A Simple Sufficient Condition for Strong Implementation

VILLE KORPELA *†
Public Choice Research Centre (PCRC),
University of Turku
Finland

October 23, 2012

Abstract

In an important step forward Maskin (1977) showed that two properties — Monotonicity and No Veto Power — form a sufficient condition for Nash implementation when conjoinet. In contrast to the vast literature that followed, this characterization has two big advantages: First, it is often easy to verify, and second, it has an elegant and simple interpretation. However, there does not exist a similar condition for social choice correspondences that are implementable in strong equilibrium. All existing characterizations are either hard to verify or apply only in comprehensive preference domains. In this paper we improve the situation by give such a condition called the Axiom of Sufficient Reason and show that this is a practical tool in some important applications.

JEL Classification: C72; D78
Keywords: Holocaust alternative; Implementation; Matching rule; Sufficient condition; Strong equilibrium

* I am grateful to all participants of the weekly seminar of Public Choice Research Center at University of Turku for helpful comments, in particular to Hannu Salonen and Marko Ahteensuu. I thank the Academy of Finland for financial support.
† E-mail: ville.korpela@utu.fi, tel. +358 2 333 6428, address: Turku School of Economics, Rehtorinpellonkatu 3, 20500 Turku, Finland.
1 Introduction

The aim of implementation theory is to characterize those social goals that can be realized in a decentralized way when information is dispersed in the economy and individuals cannot be trusted to act sincerely. In a seminal contribution Maskin (1977)\(^1\) showed that a property called Monotonicity is a necessary condition for Nash implementation, and furthermore, when conjoint with a weak property called No Veto Power it forms a sufficient condition. Roughly speaking, Monotonicity says that when an alternative is selected under one preference profile and its ranking has not dropped under another, then it should be selected also under the second profile. No Veto Power, on the other hand, says that when an alternative is top ranked for all individuals except possibly one, it should be selected. However, it was not until the work of Moore and Repullo (1990) and Dutta and Sen (1991a) that the cap between necessary and sufficient conditions was finally closed.

The logical next step after solving for the Nash implementation problem, which can be interpreted as the competitive case, was to study a cooperative equilibrium concept.\(^2\) An obvious candidate was the strong equilibrium. This is a Nash equilibrium that is not resistant only to individual deviations, but also to deviations by coalitions, all coalitions being possible a priori.\(^3\) It was discovered quite early in the literature that Monotonicity is a necessary condition also for implementation in strong equilibrium, while this time it does not form a sufficient condition when conjoint with No Veto Power (Maskin, 1979). However, a full characterization (a necessary and sufficient condition) exists now also in the cooperative case. Dutta and Sen (1991a) and Suh (1996) have given a full characterization that is based on the idea laid down by Moore and Repullo (1990) and Dutta and Sen (1991a) in the case of Nash implementation, while Fristrup and Keiding (2001) have given a full characterization that is based on the idea of Effectivity Functions.

---

\(^1\)Published later as Maskin (1999).

\(^2\)Historically this was done already before the Nash implementation problem was fully solved.

\(^3\)See Suh (1996) for a result where the set of possible coalitions can be arbitrary.
developed by Moulin and Peleg (1982) and Holzman (1987) for example. Despite of the considerable progress, there is a notable shortcoming in the strong implementation literature. Namely, there does not exist a condition that is comparable to the sufficient condition of Maskin (1977) in the case of Nash implementation. This is a substantial handicap. The original sufficient condition of Maskin has two big advantages in comparison to the vast literature that followed: First, it is often easy to verify, and second, it has a nice and intuitive interpretation. All the characterizations that are based on the idea of Moore and Repullo (1990), while illuminating, are almost as hard to verify as finding the implementing game form itself. All the characterizations that are based on the Effectivity Functions, on the other hand, apply only when the preference domain is comprehensive. Moreover, these characterizations impose conditions on the Effectivity Function of the goal rather than directly on the goal, which tends to somewhat cloud the interpretation.

In this paper we improve current situation by giving a simple and intuitive sufficient condition for strong implementation. We dub the central property of this condition as the Axiom of Sufficient Reason. For any preference relation of individual $i$ we say that the pair $(b, i)$ is a reason to select alternative $a$ if individual $i$ prefers alternative $a$ to alternative $b$. A social goal satisfies the Axiom of Sufficient Reason if, whenever an alternative $a$ is selected under one preference profile and every reason to select it is also a reason to select $c$ (possibly different from $a$) under another profile, then alternative $c$ should be selected under the second profile. Although this condition is considerably stronger than Monotonicity, as a sufficient condition should be by the results of Maskin (1979), it seems to be applicable in many interesting cases.

