathcal{C}$. A *mechanism* is a function $g : S \to X$ that

[5] The necessary condition is that the choice function be convex valued and Hausdorff continuous, which guarantees the existence of a setwise Nash equilibria of the game. Details on this proof can be furnished from the author on request.

specifies an outcome $g(s) \in X$ for a strategy profile $s \in S = \times_{i \in N} S_i$. Hence a mechanism and a choice profile C specifies a normal form game (g, C).

Let $GNE(g, C)$, $PNE(g, C)$, $SNE(g, C)$ be the set of GNE, PNE, and SNE of (g, C), respectively. Say that a mechanism g Nash implements a SCC f if for all $C \in \mathcal{C}$ we have $f(C) = g(GNE(g, C))$. Similarly define PNE and SNE implementation. Observe that under (α), if f is SNE implementable, then it is PNE implementable. The converse holds if we impose (γ) instead of (α). And, note that it is *not* the case that f is Nash implementable iff f is PNE and SNE implementable. Implementation in two different equilibrium concepts (double implementation) requires one mechanism that implements in the two equilibrium notions under consideration, while Nash implementation requires a single mechanism for a more restrictive equilibrium concept (since $GNE(g, C) = PNE(g, C) \bigcap SNE(g, C)$).

To explore the relation between SCCs defined over $\mathcal{C}$ and SCCs defined over $\mathcal{R}$, we need some notation. Given an SCC $f : \mathcal{C} \to \mathcal{X}$, let $h_f : \mathcal{B} \to \mathcal{X}$ by $h_f \equiv f \circ \phi^{-1}$ be the restriction to $\mathcal{B}$, i.e. $\forall R \in \mathcal{R},\ h_f(R) = f(\phi^{-1}(R)) = f(C_R)$. Similarly, given $h : \mathcal{B} \to \mathcal{X}$ let $f_h : \mathcal{C} \to \mathcal{B}$ by $f_h \equiv h \circ f$ over $\mathcal{C}'$, i.e. $\forall C \in \mathcal{C}',\ f_h(C) = h(R_C)$. Since h is defined only over $\mathcal{B}$, f_h can be extended non-uniquely outside of $\mathcal{C}'$.

We have two definitions of monotonicity for an SCC f. (M) is the natural generalization of (M*) to SCCs over $\mathcal{C}$:

(M*) $\forall R, R' \in \mathcal{R},\ a \in f(R)\ \&\ [\forall i \in N\ \forall b \in X,\ aR_i b \Rightarrow aR'_i b]\ \Rightarrow\ a \in f(R')$

(M) $\forall C, C' \in \mathcal{C},\ a \in f(C)\ \&\ [\forall i \in N\ \forall E \subseteq X,\ a \in C_i(E) \Rightarrow a \in C'_i(E)]\ \Rightarrow\ a \in f(C')$

It does not make sense to claim that if f satisfies (M), then f satisfies (M*), since (M) and (M*) apply to functions of different domains. However, we do have the following simple result showing that, roughly speaking, (M) is stronger than (M*).

Proposition 2 *If $f : \mathcal{C} \to \mathcal{X}$ satisfies (M), then h_f satisfies (M*).*

Proof: Let $R, R' \in \mathcal{R}, a \in h_f(R) \equiv f(C_R)$. Now $[\forall i \in N\ \forall b \in X,\ aR_i b \Rightarrow aR'_i b]\ \Rightarrow$ $[\forall i \in N\ \forall E \subseteq X,\ a \in M_i(E, R) \Rightarrow a \in M'_i(E, R)]$. And note that $C_{Ri}(E) = M_i(E, R)$ and $C_{R'i}(E) = M_i(E, R')$. (M) holds for f, so $a \in f(C'_R) = h_f(R')$. ■

One benefit of Maskin monotonicity is that it is easy to check. de Clippel [2014] also establishes the necessary condition for SNE implementation, but that condition is close to part of Condition μ of Moore and Repullo [1990], which is generally more difficult to check than (M).

4.1 Necessity

As in the standard case, monotonicity is important because it gives a necessary condition for Nash implementation. Indeed, the focus of the prior literature on SNE is exactly why we focus on GNE here. This is done to avoid overlap with the existing literature.

Theorem 1 *If f is GNE implementable, then f is monotonic.*

Proof: Let $C, C' \in \mathcal{C}, a \in f(C)$ and suppose $\forall i \in N\ \forall E \subseteq X,\ a \in C_i(E) \Rightarrow a \in C'_i(E)$. Since f is Nash implementable, there exists a mechanism (g, C) and an equilibrium $s \in GNE(g, C)$ such that $a = g(s)$. Since s is a GNE, $a \in C_i(D_i(s))\ \forall i$ and $a \in C_i(\{s, (s'_i, s_{-i})\})\ \forall i,\ \forall s'_i \in S_i$. By the hypothesis of monotonicity, $a \in C'_i(D_i(s))\ \forall i$ and $a \in C'_i(\{s, (s'_i, s_{-i})\})\ \forall i,\ \forall s'_i \in S_i$. Thus $s \in GNE(g, C')$, and since g Nash implements f, $a = g(s) \in f(C')$. ■

