Discussion Paper No.321

AN AFFIRMATIVE ACTION POLICY AS STRATEGY ON TWO-SIDED MATCHING MARKET

Taisuke Matsubae
Chuo University

Faculty of Economics
Chuo University

October 2019

INSTITUTE OF ECONOMIC RESEARCH
Chuo University
Tokyo, Japan

AN AFFIRMATIVE ACTION POLICY AS STRATEGY ON TWO-SIDED MATCHING MARKET

TAISUKE MATSUBAE

Faculty of Economics, Chuo University
Email: matsubae@tamacc.chuo-u.ac.jp

ABSTRACT. This paper investigates a market situation in which firms freely can choose whether they implement an affirmative action policy or not. We show that affirmative action policies can be abused by firms. Specifically, we demonstrate that there are market situations where a firm's implementation of affirmative action policies can employ more desirable workers, but is detrimental to the minority workers in the firm.
JEL *Classification number*: C70, D61.
Keywords: Matching, Stability, Affirmative Action, Stable Mechanism, Abuse-Proof.

1. INTRODUCTION

Matching markets arise in many economic environments. Matching markets are two-sided of the market, called men and women, workers and firms, students and colleges, and interns and hospitals and one sides want to be matched other sides. Many researchers have begun to investigate the theory of a two-sided matching market because of its theoretical appeal and the relevance of its design to real-world institutions. This paper investigates a environment in which firms can strategically choose an affirmative action policy in the context of a two sided matching market.

The two sided matchign market (also called by many to one matching problem) is pioneered by Gale and Shapley (1962). The worker-proposing deferred acceptance algorithm proposed by them has many appealing properties. A matching induced by the worker-proposing deferred acceptance algorithm is stable and Pareto dominates any other stable matching. Moreover, the woker-proposing deferred acceptance algorithm has a desirable property with regard to wokers' strategically action. Thus, it makes the truthful reporting of preferences a dominant strategy for all wokers. [1]

On the other hand, there are many studies focused on firms' strategically action. Any stable mechanism (including, in this case, the worker-proposing deferred acceptance mechanism) does not make the truthful reporting of preferences a dominant strategy for every firms. Thus, a stable mechanism can be manipulated via preference in the case where a firm will gain in terms of true preference by submitting a false preference instead of its true preference. Roth (1982) showed that any stable mechanism can be manipulable via preferences for firms. Firm's capacities are also private information; a mechanism can be manipulable in terms of capacities if a firm can gain in terms of true preference by underreporting its true capacities. Ehlers (2010) shows that manipulation via capacities can be equivalently described by two types of manipulation via capacities. He shows that no mechanism exists that is stable and excludes the possibility of non-Type II-manipulation via capacities.[2] [3] He also shows if mechanism is manipulable via capacities, then the mechanism is manipulable via preferences.

In many cases (i.e. college admission or job search), affirmative action policies have been playing an important role in the pursuit of racial desegregation. Affirmative action is the attempt to promote equal opportunities for all groups of society. It is often instituted in government and educational settings to ensure

I am thank Fuhito Kojima for discussion in early stages of this research.

[1]See Roth and Sotomayor (1990) for more detailed discussion.

[2]Non-Type II-manipulation means that firms with no vacant positions are unable to manipulate the situation by dropping some of its filled positions.

[3]Sönmez (1997) showed that when there are at least three workers and two firms, there exists no mechanism that is stable and non-manipulable via capacities.

that minority groups within a society are included in all programs. There are many studies that have examined affirmative action. (e.g. Abdulkadiroğlu (2005) and Kojima (2012)) Abdulkadiroğlu (2005) showed that, in the college preference domain, including in the context of the college admission problem, the student (worker)-proposing deferred acceptance mechanism makes the truthful revelation of preferences a dominant strategy for every student.