The rest of this paper is organized as follows: In section 2 we run through all the basic definitions needed in every implementation paper. A reader familiar with the field can easily skip this section. Although, one should take a look at our equilibrium concept, which is a bit different than usually.
In section 3 we show that when there are only two individuals, strong implementation does not much differ from Nash implementation beyond the additional requirement of strong Pareto optimality. In section 4 we move to the more difficult case of at least three individuals. Here we finally introduce the Axiom of Sufficient Reason and prove that together with a few technical conditions it is sufficient for strong implementation. Section 5 is important. We show that there are a lot of important cases where our condition can be applied. In particular, we show that our condition is useful with matching rules. This is potentially important, since the original result of Maskin (1977) in Nash implementation does not seem to work with matching rules. Usually, the result of Danilov (1991) or Moore and Repullo (1990) has to be used instead (e.g. Kara and Sönmetz, 1996). Finally, section 6 concludes the paper with a brief discussion of the results.

2 Notation and Preliminaries

We denote the set of individuals (the society) by $N = \{1, \ldots, n\}$ and the set of social alternatives by $A$. Let $\mathcal{R}_A$ be the set of all complete and transitive preference relations over $A$. This is the unrestricted domain or the full domain. We denote a typical element of $\mathcal{R}_A$ by $R$ and the preference relation of individual $i$ in this profile by $R_i$, with a strict part $P_i$ and indifference part $I_i$. As standard, $R_{-i}$ will denote an $(n-1)$-dimensional profile that specifies the preference relation of all individuals except $i$. Let $\mathcal{R} \subseteq \bigotimes \mathcal{R}_A$.

A Social Choice Correspondence (SCC) is any non-empty correspondence $f : \mathcal{R} \to A$ that picks the set of socially acceptable alternatives as a function of the preference profile. By a coalition we mean any non-empty subset of individuals. The coalition that contains all the individuals is called the grand coalition.

A game form (or mechanism) $G$, is an $(n+1)$-tuple $(S_1, \ldots, S_n; g)$ where $S_i$ is the strategy set of individual $i$ and $g : \bigotimes S_i \to A$ is the outcome function. Let $S = \bigotimes S_i$ and denote a typical element of this set by $s$. For a given preference profile $R$, this game form will define a game $\Gamma = (G, R)$ in a
normal form. For any coalition $C \subseteq N$, let $s_C = (s_i)_{i \in C} \in S_C = \bigotimes_{i \in C} S_i$ and $s_{-C} = (s_i)_{i \in N \setminus C} \in S_{N \setminus C} = \bigotimes_{i \in N \setminus C} S_i$.

Since we allow for coalition formation, a natural choice as the solution concept is the strong equilibrium. A strategy profile $s \in S$ is a strong equilibrium of the game $\Gamma = \langle G, R \rangle$, if for every coalition $C \subseteq N$ and all joint strategies of this coalition $s'_C \in S_C$, there exists an individual $i \in C$, such that either $g(s)_i P g(s'_C, s_{N \setminus C})$ or $g(s)_i I g(s'_C, s_{N \setminus C})$ for all $i \in C$. 4 We get the definition of Nash equilibrium as a special case if only coalitions of size one are allowed. Notice that according to this definition a coalition deviates if the position of no member is made worse and the position of at least one member is strictly improved. Let $SE(G, R) \subseteq S$ be the set of all strong equilibria of the game $\Gamma$. We say that the game form $G$ implements $f$ in strong equilibrium if $g(SE(G, R)) = f(R)$ for all $R \in \mathcal{R}$.

If there exists a game form $G$ that implements $f$ in strong equilibrium, then $f$ is called strongly implementable or implementable in SE.

A few properties of SCCs’ are needed. Let $L_i(a, R)$ denote the lower contour set of individual $i$ at $a \in A$ when the preference profile is $R$, that is $L_i(a, R) = \{ b \in A \mid a R_i b \}$. 5 A SCC $f$ is called Monotonic, if $a \in f(R)$ and $L_i(a, R) \subseteq L_i(a, R')$ for all $i \in N$, implies that $a \in f(R')$. A SCC $f$ satisfies No Veto Power, if for all $a \in A$ and all $R \in \mathcal{R}$, $\# \{ i \mid L_i(a, R) = A \} = n - 1$ implies that $a \in f(R)$. No Veto Power is stronger than Unanimity, which only requires that $\# \{ i \mid L_i(a, R) = A \} = n$ implies $a \in f(R)$.