Observe that this theorem uses GNE, not SNE. Proposition 2 of de Clippel [2014] establishes a necessary condition of SNE, though it leans on a condition closer to Moore and Repullo [1990]. Instead, I rely on Condition (M), which is an analogue to Maskin monotonicity. This reinforces the point made earlier, that de Clippel [2014] and Korpela [2012] seeks to replicate the Moore and Repullo [1990] arguments for choice functions, whereas I seek to work here within the paradigm of Maskin monotonicity. Finally, observe that Hurwicz [1986] provides a similar result but only for binary relations. Therefore, the theorem above generalizes that paper for more general choice sets.

4.2 Sufficiency

We now consider sufficient conditions for Nash implementation. We have two definitions of no veto power for a SCC f:

(NVP*) $f : \mathcal{R} \to X$ satisfies (NVP*) $\forall i \in N,\ \bigcap_{j \neq i} M_j(X, R) \subseteq f(R)$

(NVP) $f : \mathcal{C} \to X$ satisfies (NVP) $\forall i \in N,\ \bigcap_{j \neq i} C_j(X) \subseteq f(C)$

Maskin [1999] showed that (M*) and (NVP*) are sufficient for Nash implementation over $\mathcal{R}$. As we may expect, without imposing *any* restrictions on the choice profiles, the generalization to implementation over $\mathcal{C}$ does not hold:

Example 4. Let $f(C) = \bigcup_{i \in N} C_i(X)$. It is easy to check that (M) and (NVP) hold. Consider

	C_1	C_2	C_3	C'_1	C'_2	C'_3	C''_1	C''_1	C''_1
ab	b	a	a	a	b	a	b	a	a
ac	c	c	c	a	c	c	a	a	a
bc	b	c	c	b	c	c	b	b	b
abc	b	b	c	a	a	a	b	b	b

and $f(C) = bc, f(C') = a, f(C'') = b$. Then $(\alpha), (\gamma)$ fail, and f is not PNE implementable. Suppose that it is. Then $\exists s \in PNE(g, C) : g(s) = c$. So $c = g(s) \in C_i(D_i(g(s)))$. By inspection, $D_1(g(s)) \neq \{b, c\}$ or $\{a, b, c\} \Rightarrow$ there does not exist a s''_1 : $g(s''_1, s_2, s_3) = b$. Hence $D_1(g(s)) = \{c\}$ or $\{a, c\}$ since $c \in D_i(c)$. Suppose $D_i(g(s)) = \{a, c\}$. Let $s' = (s'_1, s_2, s_3) : g(s') = a$. $s'_{-1} = s_{-1}$, so $D_1(g(s')) = D_1(g(s)) = \{a, c\}$. And $a \in C''_1(ac)$. Now for $j = 2, 3$ $a = g(s') \in C''_i(D_i(g(s')))$ since $D_i(g(s')) \in \{a, ab, ac\}$. So $g(s') = a \in PNE(g, C'')$. But $a \notin f(C'')$, a contradiction. So $D_1(g(s)) = \{c\}$. Then $c = g(s) \in PNE(g, C')$. But $c \notin f(C') = \{a\}$, a contradiction. So f is not PNE implementable. □

Maskin [1999] showed that monotonicity and no veto power are sufficient for Nash implementation over $\mathcal{R}$. In our setting, Maskin's theorem takes the following form:

Maskin's Theorem. *If (α), (γ) hold and $f : \mathcal{C} \to X$ satisfies (M) and (NVP), then f is Nash implementable over $\mathcal{C}$.*

Under (α) and (γ), a choice function can be represented as a preference relation, so implementation over $\mathcal{C}$ is synonymous with implementation over $\mathcal{R}$.

Yet, is it possible to relax (α) and (γ) and still obtain sufficient conditions for Nash implementation? Said differently, what is the minimum amount of rationality necessary to still get sufficiency? The next two theorems show that *some* relaxations are possible, but not major ones.

The main difficulty with extending Maskin's theorem over $\mathcal{C}$ lies in the choice-analog of the lower contour set $L_i(a) \equiv \{x : aRx\}$. When dealing with choice functions instead of preferences relations, there is no single, unambiguous way to generalize the lower contour set. Two definitions, equally plausible, come to mind.

(A) $L_i^A(a, C) \equiv \bigcap\{E : a \in C_i(E)\}$.

(B) $L_i^B(a, C) \equiv \bigcup\{E : a \in C_i(E)\}$.

Consider all sets from which a is chosen. Under (A), a is chosen from the intersection of all these sets, while under (B), a is chosen from the union of all these sets. Clearly (A) is a more restrictive definition. We need additional axioms in place of contraction and expansion.