In Kojima (2012)'s investigation of the welfare effects of affirmative action policies in school choice, he found that there are certain market situations in which an affirmative action policy inevitably hurt every minority students. [4]

In this paper, we conduct an investigation, against this background, to examine whether firms can adapt their affirmative action policies to promote equal opportunity for minority workers. To examine this, we assess whether firms implement affirmative action policies to employ more desirable workers or not. Thus, we consider the case in which a firm can strategically implement affirmative action policies. We define that a firm abuses its affirmative action policies when the implementation of the affirmative action policies benefits the firm but is to the detriment of the minority workers in the firm. We define a mechanism as abuse-proof if the firm cannot benefit from the policy and minority workers are not put to a disadvantage, can benefit from the policy and minority workerss are not put a disadvantage, or can benefit from the policy and minority workers are put an advantage. The analytical approach of this paper follows in the tradition of impossibility studies from the existing literature. For example, see Roth (1982), Sönmez (1997), Sönmez (1999), and Kojima (2012).

The paper is organized as follows: Section 2 introduces the two-sided matching market-stability. Section 3 describes various affirmative action policies and defines what we mean by misuse-proof; we also present our results in terms of the stable mechanism examined. In Section 4, we provide concluding remarks.

2. Model

We consider the following college admission problem with type-specific capacities. We focus on the simple situation in which there are only two types of majority students and minority students. [5]

College admission problem with type-specific capacities, or simply a **problem** consists of the following terms:

(1) a finite set of students $S = \{s_1, s_2, \dots, s_n\}$,
(2) a finite set of colleges $C = \{c_1, c_2, \dots, c_m\}$,
(3) a students' preference profile $R_S = (R_{s_1}, R_{s_2}, \dots, R_{s_n})$, where R_s is a preference relation over C and being unmatched (being unmatched is denoted by $\varnothing$). We assume that preferences are strict. We write $c'P_s c$ if and only if $c'R_s c$ but not $cR_s c'$. If $c \succ_s \varnothing$, then c is said to be **acceptable** to s. The set of students are partitioned to two subsets; the set S^M of **majority students** and S^m of **minority students**;
(4) a colleges' preference profile $\succeq_C = (\succeq_{c_1}, \succeq_{c_2}, \dots, \succeq_{c_m})$, where $\succeq_c$ is the preference relation of college $c \in C$ over S. We assume that the preference relations are strict and **responsive with quota**. [6] [7] We write $s' \succ_c s$ if and only if $s' \succeq_c s$ but not $s \succ_c s'$. If $S' \succ_c \varnothing$, then set of students $S' \subset S$ is said **acceptable** for c ;
(5) for each $c \in C$, $\mathbf{q}_c = (q_c, q_c^M)$ is the **capacity** of c: The first component q_c represents the total capacity of college c, while the second component q_c^M represents the type-specific capacity for the majority students.

Let a problem be $G = (S, C, (R_s)_{s\in S}, (\succeq)_{c\in C}, (\mathbf{q}_c)_{c\in C})$.

[4] Hafalir, Bumin, Muhammed and Yildirim (2013) considered welfare effect of minority by anothe affimative action poicy. They showed that there is minority student who admits to better college by their affimative action policy.

[5] Note that our results do not depend on the assumption that there are more majority students than minority students.

[6] The preferences $\succeq_c$ of college c are responsive with quata q_c if (a) for all $s, s' \in S$, if $\{s\} \succ_c \{s'\}$, then for any $S' \subseteq S \setminus \{s, s'\}$, $S' \cup \{s\} \succ_c S' \cup \{s'\}$, (b) for all $s, s' \in S$ if $\{s\} \succ_c \varnothing$, then for any $S' \subseteq S$ such that $|S'| < q_c, S' \cup \{s\} \succ_c S'$, and (c) $\varnothing \succ_c S'$ for any $S' \subseteq S$ with $|S'| > q_c$.

[7] Note that while our results are showed under responsive preference, showing an impossibility with respect to the requirement immediately implies an impossibility result for a wider requirement.