An alternative $a \in A$ is called strongly Pareto optimal for the preference profile $R$ if there does not exist another alternative $b \in A$, such that $b R_i a$ for all $i \in N$ and $b P_j a$ for at least one $j \in N$. We denote the set of all those alternatives in $B \subseteq A$ that are strongly Pareto optimal for the prefer-

\footnote{This is not the solution concept that is usually used in the literature on coalition-proof implementation. However, it gives us sharp results.}

\footnote{Strictly speaking $L_i$ only depends on $R_i$. Still, it is notationally more convenient to express it as depending on $R$.}
ence profile \( R \) by \( \text{sPO}(B, R) \). A SCC \( f \) is called strongly Pareto optimal if \( f(R) \subseteq \text{sPO}(A, R) \) for all \( R \in \mathcal{R} \). In a similar way, an alternative \( a \in A \) is called \textit{weakly Pareto optimal} for the preference profile \( R \) if there does not exist another alternative \( b \in A \), such that \( b \succ_{i} a \) for all \( i \in N \). We denote the set of all those alternatives in \( B \subseteq A \) that are weakly Pareto optimal for the preference profile \( R \) by \( \text{wPO}(B, R) \). A SCC \( f \) is called weakly Pareto optimal if \( f(R) \subseteq \text{wPO}(A, R) \) for all \( R \in \mathcal{R} \). Since \( \text{sPO}(A, R) \subseteq \text{wPO}(A, R) \), every strongly Pareto optimal SCC is also weakly Pareto optimal, but not vice versa.

3 The Relatively Simple Two Player Case

In Nash implementation theory the case of two individuals \( (n = 2) \) is much harder than the case of three or more individuals \( (n \geq 3) \) (e.g. Moore and Repullo, 1990, Dutta and Sen, 1991b). Roughly speaking, this is cause an additional intersection condition must hold for lower contour sets when there are only two individuals. In contrast, the case of three or more individuals turns out to be much harder than the two individual case if we want to implement in strong equilibrium. In fact, as confirmed by the following theorem, strong implementation with two individuals does not much differ from Nash implementation beyond the additional requirement of strong Pareto optimality.\(^6\)

THEOREM 1. A SCC \( f : \mathcal{R} \rightarrow A \) is implementable in SE if, and only if, there exists a set \( B \subseteq A \) such that (i) \( f \) is strongly Pareto optimal in \( B \) and (ii) \( f \) is Nash implementable in \( B \).

Proof. \((\Rightarrow)\) Assume first that \( f : \mathcal{R} \rightarrow A \) is implementable in SE. Let \( G = (S_1, \ldots, S_n; g) \) be a game form that implements \( f \). Choose \( B = g(S) = \{g(s) \mid s \in S\} \) as the set of all alternatives that are attainable through the

\(^6\)This result holds also for the solution concept that is usually used in the literature on strong implementation, we just have to replace the requirement of strong Pareto optimality by weak Pareto optimality.
game form. For the sake of contradiction, assume that $f$ is not strongly Pareto optimal in $B$. This implies that there must exist $R \in \mathcal{R}$, $a \in f(R)$ and $b \in B$, such that $bR_i a$ for all $i \in N$ and $bP_i a$ for some $i \in N$. However, by the definition of $B$ there must then exist $s \in S$, such that $g(s) = b$, so the grand coalition $C = N = \{1, 2\}$ would deviate from any strategy profile that gives the outcome $a$ when the true preference profile is $R$. This is a contradiction with the fact that $a$ must be the outcome of some strong equilibrium when the preference profile is $R$ since $G$ implements $f$ in SE. The last claim that $f$ is also Nash implementable in $B$ follows directly from the observation that $SNE(G, R) = NE(G, R)$ must hold for all $R \in \mathcal{R}$ when $f$ is strongly Pareto optimal in $g(S)$ and $n = 2$.

$(\Leftarrow)$ Assume that $B \subseteq A$ satisfies both (i) and (ii). Let $G = (S_1, \ldots, S_n; g)$ be a game form that implements $f$ in Nash equilibrium and satisfies $g(S) \subseteq B$. This must exist by (ii). Since $f$ is strongly Pareto optimal in $B$ by (i), it must also be strongly Pareto optimal in the smaller set $g(S)$. As before, this implies that $SNE(G, R) = NE(G, R)$ must hold for all $R \in \mathcal{R}$. Therefore, $G$ implements $f$ also in SE. This concludes the proof.