(δ_j) If $C \in \mathcal{C}$, $i \in N$, and $a \in X$ then $a \in C_i(L_i^j(a, C))$

(ϵ) $C(E) \bigcap C(F) \subseteq C(E \bigcap F)$

(μ_j) For all $C, C' \in \mathcal{C}$, if $a \in C_i(E)$, then $C'_i(L_i^j(a, C)) \subseteq C'_i(E)$

In words, (δ_j) guarantees that a is indeed chosen from the lower contour set $L_i^j(a, C)$. While (ϵ) assumes that if a is chosen from E and is chosen from F, then a is chosen from E and F (provided their intersection is nonempty). Finally, (μ_j) is a technical condition required in the proof. Note that (α) and (B) imply (μ_B), and (μ_A) is a weak form of expansion. Using (A), we have our first sufficiency result.

Theorem 2 *Let $|N| \geq 3$. If (δ_A) and (μ_A) hold and $f : \mathcal{C} \to X$ satisfies (M) and (NVP), then f is Setwise Nash implementable over $\mathcal{C}$.*

The implementing mechanism is in the spirit of Maskin [1999]'s mechanism. Each player's strategy consists of an outcome $x \in X$, an integer, and a choice profile for the other players. If all announce the same strategy, the mechanism selects the common outcome. If one person announces a different strategy from the rest, the mechanism selects an outcome from his lower contour set $L_i(a, C)$, where a is the common strategy of everyone else. In this way, any single person has no incentive to deviate. And if more than one player announce different strategy, the mechanism selects the outcome of the player who announced the highest integer. See the appendix for details. The following example provides a sample choice function for which the hypothesis of the theorem holds, so the choice function satisfies (δ_A) and (μ_A). This, by definition, is not a strict preference relation.

Example 5. Consider the outcome space $X = \{g, x, y\}$ and choice function C for David and Simon given in Example 1. Under (A), the lower contour sets of x for David and Simon are:

$$L_D^A(x,C) = \bigcap\{E : x \in C_D(E)\} = \{gx\}\bigcap\{x\} = \{x\}$$
$$L_S^A(x,C) = \bigcap\{E : x \in C_S(E)\} = \{xy\}\bigcap\{x\} = \{x\}$$

In similar manner, computing all the lower contour sets and their choices from those sets for any action $a \in X$ gives the table

a	$L_D^A(a,C)$	$L_S^A(a,C)$	$C_D\left(L_D^A(a,C)\right)$	$C_S\left(L_S^A(a,C)\right)$
x	x	x	x	x
y	y	y	y	y
g	g	g	g	g

Indeed, we have that $L_i^A(a,C) = \{a\} = C_i\left(L_i^A(a,C)\right)$ for all $a \in X$ and $i \in N$. This occurs because the intersection of the lower contour sets of any element always contains the singleton set of that element itself. Another choice function $C' \in \mathcal{C}$ is defined by $C'(E) = E$, the "no-choice" option. Computing the same table above also shows that $L_i^A(a,C') = \{a\} = C_i'\left(L_i^A(a,C')\right)$.

Let $C \in \mathcal{C}$, $i \in N$, $a \in X$. Then $a \in \{a\} = C_i\left(L_i^A(a,C)\right)$ trivially, so (δ_A) holds. Also, for all C, $C'' \in \mathcal{C}$, if $a \in C_i(E)$, then $C_i'\left(L_i^A(a,C)\right) = \{a\} \subseteq C_i'(E) = E$, so (μ_A) holds, also. Example 1 shows that C is non-binary. Hence we have a choice profile that contains a non-binary choice function and satisfies (A), (δ_A), and (μ_A). $\square$

Using (B), we have our second sufficiency result.

Theorem 3 *Let $|N| \geq 3$. If (α) and (δ_B) hold and $f : \mathcal{C} \to X$ satisfies (M) and (NVP), then f is Setwise Nash implementable over $\mathcal{C}$.*

The mechanism here is identical to the one from the previous theorem. Observe that (δ_B) implies that $\forall a \in X, a \in C_i(\bigcup\{E : a \in C_i(E)\})$. This says that if $a \in C_i(E)$ and $a \in C_i(F)$, then $a \in C_i(E \bigcup F)$. But this is exactly (γ)! In other words, Theorem 3 is equivalent to Maskin's Theorem in our setting. Maskin [1999] proved his theorem under the assumption of preference relations. Here, we show that the intuitions behind that proof are more general and can drive the results of the last two theorems. In particular, Theorem 3 uses the framework of choice functions, rather than preference relations, and therefore, shows that the logic of Maskin's theorem is not limited to preference relations. Indeed, assuming conditions (δ_{A}) and (μ_{A}) instead of (α) and (γ) allow us to use choice-based arguments rather than preference-based arguments. [6]

Observe that (δ_{A}) and (μ_a) avoid reliance on conditions (α) and (γ) from Maskin's theorem. Therefore, there is a tradeoff between these two sets of conditions. In contrast, from Theorem 3 above, we see that (α) and (δ_{B}) are equivalent to (γ). In this sense, the looser definition of the lower contour sets from (B) eliminates the reliance on γ. Ultimately, the definition of the lower contour sets is vital to determining whether we need to assume γ in order to guarantee our results.