We define the notation of **matching** introduced by Kojima (2012). A matching μ is a function from the set $S \cup C$ to the set of all subsets of $S \cup C$ such that

(1) $|\mu(s)| = 1$ for every students s, and $\mu(s) = \varnothing$ if $s \notin \mu(c)$ for all $c \in C$; [8] [9]
(2) For all $s \in S$ and $c \in C$, $\mu(s) = c$ if and only if $s \in \mu(c)$;
(3) $|\mu(c)| \leq q_c$ and $\mu(c) \subseteq S$ for any college c;
(4) $|\mu(c) \cap S^M| \leq q_c^M$ for all $c \in C$.

This matching requires (4) in addition to the standard requirements (1)-(3). (4) means that the number of majority students matched to each college c represents its type-specific capacity q_c^M or fewer. Let $\mathscr{M}$ be the set of matching.

We use the definition of stable matching introduced by Kojima (2012). A matching μ is **stable** if

(1) $\mu(s) R_s \varnothing$ for all $s \in S$, and for all $c \in C$ and all $s \in \mu(c)$, $s \succeq \varnothing$, and
(2) if there exists a pair $(s,c) \in S \times C$ such that $c P_s \mu^G(s)$, then we have either
  (a) $|\mu(c)| = q_c$ and $s' \succ_c s$ for all $s' \in \mu(c)$, or
  (b) $s \in S^M$, $|\mu(c) \cap S^M| = q_c^M$, and $s' \succ_c s$ for all $s' \in (\mu(c) \cap S^M)$.

This definition is standard except for (2b), but later part of (1) is different from Kojima (2012) . Condition (2b) means that when student s is a majority student, even if she prefers college c to the college matched to her and the slots of the college are left, if the slots of c for majority students are filled by students who have higher preference than her, then she cannot block the matching.

A **mechanism** is a function ϕ that, for each problem G, associates a matching $\mathscr{M}$. Note that $\phi(G) \in \mathscr{M}$ for any problem G. A mechanism ϕ is stable if $\phi(G)$ is a stable matching for any given G. We denote college to which student s match at a problem G as $\phi_s(G)$, and the set of students to which college c is matched at a problem G as $\phi_c(G)$.

We denote the stable mechanism as $\phi^{\text{stable}}(\cdot)$. [10]

3. Abuse-proof for affirmative action policies

We consider a market situation in which a college can freely implements an affirmative action policy. However, we assume that any college cannot change quota or preference except implementing an affirmative action policies. For example, by law or rule, the preferences orderings of college is decided.

Let $G = (S, C, (R_s)_{s\in S}, (\succeq)_{c\in C}, (\mathbf{q}_c, \mathbf{q}_{-c}))$ be a problem. We say that another problem $\tilde{G} = (S, C, (R_s)_{s\in S}, (\succeq)_{c\in C}, (\tilde{\mathbf{q}}_c, \mathbf{q}_{-c}))$ is obtained via college c's implementation of an affirmative action policy if $q_c = \tilde{q}_c$, $q_c^M > \tilde{q}_c^M$ and for all $c' \in (C \setminus \{c\})$, $q_{c'} = \tilde{q}_{c'}$ and $q_{c'}^M = \tilde{q}_{c'}^M$.

Definition 1. Given a mechanism ϕ, a college c **abuses the affirmative action policy** in problem G if there exist a problem $\tilde{G}$ and minority student $s \in (\phi_c(\tilde{G}) \setminus S^M)$ such that $\phi_c(\tilde{G}) \succ_c \phi_c(G)$, $\phi_s(G) P_s c$, and $\tilde{G}$ is a problem in which the college c implements affirmative action policy.

This defines a market situation in which, when a college implements an affirmative action policy, the decision benefits the college, but is to the detriment of the new minority students who is matched to the college in the problem $\tilde{G}$ and is not matched to the college in the problem G. Clearly, the conditions of this market situation are entirely contrary to those intended by the architects of affirmative action. Hence, we consider whether there might be a mechanism that can eliminate such a market situation.

