Acyclicity, Robust Stability and Nash Implementation

Taro Kumano*  Masahiro Watabe†
Department of Economics  Department of Economics
Washington University in St.Louis  Zirve University

September 13, 2010

Abstract

Several school choice districts in the United States adopt the student-proposing deferred acceptance algorithm as a centralized matching procedure to assign students to schools. We model schools axiomatically in terms of choice functions, unlike the existing literature which assumes that schools accept students so as to maximize their priorities. The paper studies the implications of a further restriction on choice functions for school choice problems. Under substitutable and quota-filling choice functions, acyclicity of the priority structure and pairwise robust stability of the deferred acceptance algorithm are equivalent. We also show that Pareto efficiency and pairwise robust stability are equivalent for any stable and strategy-proof rules. We further establish the equivalence between acyclicity and Nash implementability of the set of stable assignments.

Keywords  Deferred acceptance algorithm · Acyclicity · Robust Stability · Nash implementation

JEL Classification  C62 · C78 · D78

1 Introduction

Several school choice districts in the United States adopt the student-proposing deferred acceptance algorithm as a centralized matching procedure to assign students to schools in NYC and Boston.1 Students are assigned to schools on the basis of priorities. A matching is stable if it is not blocked by any individual student or any student-school pair. In one-to-one matching problems, Gale and Shapley [9] show that if students list their preferred schools then the deferred acceptance algorithm finds a stable matching that any student weakly prefers to any

---

*Email: tkumano@wustl.edu
†Email: masahiro.watabe@zirve.edu.tr

1 In what follows, we use the deferred acceptance algorithm for the sake of shorthand.
other stable matching, called the student-optimal stable matching. Roth and Sotomayor [19] generalize this result to a many-to-one matching problems where priority structures satisfy substitutability.²

The analysis of strategic behavior is of importance to support the use of the deferred acceptance algorithm in practice. In one-to-one matching problems, Dubins and Freedman [6] and Roth [18] show that it is strategy-proof for the students to list their preferred schools in the deferred acceptance algorithm. Hatfield and Milgrom [11] generalize this result to many-to-one matching problems where priority structures satisfy substitutability and the law of aggregate demand. We assume that priority structures satisfy a slightly stronger notion of the law of aggregate demand, called quota-fillingness, introduced by Alkan [3] and Alkan and Gale [5]. Therefore, the deferred acceptance algorithm is well-defined and strategy-proof in our setting. Due to the nature of school choice problems, it is reasonable to assume that each object accepts students unless its quota is entirely filled.

The paper studies the implications of a further restriction on priority structures, called acyclicity. Ergin [8] introduces the acyclical priority structure as a necessary and sufficient condition for Pareto efficiency of the deferred acceptance algorithm for a subset of priorities satisfying substitutability and the law of aggregate demand. Ergin-acyclicity is a restriction to linear orders over the power set of students, rather than to choice functions. Kumano [15] generalizes the equivalence result in Ergin [8] for substitutable and quota-filling choice functions, in which each choice function is defined as the maximal set induced by its priority. In the paper, we do not assume that schools have a priority which is a linear order over the sets of students. It is natural not to assume that schools accept students so as to maximize their priorities but to model them axiomatically.

The first part of the paper examines a notion of (pairwise) robust stability introduced by Kojima [13]. One implication of eliminating a certain cycle of the priority structure is the following. The priority structure is acyclical if and only if the deferred acceptance algorithm satisfies pairwise robust stability (Theorem 1). In addition, we show that Pareto efficiency and pairwise robust stability are equivalent for any stable and strategy-proof rules (Proposition 1). The second part characterizes the set of stable assignments as Nash equilibrium outcomes of the preference revelation game induced by the deferred acceptance algorithm. Recall that the deferred acceptance algorithm produces a stable assignment for the stated preferences, which is not necessarily stable with respect to the true preferences. The direct mechanism consisted of the deferred acceptance algorithm might have multiple Nash equilibria, some of which produce an unstable assignment with respect to the true preferences. Another implication of eliminating cycle is that the priority structure is acyclical if and only if the stable correspondence is Nash-implementable by the direct mechanism associated by the deferred acceptance algorithm (Theorem 2).

² This is first introduced by Kelso and Crawford [12].
2 Model

Denote by \( N \) the finite set of agents and by \( A \) the finite set of indivisible object types. Let 
\[ q = (q_a)_{a \in A}, \] 
where \( q_a \in \mathbb{Z}_+ \), be the number of available objects of type \( a \). A preference profile is a vector of linear orders \( R = (R_i)_{i \in N} \), where \( R_i \) denotes the preference of agent \( i \) defined over \( X_i = A \cup \{\emptyset\} \). The symbol \( \emptyset \) stands for being unassigned. The asymmetric part of \( R_i \) is denoted by \( P_i \). Let \( \mathcal{R} = \prod_{i \in N} \mathcal{R}_i \) be the set of all preference profiles. Finally, an object \( a \) is acceptable to agent \( i \) if a \( P_i \emptyset \).