The message of this theorem is that a full characterization of SCCs that are strongly implementable follows directly from the equivalent characterization for Nash implementable SCCs when there are only two individuals. Since the latter is well-known (see Moore and Repullo, 1990, and Dutta and Sen, 1991b), there is not much to do here. However, the problem gets more complicated when more individuals are involved.

4 The Considerably More Difficult Case of at Least Three Players

What is it that makes the case of three or more individuals more difficult than the case of two individuals when it comes to strong implementation? When there are only two individuals, essentially two things can happen. The individuals can either compete, that is they can play a Nash equilibrium, or
cooperate against the planner, that is cooperate against the designer of the game form by forming the grand coalition. What cannot happen is that a set of individuals cooperate to compete against the rest of the individual, something that can happen only when there are at least three individuals (while in this case individuals can also compete and play a Nash equilibrium or cooperate against the planner by forming the grand coalition). Theoretically, or more precisely geometrically, the difference is again simply an intersection condition (see Dutta and Sen, 1991a, or Suh, 1996). The most natural way to get rid of all those intersection conditions it is to make the following assumption.

**Holocaust Alternative** (McKelvey, 1989).\(^7\) There exists an alternative \(a_H \in A\), with \(a_H \in f(R)\) for no \(R \in \mathcal{R}\), such that for all \(R \in \mathcal{R}\), and all \(a \in A \setminus \{a_H\}\), \(a \notin L_i(a_H, R)\) and \(a_H \in L_i(a, R)\) for all \(i \in N\). The alternative \(a_H\) is called the *holocaust*.\(^\square\)

If a holocaust exists, then the intersection of any number of lower contour sets is always non-empty. Furthermore, the existence of a holocaust does not change the incentive structure of the implementation problem in any essential way. This is simply due to the fact that it can never be part of a preference reversal.\(^8\) By assuming a holocaust we can get to the core of the strong implementation problem and do not have to mind about the “artificial” restrictions imposed by the geometry of the incentive device (here the normal form game). The idea is that in this way we can reveal something new about the implementation problem. The following property turns out useful when the environment has a holocaust.\(^9\)

**The Axiom of Sufficient Reason.** A SCC \(f : \mathcal{R} \rightarrow A\) satisfies the

---

\(^7\)Dutta and Sen (1991a) calls it the “doomsday alternative”.

\(^8\)A preference reversal is a profile \((a, b, R_i, R'_i)\), such that \(a P_i b\) while \(b R'_i a\).

\(^9\)This has some nice connections with the Principle of Sufficient Reason (PRS) discussed extensively in the philosophy literature (see Belot (2001) or Pruss (2006), for example). Roughly speaking, PSR says that everything must have a reason. I thank Marko Ahteensuu for this remark.
axiom of sufficient reason (ASR), if for all $\mathbf{R}, \mathbf{R}' \in \mathcal{R}$, $a \in f(\mathbf{R})$ and $c \in A$,

$$L_i(a, \mathbf{R}) \subseteq L_i(c, \mathbf{R}') \text{ for all } i \in N \Rightarrow c \in f(\mathbf{R}').$$

(2)

□

This property is stronger than Monotonicity, which is a necessary condition for strong implementation (Maskin, 1979). One can see this by choosing $c = a$. Moreover, this condition says that every alternative which weakly Pareto dominates $a \in f(\mathbf{R})$ should also be in $f(\mathbf{R})$. One can see this by choosing $\mathbf{R} = \mathbf{R}'$. The name of this property stems from the observation that when there is a reason to select $a$ under the preference profile $\mathbf{R}$ and the left hand side of implication (2) holds, then exactly the same reason should guarantee that $c$ is chosen under the preference profile $\mathbf{R}'$ (see Nurmi (2004) for a similar interpretation of Monotonicity). To be more precise, lets say that $(b, i)$ is a reason to select $a$ under the preference profile $\mathbf{R}$ if $a R_i b$. In a similar way, lets say that $(b, i)$ is a reason not to select $a$ under $\mathbf{R}$ if $b P_i a$. With this interpretation in mind ASR states that if every reason to select $a$ under $\mathbf{R}$ is also a reason to select $c$ under $\mathbf{R}'$ and $a$ is indeed selected under $\mathbf{R}$, the $c$ should be selected under $\mathbf{R}'$.