In a situation where, for any problem, no colleges abuse the affirmative action policy under a mechanism, then we can say the mechanism is **abuse-proof for the affirmative action policy**. We hope that a stable mechanism is abuse-proof for an affirmative action policy. Unfortunately, the next result shows any stable mechanism is not abuse-proof for affirmative action policy.

Theorem 1. *There exists no stable mechanism which is abuse-proof for the affirmative action policy.*

[8] $\mu(c) = \{s \in S \mid \mu(s) = c\}$

[9] Because each student is matched to exactly one school or no school, we will omit set brackets and write $\mu(s) = c$ instead of $\mu(s) = \{c\}$ and $\mu(s) = s$ instead of $\mu(s) = \{s\}$.

[10] We call a mechanism a stable mechanism if for any problem, the mechanism induce stable matching.

Proof. The proof is provided via a counter example. Hence, we consider a problem

$$G = (S, C, (R_s)_{s\in S}, (\succeq)_{c\in C}, (\mathbf{q}_c)_{c\in C}).$$

The problem G can be expressed as follows: Let $S = \{s_1, s_2, s_3, s_4\}$, $C = \{c_1, c_2, c_3\}$, $S^M = \{s_1, s_2, s_4\}$, $S^m = \{s_3\}$. Student preferences are given by

$$\begin{aligned} R_{s_1} &: c_1, c_2, c_3, \\ R_{s_2} &: c_1, c_2, c_3, \\ R_{s_3} &: c_2, c_1, c_3, \\ R_{s_4} &: c_1, c_2, c_3, \end{aligned}$$

where the notational convention is that colleges are listed in order of preferences and colleges not on the preference list are unacceptable. College preferences and capacities are given by

$$\begin{aligned} \succeq_{c_1} &: s_3, s_2, s_1, s_4 \qquad \mathbf{q}_{c_1} = (q_{c_1}, q^M_{c_1}) = (2,2), \\ \succeq_{c_2} &: s_1, s_2, s_3, s_4 \qquad \mathbf{q}_{c_2} = (q_{c_2}, q^M_{c_2}) = (1,1), \\ \succeq_{c_3} &: s_1, s_2, s_3, s_4 \qquad \mathbf{q}_{c_3} = (q_{c_3}, q^M_{c_3}) = (1,1), \end{aligned}$$

where the notational convention here is that students are listed in their order of preference: At college c_1, for instance, student s_3 has the highest preference, s_1 has the second highest preference, s_2 has the third highest preference, and s_4 has the lowest preference.

There exists a unique stable matching $\phi^{\text{stable}}(G)$ in this problem G , which is given by

$$\phi^{\text{stable}}(G) = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_1, s_2 & s_3 & s_4 \end{pmatrix},$$

which means that c_1 is matched to s_1 and s_2, c_2 is matched to s_3, and c_3 is matched to s_4.

Consider the case in which college c_1 implements the affirmative action policy, $(\tilde{\mathbf{q}}_c)_{c\in C} = (\tilde{\mathbf{q}}_{c_1}, \mathbf{q}_{c_2})$, where $\tilde{\mathbf{q}}_{c_1} = (2,1)$, $\mathbf{q}_{c_2} = (1,1)$, and $\mathbf{q}_{c_3} = (1,1)$. Thus, the problem

$$\tilde{G} = (S, C, (R_s)_{s\in S}, (\succeq_c)_{c\in C}, (\tilde{\mathbf{q}}_c)_{c\in C}),$$

is that college c_1 implements the affirmative action policy. In the problem $\tilde{G}$, there is a unique stable matching $\phi^{\text{stable}}(\tilde{G})$ given by

$$\phi^{\text{stable}}(\tilde{G}) = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_2, s_3 & s_1 & s_4 \end{pmatrix}.$$

The minority student s_3 in college c_1 is strictly worse off under $\tilde{G}$ than they are under G. Moreover, college c_1 is strictly better off under $\tilde{G}$ than it would be under G. Therefore, it is clear that the college c_1 has incentive to abuse the affirmative action policy. □

The result suggests that no stable mechanism can remove the incentive for a college to abuse the affirmative action policy. This result may sound that colleges always abuse affirmative action policy. However, this intuition is incorrect. We present a matching market such that stable mechanism can help minority students without allowing the college to abuse their affirmative action policy. To do show it, we borrow the example of Kojima (2012).