For each object \( a \), define \( X_a = \{ S \subseteq N \mid S \preceq q_a \} \). A choice function \( C_a(\cdot) \) is a relation of the power set of \( N \) into \( X_a \) satisfying \( C_a(S) \subseteq S \) for every \( S \subseteq N \). Each profile \( (C_a(\cdot))_{a \in A} \) of choice functions is referred to as a priority structure. There are several admissible restrictions on the class of priority structures.

**Definition 1.** A choice function \( C_a(\cdot) \) is substitutable if for every pair \( (S, T) \) of subsets of \( N \) with \( S \subseteq T \), \( C_a(T) \cap S \subseteq C_a(S) \).

Substitutability is discussed in a labor market model by Kelso and Crawford [12]. This condition simply says that if an agent is admitted by an object from a larger set of agents, then he must be admitted by the same object from any subset of agents including him. The following notion is discussed in Alkan [3] and Alkan and Gale [5].

**Definition 2.** A choice function \( C_a(\cdot) \) is quota-filling if \( | C_a(S) | = \min \{ | S |, q_a \} \) for every \( S \subseteq N \).

Throughout the paper, we assume that every choice function is substitutable and quota-filling. Alkan [4] introduces a weaker notion of quota-fillingness.

**Definition 3.** A choice function \( C_a(\cdot) \) is cardinally monotonic if for every pair \( (S, T) \) of subsets of \( N \) with \( S \subseteq T \), \( | C_a(S) | \leq | C_a(T) | \).

An assignment is a function \( \mu : N \to A \cup \{\emptyset\} \) satisfying: (i) for every agent \( i \), \( \mu(i) \in X_i \) and (ii) for every object \( a \), \( \{ i \in N \mid \mu(i) = a \} \leq q_a \). Denote by \( X \) the set of assignments. A rule \( g \) is a function of \( \mathcal{R} \) into \( X \). If \( g(R) = \mu \) for some \( R \in \mathcal{R} \), then denote \( g_i(R) = \mu(i) \) for every agent \( i \). An assignment \( \mu \) is stable for \( R \) if it satisfies the following conditions: (i) for every agent \( i \), \( \mu(i) \in P_i \emptyset \) and (ii) there does not exist \( (i, a) \in N \times A \) such that \( a \in \mu \) and \( i \in C_a(\{ k \in N \mid \mu(k) = a \} \cup \{ i \}) \).

Denote by \( \varphi_S(R) \) the set of stable assignments for \( R \). A rule \( g \) is stable if \( g(R) \in \varphi_S(R) \) for every \( R \in \mathcal{R} \). The relation \( \varphi_S \) of \( \mathcal{R} \) into \( X \) is referred to as the stable correspondence. Finally, an assignment \( \mu \) is Pareto efficient for \( R \) if there does not

---


4 Condition (i) is the individual rationality and condition (ii) is the pairwise stability. Nothing is lost by not considering larger groups (see Roth and Sotomayor [19, p.174]).
exist an assignment \(\eta\) such that \(\eta(i) R_i \mu(i)\) for every agent \(i\) and \(\eta(j) P_j \mu(j)\) for some agent \(j\). A rule \(g\) is Pareto efficient if \(g(R) \in \varphi_{PE}(R)\) for every \(R \in \mathcal{R}\). Denote by \(\varphi_{PE}(R)\) the set of Pareto efficient assignments for \(R\).

We introduce some notions regarding the properties of rules. A rule \(g\) is strategy-proof if for every \(R \in \mathcal{R}\), every agent \(i\) and every \(R_i' \in \mathcal{R}_i\), we have \(g_i(R) R_i g_i(R_{-i}, R_i')\). For each assignment \(\mu\) and each \(R_i \in \mathcal{R}_i\), define \(L_i(\mu, R_i) = \{a \in X_i \mid \mu(i) R_i a\}\). A rule \(g\) is Maskin monotonic if for every \((R, R') \in \mathcal{R} \times \mathcal{R}\), if \(L_i(g(R), R_i) \subseteq L_i(g(R), R_i')\) for every agent \(i\), then \(g(R) = g(R')\). Finally, a rule \(g\) is nonbossy if for every agent \(i\), every \(R \in \mathcal{R}\), and every \(R_i' \in \mathcal{R}_i\), if \(g_i(R) = g_i(R_{-i}, R_i')\), then \(g(R) = g(R_{-i}, R_i')\).