We are now ready to give our main result. We begin with a simple lemma.

**Lemma.** If A SCC $f : \mathcal{R} \rightarrow A$ satisfies ASR, then it is Unanimous.

**Proof.** Let $\mathbf{R} \in \mathcal{R}$ be such that $a R_i b$ for all $i \in N$ and all $b \in A$. Choose any $c \in f(\mathbf{R})$ (not necessarily $a$). Since $L_i(c, \mathbf{R}) \subseteq L_i(a, \mathbf{R})$ for all $i \in N$, we must have $a \in f(\mathbf{R})$ by ASR. □

REMARK. If the game form $G = (S_1, \ldots, S_n; g)$ implements $f$ in SE, then $f$ must be Unanimous and strongly Pareto optimal in the set $g(S)$ of all alternatives that are attainable through the game form. Otherwise the grand coalition $N$ could deviate profitably, to make a Pareto improvement, or it could get stuck into a bad equilibrium, the Unanimous alternative that is not selected by $f$. ||
THEOREM 2. Assume that there exists a holocaust \( a_H \in A \). If the SCC \( f : \mathcal{R} \to A \) is strongly Pareto optimal and satisfies ASR, then it is implementable in SE.

Proof. We construct the mechanism \( G = (S_1, \ldots, S_n; g) \) that implements \( f \). Let \( S_i = A \times \mathcal{R} \times \{F, NF\} \times \mathbb{N}_+ \) and denote a typical element of this set by \( s_i = (a^i, R^i, x^i, n^i) \). Here \( x^i \) is interpreted as individual \( i \) raising a flag \( (x^i = F) \) or not \( (x^i = NF) \). The outcome function \( g \) is defined by the following four rules:

1. If \( s_i = (a, R, NF, n^i) \) for all \( i \in N \) and \( a \in f(R) \), set \( g(s) = a \).

2. If \( \emptyset \neq C \subset N \), \( a \in f(R) \), \( s_j = (a, R, NF, n^j) \) for all \( j \in N \setminus C \) and \( s_j = (a^j, R^j, F, n^j) \) for all \( j \in C \), set

\[
g(s) = \begin{cases} 
a^i, & \text{if } i = \min \{\arg \max_{j \in C} n^j\} \text{ and } a^i \in \bigcup_{j \in C} L_j(a, R), \\
a, & \text{otherwise}. \end{cases}
\]

3. If \( x^i = F \) for all \( i \in N \), let \( i = \min \{\arg \max_{j \in N} n^j\} \) and set \( g(s) = a^i \).

4. In all other cases, set \( g(s) = a_H \).

Next we show that this game form implements \( f \) in SE. Assume that \( R \) is the true preference profile and \( a \in f(R) \). A strategy profile \( s \), such that \( s_i = (a, R, NF, n^i) \) for all \( i \in N \), is a strong equilibrium of this game form and \( g(s) = a \). The grand coalition \( N \) cannot deviate profitably since \( f \) is strongly Pareto optimal and for any other coalition \( C \subset N \) the outcome would be in the lower contour set of \( a \) for at least one individual by rule (2). Therefore, we have \( f(R) \subseteq SE(G, R) \).

To prove this inclusion the other way around, we need to check three mutually exclusive cases depending on which rule (1)-(3) is used to calculate the outcome (notice that rule (4) cannot be used since the grand coalition \( N \) would deviate by the definition of \( a_H \)). Assume that rule (1) is used. Then \( s_i = (a, R', NF, n^i) \) for all \( i \in N \) and \( a \in f(R') \) but \( R' \neq R \). The fact that also \( a \in f(R) \) must hold follows from Monotonicity, which is a special case of ASR. Assume then that rule (2) is used and let \( C \subset N \) be the
coalition that has raised a flag. Denote $s_j = (a, R', NF, n^j)$ for all $j \in N \setminus C$, $s_j = (a^j, R^j, F, n^j)$ for all $j \in C$ and $g(s) = a^j$. Remember that $a \in f(R')$.

By rule (2) and (3), we must now have $L_i(a, R') \subseteq L_i(a^i, R)$ for all $i \in N$, so that $a^i \in f(R)$ must hold by ASR as required. Finally, assume that rule (3) is used. Now the outcome $g(s) = a^i$ must be the best alternative for all individuals and therefore $a^i \in f(R)$ holds by Unanimity (see the Lemma above). This completes our proof. ■

Using a similar idea as in Theorem 1, we can get also the following slightly more general result.