Example 1 (Example 1 of Kojima (2012))**.** Let $G = (S, C, (R_s)_{s\in S}, (\succeq)_{c\in C}, (\mathbf{q}_c)_{c\in C})$. The problem G is as follows. Let $S = \{s_1, s_2, s_3, s_4\}$, $C = \{c_1, c_2\}$, $S^M = \{s_1, s_2, s_3\}$, and $S^m = \{s_4\}$. Student preferences are given by

$$\begin{aligned} R_{s_1} &: c_1, \\ R_{s_2} &: c_1, \\ R_{s_3} &: c_1, c_2, \\ R_{s_4} &: c_2, c_1. \end{aligned}$$

College preferences and capacities are given by

$$\succeq_{c_1} \; : \; s_1, s_4, s_2, s_3 \qquad \mathbf{q}_{c_1} = (q_{c_1}, q^M_{c_1}) = (2,2),$$
$$\succeq_{c_2} \; : \; s_3, s_4, s_2, s_1 \qquad \mathbf{q}_{c_2} = (q_{c_2}, q^M_{c_2}) = (1,1).$$

The unique stable matching $\phi^{\text{stable}}(G)$ in this problem G is given by

$$\phi^{\text{stable}}(G) = \begin{pmatrix} c_1 & c_2 & \varnothing \\ s_1, s_4 & s_3 & s_2 \end{pmatrix}$$

Let us consider the case in which college c_1 implements an affirmative action policy, $(\tilde{\mathbf{q}}_c)_{c\in C} = (\tilde{\mathbf{q}}_{c_1}, \mathbf{q}_{c_2})$, where $\tilde{\mathbf{q}}_{c_1} = (2,1)$ and $\mathbf{q}_{c_2} = (1,1)$. The problem is

$$\tilde{G} = (S, C, (R_s)_{s\in S}, (\succeq_c)_{c\in C}, (\tilde{\mathbf{q}}_c)_{c\in C}).$$

In the problem $\tilde{G}$, there is a unique stable matching $\phi^{\text{stable}}(\tilde{G})$ which is given by

$$\phi^{\text{stable}}(\tilde{G}) = \begin{pmatrix} c_1 & c_2 & \varnothing \\ s_1, s_4 & s_3 & s_2 \end{pmatrix}.$$

The minority student in c_1, s_4 is not strictly worse off under $\tilde{G}$ than they are under G. In addition, c_1 is not better off under $\tilde{G}$ than it would be under G. Therefore, c_1 does not abuse the affirmative action policy.

Next, we consider the case in which college c_2 implements an affirmative action policy, $(\mathbf{q}'_c)_{c\in C} = (\mathbf{q}_{c_1}, \mathbf{q}'_{c_2})$, where $\mathbf{q}_{c_1} = (2,2)$ and $\mathbf{q}'_{c_2} = (1,0)$. The problem is

$$G' = (S, C, (R_s)_{s\in S}, (\succeq_c)_{c\in C}, (\mathbf{q}'_c)_{c\in C}).$$

In the problem G', there is a unique stable assignment $\phi^{\text{stable}}(G')$, which is given by

$$\phi^{\text{stable}}(G') = \begin{pmatrix} c_1 & c_2 & \varnothing \\ s_1, s_2 & s_4 & s_3 \end{pmatrix}.$$

The minority student in c_2, s_4 is strictly better off under G' than they are under G. In contrast, college c_2 is worse off under G' than it would be under G. Therefore, c_2 dose not abuse the affirmative action policy. No colleges abuse the affirmative action policy for any stable mechanism. Thus, this example shows that a stable mechanism is abuse-proof for an affirmative action policy.