The deferred acceptance algorithm proposed by Gale and Shapley [9] has been adopted into a practical assignment procedure. At the first step, each agent applies to his most preferred acceptable object. Let \(N^1_a\) be the set of agents applying to object \(a\) at the first step. Object \(a\) tentatively accepts \(C_a(N^1_a)\) and rejects the remaining. At the \(r\)th step, each agent who was rejected at step \(r - 1\) applies to his next preferred acceptable object. Let \(N^r_a\) be the set of agents applying to object \(a\) at step \(r\). Object \(a\) tentatively accepts \(C_a(N^r_a \cup N^r_{a'}\) and rejects the remaining. The algorithm terminates when every agent is held tentatively by some object or has been rejected by every object that is acceptable for him. Each agent is assigned the object if he is tentatively held at the last step, otherwise assigned nothing. The above explanation can be found in Roth and Sotomayor [19, Chapter 5, pp. 134-5]. It is known that under any substitutable choice functions, the deferred acceptance algorithm yields a unique stable assignment that is Pareto superior to any other stable assignment, called the agent-optimal stable assignment (see Roth and Sotomayor [19, Theorem 6.8]). We denote by \(f\) the deferred acceptance algorithm.

Let us briefly summarize the properties of the deferred acceptance algorithm in our setting. Abdulkadiroğlu [1] finds that substitutability of choice functions itself is not sufficient for the existence of a strategy-proof stable rule. Hatfield and Milgrom [11] show that substitutability coupled with cardinal monotonicity are sufficient for strategy-proofness of the deferred acceptance algorithm in our setting.\(^5\) Therefore, the deferred acceptance algorithm is strategy-proof in the paper. Among stable rules, Alcalde and Barberà [2] claim that only the (agent-proposing) deferred acceptance algorithm is strategy-proof for a special case of substitutable and cardinally monotonic priority structures. Sakai [20] reports that their result holds for substitutable and cardinally monotonic priority structures.

In the existing literature, acyclical conditions of priority structures have been discussed. Ergin [8], Kojima and Manea [14], and Kumano [15] assume that the priority structure is a profile of linear orders of subsets of agents, that is, priorities are complete, transitive and antisymmetric. And then, they define the corresponding choice functions as the maximal set from a subset of agents with respect to the priority. The interpretation is that every object selects the most preferred set of agents for every set of agents. In the context of school choice, however, objects are merely indivisible goods to be consumed, not active agents. Our interpretation is that each

\(^5\) Hatfield and Milgrom [11] say that a profile of choice functions satisfies the law of aggregate demand if the priority structure \((C_a(\cdot))_{a \in A}\) is cardinally monotonic.
object accepts a set of agents as long as there is no violation of a combination of restrictions to its priority. We assume neither that each object has a linear order over the entire set of agents $N$ nor that each object maximizes its priorities as its preference.

To end this section, we summarize some properties of choice functions needed in the paper.

**Lemma 1.** Every substitutable and cardinally monotonic choice function $C_a(\cdot)$ satisfies the following: for every pair $(S, T)$ of elements of $N$,

1. $C_a(C_a(S) \cup T) = C_a(S \cup T)$. (Path-independence)
2. If $C_a(S) \subseteq T \subseteq S$ then $C_a(S) = C_a(T)$. (Consistency)

We omit the proof of the Lemma. It is well-known that both properties are satisfied if $C_a(S)$ is defined as the most preferred subset of $S$, however, this is not our case.

### 3 Implications of Acyclical Priority Structure

The following notion of acyclicity is introduced by Ergin [8]. For each object $a$, denote by $\succeq_a$ a linear order over the power set of $N$. The asymmetric part of $\succeq_a$ is denoted by $\succ_a$.

**Definition (Ergin [8]).** An Ergin-cycle is constituted of distinct $\{i, j, k\} \subseteq N$ and $\{a, b\} \subseteq A$ such that the following are satisfied:

1. Cycle condition: $i \succ_a j \succ_a k \succ_b i$, and
2. Scarcity condition: There exist $\{S_a, S_b\} \subseteq N \setminus \{i, j, k\}$ with $S_a \cap S_b = \emptyset$ such that $S_a \subseteq \{\ell \in N \mid \ell \succ_a j\}$, $S_b \subseteq \{\ell \in N \mid \ell \succ_b i\}$, $|S_a| = q_a - 1$ and $|S_b| = q_b - 1$.

We define acyclicity in terms of choice functions.