**Corollary 1.** Assume that there exists a holocaust $a_H \in A$. A SCC $f : R \rightarrow A$ is implementable in SE if there exists a set $B \subseteq A$, such that (i) $\text{Im}(f) \cup \{a_H\} \subseteq B$, (ii) $f$ is strongly Pareto optimal in $B$ and (iii) $f$ satisfies ASR in $B$.

**Proof.** Use the game form given in Theorem 2 but replace the message space $S_i = A \times R \times \{F, NF\} \times N_+$ of individual $i$ with $S_i = B \times R \times \{F, NF\} \times N_+$. The rest of the proof is exactly the same. ■

If we think more deeply about the function that holocaust performs in the game form we used to prove Theorem 2 and Corollary 1, it is evident that we can replace its existence with weaker conditions. The subsequent two are obvious.\(^{11}\)

**Global Intersection of Strict Lower Contour Sets.** We say that a SCC $f : R \rightarrow A$ satisfies the *global intersection of strict lower contour sets* (GISL) if

$$\bigcap_{R \in R} \bigcap_{a \in f(R)} \bigcap_{i \in N} SL_i(a, R) \neq \emptyset.$$  

\(^{11}\) A strict lower contour set is defined as $SL_i(a, R) = \{b \in A \mid aP_i b\}$.

\(^{10}\) If $N \setminus C = \{j\}$, then this must hold because every individual $i \in N \setminus \{j\}$ can choose any alternative from $L_i(a, R')$ by rule (2) and individual $j$ can choose any alternative from $L_j(a, R')$ by rule (3). This is because individual $j$ is now the only one who has not raised a flag. If $|N \setminus C| \geq 2$, on the other hand, then evey individual $i \in N$ can choose any alternative from $L_i(a, R')$ simply by rule (2).
Universally Pareto Dominated Alternative. We say that a SCC $f : \mathcal{R} \rightarrow A$ has a universally Pareto dominated alternative (UPDA) if there exist an alternative $a_0 \in A$, such that for all $R \in \mathcal{R}$

$$a_0 \in sPO(R) \Rightarrow a_0 \in f(R)$$

If a holocaust exists, then both GISL and UPDA are clearly satisfied. However, these can hold more generally due to the particular properties of a SCC. For example, let’s consider matching problems (see Gale and Shapley (1962) for the original treatment). Following the terminology of Kara and Sönmez (1996), this is an ordered triplet $(M, W, \mathcal{R})$, where $M$ and $W$ are two non-empty, finite and disjoint sets, and $\mathcal{R} = (R_i)_{i \in M \cup W}$ is a preference profile such that $R_i$ is a linear order on $M \cup \{i\}$ if $i \in W$ and a linear order on $W \cup \{i\}$ if $i \in M$. A matching, which is a solution to this matching problem, is any function $\mu : M \cup W \rightarrow M \cup W$ such that for all $m \in M$, if $\mu(m) \neq m$ then $\mu(m) \in W$, and for all $w \in W$, if $\mu(w) \neq w$ then $\mu(w) \in M$, and furthermore $\mu(\mu(i)) = i$ for all $i \in M \cup W$.\footnote{If $m$ is matched with $w$, then also $w$ is matched with $m$ and vice versa.} The match $\mu(i) = i$ is interpreted as individual $i$ staying a single.

Let’s denote the set of all possible matchings by $\mathcal{M}$. In the context of matching problems a SCC is any correspondence $f : \mathcal{R} \rightarrow \mathcal{M}$ that associates all matching problems $(M, W, \mathcal{R})$ with a set of matchings $f(\mathcal{R}) \subseteq \mathcal{M}$.\footnote{Notice that here $M$ and $W$ are fixed, and hence a matching problem $(M, W, \mathcal{R})$ can be identified with the preference profile $\mathcal{R}$.} To make this an implementation problem, we have to extend the preferences of individuals into the set of matchings $\mathcal{M}$. Luckily, there is a natural way to do this: For any $\mu, \mu' \in \mathcal{M}$, set $\mu R_i \mu'$ if and only if $\mu(i) R_i \mu'(i)$. In other words, individual $i$ weakly prefers matching $\mu$ to matching $\mu'$ if she weakly prefers her match under $\mu$ to her match under $\mu'$.