5 Nash Equilibria in Choices

The Nash concepts of the previous section (SNE and PNE) are the natural generalizations in games with choice functions rather than preferences. However, SNE and PNE do not make full use of the fact that choice functions map sets into sets. Here I propose a *set-based* equilibrium that exploits the special structure of choice functions that preference relations or utility functions cannot provide. Moreover, moving to a set-based concept permits some consideration of equilibrium selection.

Let i's deviation set for a set E be all possible strategy profiles that i may choose if the other players choose any other equilibrium strategy in E. In other words, $D_i(E) \equiv \{s \in S : s = (s_i, s_{-i}) \text{ for some } s_i \in S_i \text{ and some } s_{-i} \in E_{-i}\} = \bigcup_{s \in E} D_i(s) = S_i \times E_{-i}$. The interpretation is that if E is a set of equilibria, then i's equilibrium choice should allow for all other players to play *any* strategy in E. Hence E qualifies as an equilibrium if everyone chooses it when considering $D_i(E)$:

Definition 5 *$E \subseteq S$ is a set of Nash Equilibria in Choices (NEC) if $\forall i \in N,\ E \subseteq C_i(g_i(D_i(E)))$.*

What are some reasonable restrictions on the equilibrium set E? One restriction is that every $s \in E$ be a SNE. Note that this will hold if (α) does. Moreover, note that if $|E| = 1$, then E is a NEC iff E (={s}) is a SNE. Another reasonable property is Pareto Undominance: it is never the case that $\forall i \in N,\ C_i(E) \subset E$ if E is a NEC, where the set inclusion is strict.

[6] Because the conditions for **Theorem 3** are equivalent to full rationality, I do not provide an explicit proof of **Theorem 3**. Such a proof is available on request and is equivalent to an alternative proof of Maskins theorem.

Example 6. *Pareto Dominant Equilibria.* Let $X = \{a, b, 0\}$, $a = C(a, b)$, and suppose (α) and (γ) hold, so $aR_C bR_C 0$. Suppose that

E	$C(E)$
$a0$	a
$b0$	b
ab	a

	l	r
u	(a, a)	$(0, 0)$
d	$(0, 0)$	(b, b)

(u, l) and (d, r) are both GNE, but $E = \{(u, l), (d, r)\}$ is not a NEC, since $(u, l) = C_i(g_i(D_i(E))) = C_i(S)$ for $i = 1, 2$. This confirms intuition that while (u, l) and (d, r) may be equilibria individually, they cannot both be equilibria as a set since the former Pareto dominates the latter. Hence NEC can rule out Pareto dominated equilibria. □

Example 7. *Battle of the Sexes.* Consider the same setup as the previous example, but let the game be

	l	r
u	(a, b)	$(0, 0)$
d	$(0, 0)$	(b, a)

(u, l) and (d, r) are both GNE, but $E = \{(u, l), (d, r)\}$ is not a NEC, since $(u, l) = C_1(g_1(D_1(E)))$ and $(d, r) = C_2(g_2(D_2(E)))$.□

There are other possible equilibria depending on how the agents compare different subsets of deviations. For tractability, I only examine pairwise and the full set of deviations. But in principle, the agents could consider any possible subset of deviations, and this may or may not align with the other two definitions above. Of course, if a strategy profile satisfies both the pairwise and setwise definition, then this is a stronger equilibrium concept as it asks more of the strategy profile.

5.1 Nash Implementation in Choices

With a set-based equilibrium concept, we are now open to the possibility of multiple equilibria. Even in a fully rational model, multiple equilibria can affect the implementation problem. In general, implementation is easier with multiple equilibria because we have more possible ways to implement an outcome. At the same time, implementation is somewhat less satisfying if the space of equilibria is quite large because the multiplicity loosens the constraints on the mechanism. The development of this section largely follows Moore and Repullo (1990), which establishes axioms under a rational agent for Nash implementation. These conditions are indeed necessary and sufficient.

Let $NEC(g, C)$ be the set of Nash Equilibria in Choices. Note that this is a set of sets. With this definition of Nash equilibrium, we are now able to define what it means for a mechanism to implement a social choice correspondence. In words, a mechanism implements a social choice correspondence if, for every choice function and every outcome in that choice function, there exists a Nash equilibrium that arrives at that outcome. This is a natural generalization in this set-valued world.

Definition 6 *A mechanism $g : S \to X$ NEC implements a SCC $f : \mathcal{C} \to X$ if*

(i) $\forall C \in \mathcal{C},\ F \subseteq f(C) \ \Rightarrow$ *there exists* $E \in NEC(g, C) : g(E) = F$.

(ii) $\forall C \in \mathcal{C},\ E \in NEC(g, C) \ \Rightarrow\ g(E) \subseteq f(C)$.

The usual requirement for implementation would be $\forall C \in \mathcal{C},\ f(C) = g(NEC(g, C))$. The definition above breaks this condition into two pieces, and makes explicit use of our *set-based* concept.