**Definition 4.** A cycle is constituted of distinct $\{i, j, k\} \subseteq N$ and $\{a, b\} \subseteq A$ such that

- $C_a(S_a \cup \{i, j, k\}) = S_a \cup \{i\}$,
- $C_a(S_a \cup \{j, k\}) = S_a \cup \{j\}$, and
- $C_b(S_b \cup \{k, i\}) = S_b \cup \{k\}$

for some $\{S_a, S_b\} \subseteq N \setminus \{i, j, k\}$ with $S_a \cap S_b = \emptyset$.

We do not impose Scarcity condition in Ergin [8] explicitly. It is not difficult to see that under quota-fillingness, it must be the case that $|S_a| = q_a - 1$ and $|S_b| = q_b - 1$ in our definition. Studying substitutable choice functions has an advantage. Ergin-acyclicity is defined for a subset of substitutable and quota-filling priorities, called responsive priorities. Echenique [7] shows that the substitutable and consistent choice functions are exponentially more than the responsive priorities.\(^6\)

We will employ the following properties of the deferred acceptance algorithm and our notion of acyclicity of a priority structure to establish our main results below.

---

\(^6\) Echenique [7] refers to consistency as *independence of irrelevant alternatives.*
Remark 1. For every substitutable and quota-filling choice functions, the following are equivalent:
(1) the priority structure is acyclical.
(2) the deferred acceptance algorithm is Pareto efficient.
(3) the deferred acceptance algorithm is Maskin monotonic.
(4) the deferred acceptance algorithm is nonbossy.

Proof. Firstly, Proposition 1 in Kumano [15] establishes the equivalence between Pareto efficiency of the deferred acceptance algorithm and acyclicity of priority structure under substitutable and quota-filling choice functions when each choice function is defined as the maximal set induced by a linear order over the power set of agents. Our acyclicity and his acyclicity are identical in our setting. His proof relies on the axiom of path-independence of choice functions, which is satisfied in our setting by Lemma 1 (1). The remaining properties of the deferred acceptance algorithm follow from Proposition 1 in Kojima and Manea [14] and Lemma 1 in Pápai [17]. This establishes the remark.

Watabe [21] finds that the equivalence between our notion of acyclicity of priority structures and Pareto efficiency of the deferred acceptance algorithm crucially depends on the class of priority structures. We are not able to weaken quota-fillingness to cardinal monotonicity.

3.1 Robust Stability

In a restricted class of priority structures than ours, Kojima [13] considers the following notion.

Definition (Kojima [13]). A rule \( g \) is robustly stable if the following conditions are satisfied:
(1) \( g \) is stable,
(2) \( g \) is strategy-proof, and
(3) there exists no \( i \in N, a \in A, R \in \mathcal{R} \) and \( R_i \in \mathcal{R}_i \) such that (i) \( a \succ_i g_i(R) \) and (ii) \( i \sim_a j \) for some \( j \in \{k \in N \mid g_k(R_{-i}, R'_i) \} \) or \( \{k \in N \mid g_k(R_{-i}, R'_i) \} < q_a \).

We rewrite robust stability in terms of choice function to our rich class of priority structures. It is not difficult to see that the following notion encompasses condition (3) in Kojima [13].

Definition 5. A rule \( g \) is immune to a pairwise counter proposal if for every \( R \in \mathcal{R} \), there is no \( (i, a) \in N \times A \) and \( R'_i \in \mathcal{R}_i \) such that \( a \succ_i g_i(R) \) and \( i \in C_a(\{k \in N \mid g_k(R_{-i}, R'_i) = a\} \cup \{i\}) \).

Definition 6. A rule \( g \) satisfies pairwise robust stability if it is stable, strategy-proof, and immune to any pairwise counter proposal.

The following theorem generalizes Theorem 2 in Kojima [13].

Theorem 1. For every substitutable and quota-filling choice functions, the priority structure is acyclical if and only if the deferred acceptance algorithm satisfies pairwise robust stability.
Proof. Suppose first that the deferred acceptance algorithm \( f \) satisfies pairwise robust stability. Suppose, by way of contradiction, that the priority structure is acyclical. There exists distinct \( \{i, j, k\} \subseteq \mathcal{N} \) and \( \{a, b\} \subseteq \mathcal{A} \) such that \( C_a(S_a \cup \{i, j, k\}) = S_a \cup \{i\} \), \( C_a(S_a \cup \{j, k\}) = S_a \cup \{j\} \), and \( C_b(S_b \cup \{k, i\}) = S_b \cup \{k\} \) for some \( \{S_a, S_b\} \subseteq \mathcal{N} \setminus \{i, j, k\} \) with \( S_a \cap S_b = \emptyset \).