There are some properties that a good matching should satisfy. An obvious one is that a matching should be individually rational. That is, $\mu(i) R_i i$ should hold for all $i \in M \cup W$. If a matching is not individually rational,
then some individual is forced into a match although she would rather stay single. We are finally ready to make the point. Assume that the matching environment is such that under any admissible preference profile there exists a pair \((i, j) \in M \times W\) of individuals who both think that they should rather be matched together than stay single. Then, if a SCC selects only individually rational matchings, UPDA is satisfied by the identity matching \(id : M \cup W \rightarrow M \cup W, id(i) = i\). However, GISL does not necessarily hold even in this restricted environment. This is so because \(id\) does not belong to the set defined in Equation (3) unless everyone is matched under any preference profile. In particular, this would require that \(|M| = |W|\) must hold.

It's time to state our final result.

**Corollary 2.** A SCC \(f : \mathcal{R} \rightarrow A\) is implementable in SE if there exists a set \(B \subseteq A\), such that (i) \(\text{Im}(f) \subseteq B\), (ii) \(f\) is strongly Pareto optimal in \(B\), (iii) \(f\) satisfies GISL or UPDA in \(B\) and (iv) \(f\) satisfies ASR in \(B\).

**Proof.** The game form used in Corollary 1 works if we replace the holocaust \(a_H\) with any alternative from the set

\[
\bigcap_{R \in \mathcal{R}} \bigcap_{a \in f(R)} \bigcap_{i \in N} \text{L}_i(a, R) \bigcap B \neq \emptyset
\]

when GISL holds and replace it with the alternative \(a_0\) when UPDA holds. If we do this, there still cannot be any equilibria under rule (4) of the game form. Everything else follows directly from Corollary 1. ■

5 A Few Examples

In this final section we give a few examples that demonstrate the applicability of these results in some important cases. These confirm that the conditions really have a bite.

**EXAMPLE 1 (Individually Rational and Pareto Optimal SCCs).** Fix an alternative \(a_0 \in A\) and define \(\text{IR}_i(a_0, R) = \{a \in A \mid a R_i a_0\}\). This is the
set of all alternatives that are at least as good as $a_0$ for individual $i$. The set

$$\text{IR}(a_0, R) \equiv \bigcap_{i \in N} \text{IR}_i(a_0, R)$$

is called the \textit{individually rational correspondence} with respect to $a_0$. This correspondence satisfies UPDA, since $a_0 \in \text{sPO}(R)$ implies $\text{IR}(a_0, R) = \{a_0\}$. It is easy to verify that this correspondence satisfies also ASR. However, it does not satisfy strong Pareto optimality, therefore it is not implementable in SE (see the Remark after Theorem 2). Now let's define a new SCC $f : \mathcal{R} \to A$ by the rule

$$f(R) = \text{sPO}(R) \cap \text{IR}(a_0, R).$$

Notice that $f$ is really a SCC. Indeed, if $a_0$ is strongly Pareto optimal under $R$, then $f(R) = \{a_0\}$. Hence $f$ is always non-empty. It is also easy to verify that sPO satisfies ASR. Assume that $a \in \text{sPO}(R)$ and $L_i(a, R) \subseteq L_i(c, R')$ for all $i \in N$. If $c \in \text{sPO}(R')$ would not hold, then $a$ was certainly not strongly Pareto optimal after all. Therefore, $f$ must satisfy ASR because both IR and sPO satisfy it. On top of this, $f$ must satisfy also UPDA for exactly the same reason as IR did. The fact that $f$ is implementable in SE follows now from Corollary 2 (where $B = A$). □

EXAMPLE 2 (The Core Correspondence). A binary relation $B$ between $2^N$ and $2^A$ is called a \textit{blocking} if it satisfies the following three conditions:

(i) if $KBX$, then $K'BX'$ for all sets $K'$ and $X'$ such that $K \subseteq K'$ and $X' \subseteq X$, (ii) if $KBX$, $K'BX'$ and $K \cap K' = \emptyset$, then $(K \cup K')B(X \cup X')$ and (iii) every coalition blocks the empty set, while no coalition blocks $A$ (see Danilov and Sotskov (2002) for more details).