So far, we have seen that colleges will abuse their affirmative action policy under any stable mechanism. Kojima (2012) defines an alternative affirmative action policy, which he calls **preference-based affirmative action policy**. This situation: $\tilde{G} = (S, C, (R_s)_{s\in S}, (\tilde{\succeq})_{c\in C}, (\mathbf{q}_c)_{c\in C})$ is said to have a stronger preference-based affirmative action policy than $G = (S, C, (R_s)_{s\in S}, (\succeq)_{c\in C}, (\mathbf{q}_c)_{c\in C})$ if, for every $c \in C$ and $s, s' \in S$, $s \succeq_c s'$ and $s \in S^m$ imply $s \tilde{\succeq}_c s'$.

We consider the situation in which a college can freely implement preference-based affirmative action policy under any stable mechanism in the following way: Let $\tilde{G}$ be $(S, C, (R_s)_{s\in S}, (\tilde{\succeq}_c, \succeq_{-c}), (\mathbf{q}_c)_{c\in C})$, where $\tilde{\succeq}_c$ means that college c's preference improve one minority student who is ranked lower than a majority student, while the relative ranking of each student within her own group remains fixed. [11] We define abuse for the preference-based affirmative action policy by a college as we did before.

Definition 2. A college c **abuses the preference-based affirmative action policy** in problem G if there exist a problem $\tilde{G}$ and $s \in (\phi^{\text{stable}}_c(\tilde{G}) \setminus S^M)$ such that $\phi^{\text{stable}}_c(\tilde{G}) \succ_c \phi^{\text{stable}}_c(G)$, $\phi^{\text{stable}}_s(G) P_s c$, and $\tilde{G}$ is a problem in which the college c implements preference-based affirmative action policy.

If, for any problem, no colleges are found to abuse the preference-based affirmative action policy under a certain mechanism, then we say the mechanism is **abuse-proof for the preference-based affirmative action policy**. Although the current definition considers whether there is a strategy-proof for college preferences, our requirements are different from those of the mechanism because our conditions also include the minority student's welfare in the college. If a mechanism holds this definition, then the mechanism naturally holds the strategy-proof for college preferences, but the inverse dose not hold.

[11] $\succeq_{-c} = (\succeq_{c'})_{c'\in C\setminus\{c'\}}$.

Theorem 2. *There exists no stable mechanism which is abuse-proof for the preference-based affirmative action policy.*

Proof. We will prove this theorem with the example.

Let $G = (S, C, (R_s)_{s\in S}, (\succeq)_{c\in C}, (\mathbf{q}_c)_{c\in C})$. The problem G is is expressed thus: Let $S = \{s_1, s_2, s_3, s_4, s_5\}$, $C = \{c_1, c_2, c_3\}$, $S^M = \{s_1, s_3\}$, $S^m = \{s_2, s_4, s_5\}$. Student preferences are given by

$$\begin{aligned} R_{s_1} &: c_3, c_2, c_1, \\ R_{s_2} &: c_2, c_1, c_3, \\ R_{s_3} &: c_1, c_3, c_2, \\ R_{s_4} &: c_1, c_2, c_3, \\ R_{s_5} &: c_1. \end{aligned}$$

College preferences are as follows.

$$\begin{aligned} \succeq_{c_1} &: s_1, s_2, s_3, s_4, s_5, && \mathbf{q}_{c_1} = (q_{c_1}, q^M_{c_1}) = (2,2), \\ \succeq_{c_2} &: s_1, s_2, s_3, s_4, s_5, && \mathbf{q}_{c_2} = (q_{c_2}, q^M_{c_2}) = (1,1), \\ \succeq_{c_3} &: s_3, s_1, s_2, s_4, s_5, && \mathbf{q}_{c_3} = (q_{c_3}, q^M_{c_3}) = (1,1). \end{aligned}$$