Consider the following preference profile. For those three agents, \( R_i : b P_i a P_i \emptyset, R_j : a P_j \emptyset, \) and \( R_k : a P_k b P_k \emptyset \). Agents in \( S_a \) and \( S_b \) respectively rank \( a \) and \( b \) as the only acceptable object. Finally, agents in \( I = \mathcal{N} \setminus \{\{i, j, k\} \cup S_a \cup S_b\} \), set \( \emptyset \) as their top choice. In addition, consider the preference profile \( (R_{i, j}^f, R_j^f) \), where \( R_j^f : \emptyset \). It is not difficult to see that

\[
f(R) = \begin{pmatrix} i & j & k & S_a & S_b \emptyset \\ a & \emptyset & b & a & b \emptyset \end{pmatrix}
\]

and

\[
f(R_{i, j}^f, R_j^f) = \begin{pmatrix} i & j & k & S_a & S_b \emptyset \\ b & \emptyset & a & a & b \emptyset \end{pmatrix}.
\]

Then, \( a P_j f_j(R) \) and \( f_j(R_{i, j}^f, R_j^f) = \emptyset \). Furthermore, \( S_a \cup \{j, k\} = \{\ell \in \mathcal{N} | f_\ell(R_{i, j}^f, R_j^f) = a\} \cup \{j\} \) and \( |\{\ell \in \mathcal{N} | f_\ell(R_{i, j}^f, R_j^f) = a\} | = (S_a \cup \{k\}) | = q_a \) since \( S_a | = q_a - 1 \) and \( k \notin S_a \). By the hypothesis, \( C_a(S_a \cup \{j, k\}) = S_a \cup \{j\} \), and hence \( j \in C_a(\{\ell \in N | f_\ell(R_{i, j}^f, R_j^f) = a\} \cup \{j\}) \).

This, together with \( a P_j f_j(R) \), yields that \( f \) is not immune to pairwise counter proposal, a contradiction.

It remains to show the converse. Suppose next that the priority structure is acyclical. Suppose, by way of contradiction, that \( f \) is not immune to pairwise counter proposal. Consider any \( R \in \mathcal{R} \). Let \( (i, a) \in \mathcal{N} \times \mathcal{A} \) such that \( a P_i f_i(R) \) and \( i \in C_a(\{\ell \in N | f_\ell(R_{i, j}^f, R_j^f) = a\} \cup \{i\}) \) for some \( R_j^f \in \mathcal{R}_j \). There are two possibilities to be considered.

Case 1. \( f_i(R_{i, j}^f, R_j^f) \neq \emptyset \).

Proof of Case 1. Since \( f \) is strategy-proof and \( a P_i f_i(R) \), it follows that \( f_i(R_{i, j}^f, R_j^f) = a \). Let \( R_j^f : a P''_j f_j(R_{i, j}^f, R_j^f) \emptyset \). By strategy-proofness of \( f \), it must be \( f_i(R_{i, j}^f, R_j^f) = \emptyset \}. If \( f_i(R_{i, j}^f, R_j^f) = f_i(R_{i, j}^f, R_j^f) \), then nonbossiness implies that \( f(R_{i, j}^f, R_j^f) = f(R_{i, j}^f, R_j^f) \). This, together with our hypothesis, yields that \( i \in C_a(\{\ell \in N | f_\ell(R_{i, j}^f, R_j^f) = a\} \cup \{i\}) \). This is a contradiction to stability of \( f(R_{i, j}^f, R_j^f) \). If \( f_i(R_{i, j}^f, R_j^f) = \emptyset \), then \( f_i(R_{i, j}^f, R_j^f) P''_j \emptyset = f_i(R_{i, j}^f, R_j^f) \), which contradicts the fact that \( f \) is strategy-proof. This establishes the case.

Case 2. \( f_i(R_{i, j}^f, R_j^f) = \emptyset \).

Proof of Case 2. Immediate from an adaptation of the proof of Theorem 2 in Kojima [13].

This establishes the theorem.

Corollary 1. For every substitutable and quota-filling choice functions, the priority structure is acyclical if and only if any stable and strategy-proof rule is immune to pairwise counter proposal.

Proof. Immediate from Theorem 1 and the fact that the deferred acceptance algorithm is the unique stable and strategy-proof rule under the hypothesis.
Proposition 1. For every substitutable and quota-filling choice functions, any stable rule is strategy-proof and Pareto efficient if and only if it satisfies pairwise robust stability.