The Nash equilibrium in choices helps address equilibrium selection. Even with a restrictive definition like GNE (which itself is stricter than either PNE or SNE alone), we still may have multiple equilibria. NEC effectively requires that a set of equilibrium strategies must be a best response to all other deviations, for each strategy in the equilibrium set. For example, if s_1 is a best response against $D_1(s_1)$, and s_2 is a best response against $D_2(s_2)$, then $\{s_1, s_2\}$ is a Nash equilibrium in choices if s_2 is a best response against $D_1(s_1)$ and $D_2(s_2)$. This stronger definition of equilibria allows us to generate necessary and sufficient conditions to characterize Nash implementation, to which we now turn.

Definition 7 *$f : \mathcal{C} \to X$ satisfies **(Z)** if there is a set $B \subseteq X$ and $\forall i \in N$, $\forall C \in \mathcal{C}$, and $\forall F \subseteq f(C)$ there is a set $D_i(F, C) \subseteq B$ such that $F \subseteq C_i(g_i(D_i(F, C)))$ and $\forall C' \in \mathcal{C}$*

(Z1) $F \subseteq \bigcap_{i \in N} C'_i(g_i(D_i(F, C))) \ \Rightarrow\ F \subseteq f(C')$

(Z2) $d \subseteq C'_i(g_i(D_i(F, C)))$ *and* $d \subseteq \bigcap_{j \neq i} C'_i(B)$ *for some* $i \ \Rightarrow\ d \subseteq f(C')$

(Z3) $d \subseteq C'_i(B)\ \forall i \ \Rightarrow\ d \subseteq f(C')$

The following theorem generalizes Moore and Repullo [1990] to our new setting.

Theorem 4 *Let $|N| \geq 3$ and (α) hold. f is NEC implementable iff (Z) holds.*

The proof of the theorem follows the same intuition as the proof of Maskin's theorem. The mechanism is more complex to accomodate our set-based definition of implementation. Recall that in Maskin's theorem, any Nash equilibria s falls into three cases: (i) either everyone's action is the same ($s_i = s_j \ \forall i \neq j$), (ii) one person deviates ($s_i = s \ \forall i \neq j$, for some j), or (iii) many people deviate. Since now we are considering a set of equilibria E, at first it seems that there are many possibilities. For example, suppose E contains two equilibria of the second (ii) type; there is no guarantee that the same person deviates in each of the equilibria. However, the special structure of the mechanism permits us to classify E into three cases: (a) either every equilibria in E is of type (i) only; or (b) E contains equilibria of type (ii), possibly in addition to equilibria of type (i); or (c) E contains equilibria of type (iii), possibly in addition to equilibria of type (i), (ii), or both. These cases are exhaustive — see the proof for details.

6 Conclusion

The main message of this paper is that it is worthwhile to supplement implementation theory with a behavioral component. Implementation and mechanism design theory as they now exist rest on the usual standard assumptions of rationality that characterize mainstream economics. As psychologists and behavioral economists have shown, such assumptions lack experimental support and descriptive content.

The implementation results of this paper generalize the well-established results from a preference-based environment. Our intuition of implementation derives from thinking from the preference perspective. The next step is to think entirely from a choice perspective, perhaps reformulating the problem in fundamental ways. I have tried to do this a bit with NEC implementation, but clearly there is still more to do. While choices allow for richer behavior, they also provide less analytical discipline than preferences. Agents whose choice functions fail (α) will react differently to one another than those whose choice functions fail (γ). And if we add uncertainty, the choices of agents may be fully rational in some states and not as rational in others. What happens when they interact? These are all open questions.

7 Appendix

Proof of Proposition 1: Let $\psi : \mathcal{C} \to \mathcal{B}$ be defined by $\psi(C) = R_C$. Let $\phi : \mathcal{C}' \to \mathcal{B}$ be defined by $\phi = \psi|_{\mathcal{C}'}$. So for all $C \in \mathcal{C}', \phi(C) = R_C$. It is enough to show that ϕ is well-defined, one-to-one, and onto.

If $C_1 = C_2$, then $xR_{C_1}y \Leftrightarrow x \in C_1(xy) \Leftrightarrow x \in C_2(xy) \Leftrightarrow xR_{C_2}y$, so $\phi(C_1) = R_{C_1} = R_{C_2} = \phi(C_2)$ and hence ϕ is well-defined. Take $C_1, C_2 \in \mathcal{C}'$ such that $\phi(C_1) = \phi(C_2)$, i.e. $R_{C_1} = R_{C_2}$. Then $\forall E \in \mathcal{X},\ C_1(E) = M(E, R_{C_1}) = M(E, R_{C_2}) = C_2(E)$. Hence ϕ is one-to-one. Finally, let $R \in \mathcal{B}$. Take $C_R \in \mathcal{C}'$ by $C_R(E) = M(E, R)\ \forall E \in \mathcal{X}$. Then $\phi(C_R) = R_{C_R} = R$ since $xR_{C_R}y \Leftrightarrow x \in C_R(xy) = M(xy, R) \Leftrightarrow xRy$. Hence ϕ is onto.