In this problem, there is a unique stable matching $\phi^{\text{stable}}(G)$ given by

$$\phi^{\text{stable}}(G) = \begin{pmatrix} c_1 & c_2 & c_3 & \varnothing \\ s_3, s_4 & s_2 & s_1 & s_5 \end{pmatrix}.$$

In the case, college c_1 implements the preference-based affirmative action policy as follows:

$$\succeq'_{c_1} \;:\; s_1, s_2, s_4, s_5, s_3 \qquad \mathbf{q}_{c_1} = (q_{c_1}, q^M_{c_1}) = (2,2).$$

In problem $\tilde{G}$, a unique stable matching $\phi^{\text{stable}}(\tilde{G})$ is given by

$$\phi^{\text{stable}}(\tilde{G}) = \begin{pmatrix} c_1 & c_2 & c_3 & \varnothing \\ s_1, s_4 & s_2 & s_3 & s_5 \end{pmatrix}.$$

College c_1 is better off under $\tilde{G}$ it would be than under G, but the minority student s_2 is worse off under $\tilde{G}$ they would be than under G. Thus, $\phi^{\text{stable}}(\tilde{G})$ is found not to be an abuse-proof for the preference-based affirmative action policy. □

Note that this result dose not imply that a college has incentive to abuse the preference-based affirmative action policy under stable mechanism. The following example shows that there exists a stable mechanism which is abuse-proof for the preference-based affirmative action policy.

Example 2. Let $G = (S, C, (R_s)_{s\in S}, (\succeq)_{c\in C}, \mathbf{q}_c)_{c\in C})$. The problem, G is as follows. Let $S = \{s_1, s_2, s_3\}$, $C = \{c_1, c_2\}$, $S^M = \{s_1\}$, $S^m = \{s_2, s_3\}$. Student preferences are given by

$$\begin{aligned} R_{s_1} &: c_2, c_1, \\ R_{s_2} &: c_1, c_2, \\ R_{s_3} &: c_2, c_1, \end{aligned}$$

College preferences are as follows.

$$\begin{aligned} \succeq_{c_1} &: s_1, s_2, s_3 && \mathbf{q}_{c_1} = (q_{c_1}, q^M_{c_1}) = (1,1), \\ \succeq_{c_2} &: s_1, s_3, s_2 && \mathbf{q}_{c_2} = (q_{c_2}, q^M_{c_2}) = (1,1). \end{aligned}$$

In this problem, there is a unique stable matching $\phi^{\text{stable}}(G)$which is given by

$$\phi^{\text{stable}}(G) = \begin{pmatrix} c_1 & c_2 & \varnothing \\ s_2 & s_1 & s_3 \end{pmatrix}$$

The case in which a college c_1 implements the preference-based affirmative action policy is expressed as follows,

$$\succeq'_{c_1} \;:\; s_2, s_1, s_3 \qquad \mathbf{q}_{c_1} = (q_{c_1}, q^M_{c_1}) = (1,1).$$

Specifically, it is only preference treatment in which the college c_1 implements the preference-based affirmative action policy.

In problem $\tilde{G}$, there is a unique stable matching $\phi^{\text{stable}}(\tilde{G})$ which is given by

$$\phi^{\text{stable}}(\tilde{G}) = \begin{pmatrix} c_1 & c_2 & \varnothing \\ s_2 & s_1 & s_3. \end{pmatrix}$$

The minority student in college c_1 is not found to be strictly better off under $\tilde{G}$ than they would be under G.

Next, we consider the case in which college c_2 implements the preference-based affirmative action policy as follows,

$$\succeq'_{c_2}: s_3, s_1, s_2 \qquad \mathbf{q}_{c_2} = (q_{c_2}, q^M_{c_2}) = (1,1).$$

Note that this is the only preference treatment in which c_2 implements the preference-based affirmative action policy.