Proof. Immediate from Remark 1 and Theorem 1. □

3.2 Preference Revelation Game and Nash Implementation

Up to now, we have analyzed the properties of the deferred acceptance algorithm with respect to the true preferences. But in reality, the true preferences of agents are private information. We calculate an actual assignment from the deferred acceptance algorithm with respect to the submitted preferences. In this subsection, we examine all equilibrium assignments of the preference revelation game induced by the deferred acceptance algorithm in Nash equilibrium. The mechanism designer desires the outcomes described by a rule but does not know preferences that are private information of the agents. The task of the mechanism designer is to construct a procedure independent of private information in order to achieve the prescribed desirable assignment. An ordered pair \((M, h)\) is called a mechanism if \(h\) is a function of \(M\) into \(X\), and \(M = \prod_{i \in I} M_i\), where \(M_i\) is a non-empty set for each agent \(i\). The Cartesian product \(M\) is called the strategy space. Each element \(m \in M\) is called a strategy profile. A triplet \((M, h, R)\) is called a game if \((M, h)\) is a mechanism and \(R \in \mathcal{R}\). We restrict our attention to the class of mechanisms, where \(M_i = R_i\) for every agent \(i\). The resulting games are referred to as preference revelation games.

Given a game \((M, h, R)\), a message profile \(m = (m_{-i}, m_i)\) is called a Nash equilibrium for \((M, h, R)\) if for every agent \(i\), \(h_i(m_{-i}, m_i)R_i h_i(m_{-i}, \tilde{m}_i)\) for every \(\tilde{m}_i \in M_i\). The set of Nash equilibria for \((M, h, R)\) is denoted by \(\mathcal{N}_{(M, h)}(R)\). We can decompose the set of equilibrium messages as \(\mathcal{N}_{(M, h)}(R) = \bigcap_{i \in N} \mathcal{N}_{(M, h)}^i(R_i)\) for every \(R \in \mathcal{R}\), where \(\mathcal{N}_{(M, h)}^i(R_i) = \{m \in M \mid h_i(m_i, m_{-i})R_i h_i(m_{-i}, \tilde{m}_i)\text{ for every }\tilde{m}_i \in M_i\}\) is the graph of agent \(i\)’s best response correspondence at \(R_i\). Notice that each correspondence \(\mathcal{N}_{(M, h)}^i(R_i)\) of \(R_i\) into \(M\) depends only on his own type \(R_i\). In other words, the correspondence \(\mathcal{N}_{(M, h)}\) of \(\mathcal{R}\) into \(M\) is a coordinate correspondence.\(^7\)

Given a mechanism \((M, h)\), we want to identify the composite correspondence \(h \circ \mathcal{O}_{(M, h)}\) of \(\mathcal{R}\) into \(M\) as the actual market outcomes, where the solution concept is \(\mathcal{O}\):

\[
(h \circ \mathcal{O}_{(M, h)})(R) = h(\mathcal{O}_{(M, h)}(R)) = \{h(m) \mid m \in \mathcal{O}_{(M, h)}(R)\}.
\]

The following figure depicts this formulation.

\(^7\) This is a privacy requirement on the equilibrium message correspondence (see Mount and Reiter [16]).
A mechanism \((M, h)\) implements the relation \(\varphi\) of \(\mathcal{R}\) into \(X\) in \(\mathcal{O}\) equilibria if \(h(\mathcal{O}(M,h)(R)) = \varphi(R)\) for every \(R \in \mathcal{R}\).

Haeringer and Klijn [10] show that Ergin-acyclicity is a necessary and sufficient condition for Nash implementation of the stable correspondence \(\varphi_S\). Our theorem generalizes their result to a less restrictive class of priority structures. Nash equilibrium assignments in the preference revelation game induced by the deferred acceptance algorithm coincide with the set of stable assignments for the true preferences.

**Theorem 2.** For every substitutable and quota-filling choice functions, the priority structure is acyclical if and only if the stable correspondence \(\varphi_S\) is Nash-implementable by the direct mechanism \((\mathcal{R}, f)\).

**Proof.** Again, we can pay attention to the deferred acceptance algorithm \(f\) as the outcome function by the uniqueness of stable and strategy-proof rule.

Suppose first that the priority structure is acyclical. We shall show that \(f(\mathcal{N}(\mathcal{R},f)(R)) = \varphi_S(R)\) for every \(R \in \mathcal{R}\). Consider any \(R \in \mathcal{R}\).