Thus ϕ is a bijection, and so it has an inverse $\phi^{-1} : \mathcal{B} \to \mathcal{C}'$. By the same arguments above, it can be checked that if we write $\phi^{-1}(R) = C_R = M(E, R)$, then $\phi(\phi^{-1}(R)) = R\ \forall R \in \mathcal{B}$ and $\phi^{-1}(\phi(C)) = C\ \forall C \in \mathcal{C}'$.

To show that $\mathcal{C}$ is not equivalent to $\mathcal{R}$, it is enough to show that ψ is not one-to-one. Indeed, let $X = \{a, b, c\}$ and consider $C_1, C_2 \in \mathcal{C}$:

E	$C_1(E)$	$C_2(E)$	R
ab	a	a	a
ac	a	a	b
bc	b	b	c
X	a	b	

Here, $C_1 \neq C_2$ but $\psi(C_1) = R_{C_1} = R = R_{C_2} = \psi(C_2)$. ■

Proof of Theorem 2: Since we operate under (A), write μ, δ for μ_A, δ_A to save notation. Recall that under (α), if s is a SNE then it is a PNE. Hence it suffices to show that f is SNE implementable over $\mathcal{C}$. For each $i \in N$ let

$$S_i = \{(C_i, a_i, n_i) \in \mathcal{C} \times X \times \mathbb{N}\ |a_i \in f(C_i)\}$$

where $\mathbb{N}$ is the set of nonnegative integers. Construct the mechanism $g : S \to X$ as follows. Define $g(s)$ according to three mutually exclusive cases:

(1) If there exists $(C, a, n) \in \mathcal{C} \times X \times \mathbb{N}$ such that $s_i = (C, a, n)\ \forall i \in N$, then $g(s) = a$.

(2) If there exists $(C, a, n) \in \mathcal{C} \times X \times \mathbb{N}$ such that $s_j = (C, a, n)\ \forall j \neq i$ and $s_i = (C'_i, a'_i, n_i) \neq (C, a, n)$, then let

$$g(s) = \begin{cases} a'_i & \text{if } a'_i \in L_i(a, C) \\ a & \text{otherwise} \end{cases}$$

(3) If neither of the above apply, let $g(s) = b_l$, where $l = \min\{i | m^i = \max m^j\}$. Whoever shouts loudest wins.

Now we need to show: $\forall C \in \mathcal{C},\ f(C) = g(SNE(g, C))$.

($\subseteq$) Let $C \in \mathcal{C},\ a \in f(C)$. Let $s_i = (C, a, 0)\ \forall i \in N$. Then $D_i(g(s)) = L_i(a, C)$, and $g(s) = a$. Hence $(\delta) \ \Rightarrow\ \forall i\ a = g(s) \in C_i(D_i(g(s))$, so $s \in SNE(g, C)$.

($\supseteq$) Let $C' \in \mathcal{C},\ s \in SNE(g, C')$. There are three cases.

(1) holds, so $s_i = (C, a, n)\ \forall i$ and $a \in f(C)$. Now $D_i(g(s)) = L_i(a, C)$, so $s \in SNE(g, C') \ \Rightarrow\ a = g(s) \in C'_i(L_i(a, C))$. Take any set E such that $a \in C_i(E)$. Then $(\mu) \ \Rightarrow\ a = g(s) \in C'_i(E)$. And by (M), $a = g(s) \in f(C')$.

(2) holds, so $s_j = (C, a, n)\ \forall j \neq i$ and $a \in f(C)$. Now $D_j(g(s)) = X$, so $s \in SNE(g, C') \ \Rightarrow\ g(s) \in C'_j(X)$, and thus by (NVP), $g(s) \in f(C')$.

(3) holds, so $D_i(g(s)) = X\ \forall i$. By the same argument from Case (2), (NVP) implies $g(s) \in f(C')$.

■

Proof of Theorem 3: Write δ for δ_B. The mechanism is identical to the one in Theorem 2. The proof is almost identical, with one small change. To show $\forall C \in \mathcal{C},\ f(C) \supseteq g(SNE(g, C))$, Case (1) becomes

(1) holds, so $s_i = (C, a, n)\ \forall i$ and $a \in f(C)$. Now $D_i(g(s)) = L_i(a, C)$, so $s \in SNE(g, C') \ \Rightarrow\ a = g(s) \in C'_i(L_i(a, C))$. Take any set E such that $a \in C_i(E)$. Then $E \subseteq L_i(a, C)$, so $(\alpha) \ \Rightarrow\ a = g(s) \in C'_i(E)$. And by (M), $a = g(s) \in f(C')$.