In problem $\tilde{G}'$, there is a unique stable matching $\phi^{\text{stable}}(\tilde{G}')$ which is given by

$$\phi^{\text{stable}}(\tilde{G}') = \begin{pmatrix} c_1 & c_2 & \varnothing \\ s_2 & s_3 & s_1 \end{pmatrix}$$

The minority student s_3 in college c_2 is found to be strictly better off under $\phi^{\text{stable}}(\tilde{G}')$ than they would be under $\phi^{\text{stable}}(G)$.

The minority student in a college is never worse off when their college implements the preference-based affirmative action policy, regardless of the college. Thus, in this problem, the stable mechanism is found to be abuse-proof for the preference-based affirmative action policy.

4. Conclusion

This paper investigated whether affirmative action policies work to accomplish its intended purpose in the context of a college admission problem. We attempted to identify a stable mechanism that would be an abuse-proof for affirmative action policies. Unfortunately, our results suggest that this does not exist. This implies that college affirmative action policies do not necessarily promote equal opportunities for minority students. This finding correlated with the findings of Kojima (2012), who also found that affirmative action policies were likely to be harmful to minority students. He identified certain environments in which the implementation of an affirmative action policy was inevitably to the detriment of every minority student under any stable mechanism. Our results imply that this may be an outcome of colleges' abuse of affirmative action policies.

We find that a market situation exists in which, once an affirmative action policy is implemented, a college benefits but the minority students in the school are worse off. More specifically, we establish that this applies under any stable mechanism. Moreover, this misfortune is found to be unavoidable regardless of whether alternative affirmative action policies are implemented; for instance, this still applies in situations where minority students are given preferential treatment. Indeed, we establish that, in situations where colleges give minority students preferential treatment, the college benefits, while minority students in the college are worse off; this applies under any stable mechanism.

As both our results and those of Kojima (2012) suggest, caution should be exercised when employing affirmative action policies. While affirmative action may be an attempt to promote equal opportunity for the minority, clearly, the policy can be abused by a college.

References

Abdulkadiroğlu, A. (2005) "College admissions with affirmative action" *International Journal of Game Theory* 33(4), 535–549.

Abdulkadiroğlu, A. and T. Sönmez (2003) "School choice: A mechanism design approach" *American Economic Review* 98(3) 729–747.

Ehlers, L.(2010)"Manipulation via capacities revisited" *Games and Economic Behavior* 69(2), 302–311.

Gale, D. and L. S. Shapley (1962) “College admissions and the stability of marriage” *The American Mathematical Monthly* 69 (1), 9–15.

Hafalir Isa E. and M. Bumin Yenmez and Muhammed and A. Yildirim (2013) “Effective affirmative action in school choice” *Theoretical Economics* 8 325–363

Kojima, F. (2012) “School Choice: Impossibility For Affirmative Action” *Games and Economic Behavior* 75(2) 685–693.

Roth, A. E. (1982) “The Economics of Matching: Stability and Incentives” *Mathematics of Operations Research*, 7, 617–628.

Roth, A. E. and M. A. O. Sotomayor(1990) Two-sided matching: a study in game-theoretic modeling and analysis Econometric Society monographs, Cambridge.

Roth, A. E. (2008) “Deferred acceptance algorithms: history, theory, practice, and open questions” *International Journal of Game Theory*, 6(3), 537–569.

Sönmez, T. (1997) “Manipulation via Capacities in Two-Sided Matching Markets” *Journal of Economic Theory*, 77, 197–204.

Sönmez, T. (1999) “Can Pre-arranged Matches be Avoided in Two-Sided Matching Markets?” *Journal of Economic Theory*, 86, 148–156.

Sönmez, T. and U. Ünver (2009) “Matching, Allocation, and Exchange of Discrete Resources”, Handbook of Social Economics, eds. Jess Benhabib, Alberto Bisin, and Matthew Jackson, Elsevier.