**Step 1.** \(f(\mathcal{N}(\mathcal{R},f)(R)) \subseteq \varphi_S(R)\).

**Proof of Step 1.** Consider any \(\mu \in f(\mathcal{N}(\mathcal{R},f)(R))\). Then, \(\mu = f(m)\) for some \(m \in \mathcal{N}(\mathcal{R},f)(R)\). Suppose, by way of contradiction, that \(\mu \not\in \varphi_S(R)\). Obviously, \(\mu(i) R_i \emptyset\) for every agent \(i\), and so there must exist \((i, a) \in N \times A\) such that \(a P_i \mu(i)\) and \(i \in C_a(\{k \in N \mid \mu(k) = a\} \cup \{i\})\).

**Claim 1.** \(f_i(m_{-i}, R_i) = f_i(m_{-i}, R'_i)\).

**Proof of Claim 1.** Set \(R'_i : f_i(m_{-i}, R_i) P'_i \emptyset\). Then, it is not difficult to see that \(f(m_{-i}, R_i) \in \varphi_S(m_{-i}, R'_i)\). If \(f_i(m_{-i}, R_i) = \emptyset\), then it must be the case that \(f_i(m_{-i}, R_i) = f_i(m_{-i}, R'_i)\).

Suppose that \(f_i(m_{-i}, R_i) \neq \emptyset\). If \(f_i(m_{-i}, R'_i) = \emptyset\) then \(f_i(m_{-i}, R_i) P'_i f_i(m_{-i}, R'_i)\). Since \(f(m_{-i}, R'_i)\) is the agent-optimal stable assignment for \((m_{-i}, R'_i)\) and \(f(m_{-i}, R_i)\) is also stable for \((m_{-i}, R'_i)\), it follows that \(f_i(m_{-i}, R'_i) R'_i f_i(m_{-i}, R_i)\). By transitivity of \(R'_i\), \(f_i(m_{-i}, R_i) P'_i f_i(m_{-i}, R_i)\), a contradiction. Therefore, we must have \(f_i(m_{-i}, R'_i) \neq \emptyset\). By definition of \(R'_i\), it must be the case that \(f_i(m_{-i}, R_i) = f_i(m_{-i}, R'_i)\). In either case, \(f_i(m_{-i}, R_i) = f_i(m_{-i}, R'_i)\).

It remains to show that \(f_i(m_{-i}, R'_i) = f_i(m_{-i}, R_i)\). Since \(f\) is strategy-proof, it follows that \(f_i(m_{-i}, R_i) R_i f_i(m_{-i}, m_i)\), and hence \(f_i(m_{-i}, R'_i) R_i f_i(m_{-i}, m_i)\). If \(f_i(m_{-i}, R'_i) \neq \emptyset\), then

\[
\varphi(R) = (h \circ \mathcal{O}(M,h))(R)
\]

\[
\mathcal{O}(M,h)(R)
\]

\[
m \in M
\]

**Figure 1:** Implementation of \(\varphi\) in \(\mathcal{O}\) equilibria
Claim 2. \( f(m_{\cdot i}, R_i) = f(m_{\cdot i}, m_i) \).

Proof of Claim 2. Maskin monotonicity of the deferred acceptance algorithm is essential to establish the claim. Eventually, Maskin monotonicity and acyclicity are equivalent (see Theorem 1 in Kumano [15]).

Denote \( R''_i = m_i \) and \( P''_i \) is the strict part of \( R''_i \). Consider the binary relation \( \tilde{R}_i \) over \( X_i \) that ranks only elements of \( \{ b \in X_i \mid b P_i f_i(m_{\cdot i}, R_i) \} \cap \{ b \in X_i \mid b P''_i f_i(m_{\cdot i}, R''_i) \} \) above \( f_i(m_{\cdot i}, R_i) = f_i(m_{\cdot i}, R''_i) \). That is, \( \tilde{R}_i \) is a monotone transformation of \( R_i \) and \( R''_i \) at \( f_i(m_{\cdot i}, R_i) = f_i(m_{\cdot i}, R''_i) \). Then, \( L_i(f(m_{\cdot i}, R_i), R_i) \subseteq L_i(f(m_{\cdot i}, R_i), \tilde{R}_i) \). By Maskin monotonicity of \( f \), we obtain that \( f(m_{\cdot i}, R_i) = f(m_{\cdot i}, \tilde{R}_i) \). Similarly, \( L_i(f(m_{\cdot i}, R''_i), R''_i) \subseteq L_i(f(m_{\cdot i}, R''_i), \tilde{R}_i) \) yields that \( f(m_{\cdot i}, R''_i) = f(m_{\cdot i}, \tilde{R}_i) \). As a result, \( f(m_{\cdot i}, R_i) = f(m_{\cdot i}, R''_i) = f(m_{\cdot i}, m_i) \). This establishes the claim.