The rest of the argument is the same. ■

Proof of Theorem 4: Recall that $D_i(s) = \{s|s_i \in S_i\}$, so $D_i(E) = \bigcup_{s\in E} D_i(s)$. Of course, $E \subseteq D_i(E)\ \forall E \subseteq X$. Note that E and E' that agree on S_{-i} have the same deviation sets, i.e. $E_{-i} = E'_{-i} \Rightarrow D_i(E) = D_i(E')$. In general, $E_{-i} \subseteq E'_{-i} \Rightarrow D_i(E) \subseteq D_i(E')$. Moreover, note that $(D_i(E))_{-i} = E_{_i}$, so $D_i(D_i(E)) = D_i(E)$. Finally, $d \subseteq D_i(E) \Rightarrow d_{-i} \subseteq E_{-i} \Rightarrow D_i(d) \subseteq D_i(E)$.

(Necessity). Let $B = g(S)$. Let $C \in \mathcal{C},\ F \subseteq f(C)$. Let g NEC implement f. Then there exists $E \in NEC(g, C)$ such that $g(E) = F$. Note that E depends on F and C (so $E = E(F, C)$), but I will suppress those arguments for clarity. Let $D_i(F, C) = D_i(g(E))$.

Take $C' \in \mathcal{C}$, and suppose $\forall i \in N,\ F \subseteq C'_i(D_i(F, C))$. Then $E \in NEC(g, C')$, and so $g(E) = F \subseteq f(C')$ since g NEC implements f. So (Z1) holds.

Let $i \in N,\ d \subseteq D_i(F, C)$. Pick $C' \in \mathcal{C} : d \subseteq C'_i(D_i(F, C))$ and $d \subseteq \bigcap_{j\neq i} C'_j(B)$. Now $d \subseteq D_i(F, C) \Rightarrow D_i(d) \subseteq D_i(F, C)$. So $(\alpha) \Rightarrow d \subseteq C'_i(D_i(d))$. And $\forall j \neq i\ d \subseteq D_j(d) \subseteq g(S) = B$. So $(\alpha) \Rightarrow d \subseteq C'_j(D_i(d))$. Now $d \subseteq B$, so there exists $S' \subseteq S : g(S') = d$. Since g NEC implements f, $S' \in NEC(g, C') \Rightarrow d = g(S') \subseteq f(C')$. So (Z2) holds.

Take $C' \in \mathcal{C}$. Let $d \subseteq C'_i(B)\ \forall i \in N$. Now $d \subseteq D_i(d) \subseteq g(S) = B$, so $(\alpha) \Rightarrow d \subseteq C'_i(D_i(d))$. Since g NEC implements f, there exists $S' \subseteq S : S' \in NEC(g, C')$ and $d = g(S') \subseteq f(C')$. So (Z3) holds.

(Sufficiency). Construct the mechanism $g : S \to X$ as follows. For each $i \in N$ let

$$S_i = \{(C_i, a_i, b_i, F_i, n_i) \in \mathcal{C} \times X \times B \times \mathcal{X} \times \mathbb{N} | a_i \in F_i \subseteq f(C_i)\}$$

where $\mathbb{N}$ is the set of nonnegative integers. Classify $s \in S$ into three mutually exclusive types, and define $g(s)$ according to the type of s:

(T1) If there exists $(C, a, b, F, n) \in \mathcal{C}\times X\times B\times\mathcal{X}\times\mathbb{N}$ such that $s_i = (C, a, b, F, n)\ \forall i \in N$, then s is (T1). Let $g(s) = a$.

(T2) If there exists $(C, a, b, F, n) \in \mathcal{C}\times X\times B\times\mathcal{X}\times\mathbb{N}$ such that $s_j = (C, a, b, F, n)\ \forall j \neq i$ and $s_i \neq (C, a, b, F, n)$, then s is (T2). Let

$$g(s) = \begin{cases} b_i & \text{if } b_i \in D_i(F, C) \\ a & \text{otherwise} \end{cases}$$

(T3) If neither of the above apply, s is (T3). Let $g(s) = b_l$, where $l = \min\{i|m^i = \max m^j\}$. Whoever shouts loudest wins.

(i) holds. Let $C \in \mathcal{C},\ F \subseteq f(C)$. Let $\forall a \in F s_i(a) = (C, a, a, F, 0)\ \forall i \in N$. Now $E = \bigcup_{a\in F} s(a) = \{s(a)|a \in F\}$ and $g(s(a)) = a$, so $g(E) = F$. And $D_i(g(s(a)) = D_i(F, C)\ \forall a \in F \ \Rightarrow\ D_i(g(E)) = \bigcup_{a\in F} D_i(g(s(a))) = D_i(F, C)$. Moreover, $g(E) = F \subseteq C_i(D_i(F, C)) = C_i(D_i(g(E))) \ \Rightarrow\ E \in NEC(g, C)$. So (i) holds.

(ii) holds. Let $C' \in \mathcal{C}$ and $E \in NEC(g, C')$. Note that

(#1) $g(E) \subseteq C'_i(D_i(g(E)))$.

