Misuse-proof for Affirmative Action in College Admission Problem

Taisuke Matsubae

1. Introduction

This paper investigates a college admission problem with affirmative action policy. A college ad- mission problem is a many to one, two-sided matching problem. The theory of two-sided matching market has interested researchers for its theoretical appeal and its relevance to the design of real-world institutions. College admission problem is pioneered by Gale and Shapley (1962). There is Sönmez and Unver (2009) as a survey in this field. The deferred acceptance algorithm proposed by them has many appealing properties. The student- proposing deferred acceptance allocation Pareto dominates any other stable matching. Moreover, the student-proposing deferred acceptance rule makes truthful reporting of preferences a dominant strategy for every student. (see Roth and Sotomayor (1990), Roth (2008)) However, any stable mechanism (i.e. the student-proposing deferred acceptance mechanism) does not make truthful reporting of preferences a dominant strategy for every colleges. Thus, a stable mechanism is manipulable via preference if a college can gain in terms of true preference by submitting a false preference instead of its true preference. Roth (1982) showed that any stable mechanism is manipulable via preferences for colleges. Colleges' capacities also are private information and mechanism is manipulable via capacities if a college can gain in terms of true preference by underreporting its true capacities. For example, see Sönmez (1997). Ehlers (2010) showed that manipulation via capacities can be equivalently described by two types of manipulation via capacities. He shows that there exists no mechanism that is stable and non-Type II-manipulation via capacities. In college admission, affirmative action policies have been playing an important role in achieving racial desegregation. Affirmative action is an attempt to promote equal opportunity. It is often instituted in government and educational settings to ensure that minority groups within a society are included in all programs. There are many studies regarding affirmative action. For example, Abdulkadiroglu (2003), Abdulkadiroglu (2005) and Kojima (2012), Hafalir et al. (2013) etc. Abdulkadiroglu (2005) showed colleges' preference domain such that in the context of the college admission problem, student-proposing deferred acceptance mechanism makes truthful revelation of preferences a dominant strategy for every student. Kojima (2012) investigated the welfare effects of affirmative action policies in school choice. He showed that there are market situations in which affirmative action policy inevitably hurt every minority student.

We ask: Does a college adapt affirmative action policies to promote equal opportunity for the minority students? This paper study whether a college can misuse affirmative action policies. Thus, we consider the case in which a college strategically can implement affirmative action policies. We say that a college misuse affirmative action policies if when the college implements affirmative action policies, the college can gain but the minority students in the college is worse off. We refer a mechanism as misuse-proof if the college cannot gain or the minority students in the college are not worse off. Our main finding is that there exists the market situation in which a college can gain but the minority students in the school are worse off. More specially, we establish impossibility theorems stating that there are situations where when a college implements affirmative action policies, the college can gain but the minority students in the college is worse off under any stable mechanism. Moreover, this mis-fortune is unavoidable under alternative ways to implement the affirmative action policy: a college gives the minority students preference treatment. Thus, we also establish impossibility theorems stating that there are situations where when a college gives the minority students preference treatment., the college can gain but the minority students in the college is worse off under any stable mechanism. However, we also demonstrate that implementing affirmative action policies, the minority students in the college is strictly better off.

The analytical approach of this paper follows the tradition of impossibility studies in the matching 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 defines affirmative action policies and misuse-proof, and show our results for stable mechanism. Section 4 concludes.

2. A Model

We consider the following college admission problem. We focus on the simple situation in which there are only two types of students' majority and minority. *College admission problem* or simply *problem* consists of the following:

1. a finite set of students: $S = \{s_1, s_2, \ldots, s_n\}$, $n \in (\mathbb{N}$ (the set of natural numbers)$\setminus \{0\})$, where let s_i be a represented student, and the set of students are partitioned to two subsets; the set S^M of *majority students* and S^m of *minority students*,
2. a finite set of colleges: $C = \{c_1, c_2, \ldots, c_m\}$, $m \in (\mathbb{N} \setminus \{0\})$, where let c_k be represented college,
3. a students' preference profile$R_S = (R_{s_1}, R_{s_2}, \cdots, R_{s_n})$, where R_{s_i} is a preference relation over C and being unmatched (being unmatched is denoted by $\emptyset$). We assume that preferences are strict. We write $cP_{s_i}c'$ if and only if $cR_{s_i}c'$ but not $cR_{s_i}c'$. If $cP_{s_i}\emptyset$, then c is said to be *acceptable* to s_i,
4. a colleges' preference profile $\succeq_C = (\succeq_{c_1}, \succeq_{c_2}, \ldots, \succeq_{c_m})$, where $\succeq_{c_j}$ is the preference relation of college $c_j \in C$ over $S \cup c_j$. We assume that the preference relations are strict and responsive[1]. We write $s' \succ_c s$ if and only if $s' \succeq_c s$ but not $s \succeq_c s'$,
5. for each $c \in C$, $q_c = (q_c, q_c^M)$ 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 majority students.