By Claim 2, since \( \mu = f(m_{\cdot i}, m_i) \), we obtain that \( a P_i \mu(i) = f_i(m_{\cdot i}, R_i) \) and \( i \in C_a(\{ k \in N \mid \mu(k) = a \} \cup \{ i \}) = C_a(\{ k \in N \mid f_k(m_{\cdot i}, R_i) = a \} \cup \{ i \}) \). This yields that \( f(m_{\cdot i}, R_i) \not\in \varphi_S(m_{\cdot i}, R_i) \), a contradiction. This establishes the step.

Step 2. \( \varphi_S(R) \subseteq f(\mathcal{N}_{(R,f)}(R)) \).

Proof of Step 2. Consider any \( \mu \in \varphi_S(R) \). For every agent \( i \), set \( R'_i : \mu(i) R'_i \emptyset \). Denote \( m = R' \). By definition of the deferred acceptance algorithm, we have \( f(m) = \mu \). We shall show that \( m \in \mathcal{N}_{(R,f)}(R) \). Consider any agent \( i \). Suppose, by way of contradiction, that there exists \( m'_i \in R_i \) such that \( f_i(m_{\cdot i}, m'_i) P_i f_i(m_{\cdot i}, m_i) \). Since \( f_i(m_{\cdot i}, m_i) R_i \emptyset \), it follows that \( f_i(m_{\cdot i}, m'_i) = a \neq \mu(i) \) for some \( a \in A \). Moreover, since any agent \( k \) other than agent \( i \) ranks \( \mu(k) \) as the top choice under \( m_k \), it follows that \( i \in C_a(\{ k \in N \mid \mu(k) = a \} \cup \{ i \}) \).

But \( a = f_i(m_{\cdot i}, m'_i) P_i f_i(m_{\cdot i}, m_i) = \mu(i) \), which contradicts the fact that \( \mu \in \varphi_S(R) \). This establishes the step.

Steps 1 and 2 yield that \( f(\mathcal{N}_{(R,f)}(R)) = \varphi_S(R) \).

In what follows, we assume that \( f(\mathcal{N}_{(R,f)}(R)) = \varphi_S(R) \) for every \( R \in \mathcal{R} \). We shall show that the priority structure is acyclical. Suppose, by way of contradiction, that there exists distinct \( \{ i, j, k \} \subseteq N \) and \( \{ a, b \} \subseteq A \) such that \( C_a(S_a \cup \{ i, j, k \}) = S_a \cup \{ i \}, C_a(S_a \cup \{ j, k \}) = S_a \cup \{ j \} \), and \( C_b(S_b \cup \{ k, i \}) = S_b \cup \{ k \} \) for some \( S_a, S_b \subseteq N \setminus \{ i, j, k \} \) with \( S_a \cap S_b = \emptyset \).

Consider the preference profile defined in the proof of Theorem 1. Recall that \( a P_j f_j(R_j) \) and \( j \in C_a(\{ \ell \in N \mid f_\ell(R_{\cdot j}, R'_j) = a \} \cup \{ j \}) \). This, together with \( a P_j f_j(R_{\cdot j}, R'_j) \), yields that \( f(R_{\cdot j}, R'_j) \not\in \varphi_S(R) \).

Every agent \( \ell \) other than agent \( j \) has no incentive to deviate from \( R'_\ell \) because he sets \( f_\ell(R_{\cdot j}, R'_j) \) as his top choice under his true preference \( R_\ell \). It remains to consider a possibility of a profitable deviation for agent \( j \) under \( R_j \). It suffices to consider the following messages: \( R''_j : a P''_j \emptyset, b P''_j \emptyset, a P''_j b P''_j \emptyset, \) and \( b P''_j a P''_j \emptyset \). It is not difficult to see that \( f_j(R_{\cdot j}, R''_j) = \emptyset \) in either case. Therefore, \( (R_{\cdot j}, R'_j) \in \mathcal{N}_{(R,f)}(R) \). Since the stable correspondence \( \varphi_S \) is Nash
implementable by the hypothesis, it follows that \( f(R_{-j}, R'_j) \in \varphi_S(R) \), a contradiction. This establishes the theorem.

**Corollary 2.** Let \( g \) be any stable and strategy-proof rule. For every substitutable and quota-filling choice functions, the priority structure is acyclical if and only if the stable correspondence \( \varphi_S \) is Nash-implementable by the direct mechanism \((R, g)\).

**Proof.** Immediate from Theorem 2 and the fact that \( g = f \) under the hypothesis.

**Acknowledgement** We are grateful to Haluk Ergin and an anonymous referee for their helpful comments. All errors are ours.

**References**