(#2) $D_i(g(E)) = \bigcup_{s\in E} D_i(g(s))$

s is	$D_i(g(s))$	$D_j(g(s))\ \forall j \neq i$
(T1)	$D_i(F, C)$	$D_j(F, C)$
(T2)	$D_i(F, C)$	B
(T3)	B	B

Case 1. Every $s \in E$ is (T1). Then $g(E) = F$ and $D_i(g(s)) = D_i(F, C)\ \forall s \in E,\ \ \forall i$, so $D_i(g(E)) = D_i(F, C)$. Thus (#1) and (#2) $\Rightarrow$ $F = g(E) \subseteq C'_i(D_i(F, C))$, and by (Z1), $g(E) \subseteq f(C')$.

Case 2. Let $E = G \bigcup H$ where $s \in G \ \Rightarrow\ s$ is (T1) and $s \in H \ \Rightarrow\ s$ is (T2). For each $s \in H$, one guy deviates. Denote him by $k(s)$.

Case 2a. $k(s) = k\ \forall s \in H$,i.e. the same guy deviates throughout H. Then $\forall s \in H,\ D_k(g(s)) = D_k(F, C)$ and $D_j(g(s)) = B\ \forall j \neq k$. And $\forall s \in G,\ \ \forall i \in N,\ D_i(g(s)) = D_i(F, C)$. By (#2), $D_k(g(E)) = D_k(g(G)) \bigcup D_k(g(H)) =$ $(\bigcup_{s\in G} D_k(g(s))) \bigcup (\bigcup_{s\in H} D_k(g(s))) = D_k(F, C)$. Similarly, $\forall j \neq k,\ D_j(g(E)) =$ $D_j(F, C) \bigcup B = Bi$. Now (#1) and (Z2) $\Rightarrow$ $g(E) \subseteq f(C')$.

Case 2b. $k(s) \neq k\ \forall s \in H$, i.e. different guys deviate for different strategies in H. Take any $s, t \in E : k(s) \neq k(t)$. Then $D_{k(s)}(g(s)) = D_{k(s)}(F, C)$ and $D_j(g(s)) = B\ \forall j \neq k(s)$, in particular for $j = k(t)$. Similarly, $D_{k(t)}(g(t)) = D_{k(t)}(F, C)$ and $D_j(g(t)) = B\ \forall j \neq k(t)$. Let $Q = \{g(s), g(t)\}$. Then

$$D_{k(s)}(Q) = D_{k(s)}(F, C) \bigcap B = B$$
$$D_{k(t)}(Q) = D_{k(t)}(F, C) \bigcap B = B$$
$$D_j(Q) = B \bigcap B = B\ \forall j \neq k(s), k(t)$$

Repeat this process by induction, throughout H. Then $D_i(g(H)) = B\ \forall i \in N$. So $D_i(g(E)) = D_i(g(H)) \bigcap D_i(g(G)) = B \bigcap D_i(F, C) = B$. So (#1) and (Z3) $\Rightarrow$ $g(E) \subseteq f(C')$.

Case 3. There exists $s \in E : s$ is (T3). Let s be one such (T3) strategy. Then $D_i(g(s)) = B$, hence $D_i(g(E)) = B$ since $D_i(g(t)) = D_i(F, C)$ or $B \; \forall t \in E \backslash \{s\}$. Thus, (#1) and (Z3) $\Rightarrow g(E) \subseteq f(C')$.

These cases are exhaustive. Hence (ii) holds. ■

References

Austen-Smith, D.; and J. S. Banks. *Positive Political Theory I: Collective Preference*. Michigan Studies in Political Analysis. University of Michigan Press, 1999.

de Clippel, G. "Behavioral implementation." *American Economic Review* 104 (2014): 2975–3002.

Eliaz, K. "Fault tolerant implementation." *Review of Economic Studies* 69 (2002): 589–610.

Glazer, J.; and A. Rubinstein. "Motives and implementation: On the design of mechanisms to elicit opinions." *Journal of Economic Theory* 79 (1998): 157–173.

Hurwicz, L. *On the Implementation of Social Choice Rules in Irrational Societies*. Essays in Honor of Kenneth J. Arrow: Volume 1, Social Choice and Public Decision Making. Cambridge University Press, 1986.

Kahneman, D.; and A. Tversky (editors). *Choices, Values, and Frames*. Cambridge University Press, 2000.

Korpela, V. "Implementation without rationality assumptions." *Theory and Decision* 72 (2012): 189–203.

Kreps, D. *Notes on the Theory of Choice*. Westview Press, 1988.

Maskin, E. "Nash Equilibrium and Welfare Optimality." *Review of Economic Studies* 66 (1999): 23–38.

Moore, J.; and R. Repullo. "Nash Implementation: A Full Characterization." *Econometrica* 58 (1990): 1083–1099.

Saran, R. "Menu-depedent preferences and revelation principle." *Journal of Economic Theory* 146 (2011): 1712–1720.

Tversky, A.; and I. Simonson. "Context-dependent preferences." *Management Science* 39 (1993): 1179–1189.