We say that coalition $K$ \textit{rejects} an alternative $a$ if it is able to block the set $\bigcup_{i \in K} L_i(a, R)$. An alternative which is not rejected by any coalition is called \textit{stable}. The set of all stable outcome under $R$ is called the \textit{core} and denoted by $C(B, R)$. Using this notation we define the \textit{core correspondence} by the rule:

$$C : \mathcal{R} \to A, \quad C(R) = C(B, R).$$
This correspondence does not always qualify as a SCC since it can sometimes be empty. However, it satisfies ASR on any domain $\mathcal{R}$. Indeed, assume that $a \in C(B, R)$ and $L_i(a, R) \subseteq L_i(c, R')$ for all $i \in N$. If coalition $K$ rejects $c$, then it must be able to block $\bigcup_{i \in K} L_i(c, R)$. By condition (ii) coalition $K$ must then be able to block also the set $\bigcup_{i \in K} L_i(a, R)$, since

$$
\bigcup_{i \in K} L_i(a, R) \subseteq \bigcup_{i \in K} L_i(c, R').
$$

This cannot be the case since $a$ is stable under $R$. Therefore, no coalition $K$ can reject the set $\bigcup_{i \in K} L_i(c, R')$ and hence $c \in C(B, R')$ as required by Equation (2).

So if the environment is such that $C(B, \cdot)$ is really a SCC (non-empty), strongly Pareto optimal (often the case but can fail) and there exist a holocaust (or either GISL or UPDA holds), then $C(B, \cdot)$ is implementable in SE by Theorem 2 or Corollary 2. □

**EXAMPLE 3 (Stable Matching Rules).** A matching $\mu$ is blocked by the pair $(m, w) \in M \times W$ under $R$ if

$$
w R_{m, \mu}(m) \quad \text{and} \quad m R_{w, \mu}(w)
$$

while the relation is strict for at least one individual. This is a little bit different from the one used in the matching literature, one could call this a **blocking with side payments** (see Kara and Sönmetz (1996) or Kara and Sönmetz (1997) for the standard definition). A matching $\mu$ is called stable if it is individually rational and not blocked by any pair $(m, w) \in M \times W$. Lets denote the set of all stable matchings under $R$ by $\mathcal{S}(R)$. Remember that preferences are now linear orders. Gale and Shapley (1962) has shown that the set $\mathcal{S}(R)$ in non-empty for all $R \in \mathcal{R}$.

Let $\mathcal{R}$ be the set of all admissible preference profiles for a matching problem. The **stable matching rule** is a SCC $S : \mathcal{R} \rightarrow \mathcal{M}$ such that $f(R) = \mathcal{S}(R)$ for all $R \in \mathcal{R}$. It is easy to see that the stable matching rule is implementable in

\[14\] The deferred acceptance algorithm gives a stable matching also in this case.
SE when UPDA holds. We have already discussed a simple domain restriction that would guarantee this. First of all, $S$ is strongly Pareto optimal. For the sake of contradiction, assume not. Let $\mu, \mu' \in \mathcal{M}$ be two matchings, such that $\mu \in f(R)$ while $\mu'(i) R_i \mu(i)$ for all $i \in M \cup W$ and $\mu'(i) P_i \mu(i)$ for at least one $i \in M \cup W$. Choose any $m \in M$, such that $\mu'(m) \neq m$. Since $\mu'$ is a matching, we must have $\mu'(w) = m$ for $w = \mu'(m)$. But then $\mu$ is blocked by the pair $(m, w)$ — a contradiction with the stability of $\mu$. Therefore, $\mu$ has to be strongly Pareto optimal. Second, it satisfies also ASR. Assume that $\mu \in f(R)$ and $L_i(\mu, R) \subseteq L_i(\mu', R')$ for all $i \in M \cup W$, so that the right hand side of the implication in Equation (2) holds. It is obvious that $\mu'$ must be individually rational, since $\mu$ is, and stable, otherwise $\mu$ could not be stable either (a blocking for $\mu'$ constitutes a blocking for $\mu$). The fact that stable matching rule is implementable in SE follows now from Corollary 2 (where we have chosen $B = \mathcal{M}$).

6 Concluding Comments

We have derived a new sufficient condition for a social choice correspondence to be implementable in strong equilibrium. The property starring this condition, called the Axiom of Sufficient Reason, is often easy to verify and has a nice and intuitive interpretation. In addition, it is appealing also as a normative criterion, and therefore one may speculate whether a good SCC should satisfy it quite irrespective of the implementation goal. In other words, one might claim that the goal of a society should satisfy ASR whether individuals are acting sincerely or not.

As with any sufficient condition, only the scope of applications can work as a final test for ASR. We have shown that ASR passes the test easily, although the preference domain usually has to be restricted somehow. However, the set of all social alternatives $A$ is often not fixed, and hence there is a possibility to design also new alternatives. If and when this is the case, our restrictions on the preference domain may not be that severe.
References