We define the notation of matching introduced by Kojima (2012).

An matching μ is a function from the set $S \cup C$ to the set of all subsets of $S \cup C$ such that

M.1. $|\mu(s)| = 1$ for every student s, and $\mu(s) = s$ if $s \notin \mu(c)$. By this definition, 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 $\mu(s) = \{s\}$;

M.2. For all $s \in S$ and $c \in C$, $\mu(s) = c$ if and only if $s \in \mu(c)$;

M.3. $\mu(c)| \leq q_c$ and $\mu(c) \subseteq S$ for any college c;

M.4. $|\mu(c) \cap S^M| \leq q_c^M$ for all $c \in C$.

This matching requires M.4 in addition to standard requirements M.1-M.3. M.4 means that the number of majority students matched to each school c is its type-specific capacity q_c^M and fewer.

Moreover, we also use the definition of stable introduced by Kojima (2012).

A matching μ is sable if

S.1 $\mu(s)\ R_s\ \emptyset$ for all student $s \in S$, and

S.2 if $cP_s\ \mu(s)$, then either

S.2a $|\mu(c)| = q_c$ and $s' \succ_c s$ for all $s' \in \mu(c)$, or

S.2b $s \in S^M$, $|\mu(c) \cap S^M| = q_c^M$, and $s' \succ_c s$ for all $s' \in \mu(c) \cap S^M$.

---

1) $\succeq_c$ is responsive on students if for all subset $S' \subset S$ and all student $s, s' \in S \setminus S'$, (i) $S' \cup \{s\} \succ_c S' \cup \{s'\} \Leftrightarrow \{s\} \succ_c \{s'\}$, and (ii) $S' \cup \{s\} \succ_c S' \Leftrightarrow \{s\} \succ_c \emptyset$.

This definition is standard except for S.2b. The condition S.2b means that when student s is a majority student, even if she prefers college c to the college matched to her and slot of the college are left, if the slots of c for majority students are filled by students who have higher priority than her, then she cannot block the matching.

A *mechanism* is a function ϕ that, for each problem G, associates a matching $\phi(G)$. A mechanism ϕ is stable if $\phi(G)$ is a stable matching for any given G.

3. Misuse-proof for affirmative action policy

We consider a market situation in which a college implements affirmative action policy. Let us introduce the affirmative action policy. We define that a problem $\widetilde{G} = (S, C, (R_s)_{s\in S}, (\succeq)_{c\in C}, (\widetilde{q_c})_{c\in C})$ is said to *for a college c implement affirmative action policy* if $q_c = \widetilde{q_c}$ and $= q_c^M > \widetilde{q_c^M}$, and other colleges are $q_{-c} = \widetilde{q_{-c}}$ and $q_{-c}^M = \widetilde{q_{-c}^M}$.

Definition 1.

A college *misuses the affirmative action policy* if a college c prefers $\mu_c(\widetilde{G})$ *to* $\mu_c(G)$ but the minority student s in c prefers $\mu_s(G)$ *to* $\mu_s(\widetilde{G})$.

This definition considers a market situation in which a college implements the affirmative action policy, the college is made better off but the minority student in the college is made worse off. However, such market situation is not affirmative action's original goal. Hence, we consider a mechanism which eliminates such market situation.

If for any problem all colleges do not misuse the affirmative action policy under a mechanism, then we say the mechanism is *misuse-proof for the affirmative action policy*. We hope that a stable mechanism such that the deferred acceptance mechanism is misuse-proof for affirmative action policy. Unfortunately, the next result shows any stable mechanism is not misuse-proof for affirmative action policy.

Theorem 1.

There exists no stable mechanism which is misuse-proof for the affirmative action policy.

Proof

The proof is via a counter example.

Hence, we consider a problem $G = (S, C, (R_s)_{s\in S}, (\succeq)_{c\in C}, (q_c)_{c\in C})$.

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

$R_{s_1}: c_1, c_2,$
$R_{s_2}: c_1, c_2,$
$R_{s_3}: c_2, c_1,$
$R_{s_4}: c_1, c_2,$

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

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

where the notational convention here is that students are listed in order of preferences: At college c_1, for instance, student s_1 has the highest priority, s_5 has the second highest priority, s_3 has the third highest priority, s_4 has the fourth highest priority, and s_2 has the lowest priority.

There exists a unique stable matching μ in this problem given by

$$\mu(G) = \begin{pmatrix} c_1 & c_2 & \emptyset \\ s_1, s_2 & s_3 & s_4 \end{pmatrix}$$

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

Consider the case in which college c_1 implements the affirmative action policy, $(\widetilde{q_c})_{c\in C} = (\widetilde{q_{c_1}}, q_{c_2})$, where $\widetilde{q_{c_1}} = (2, 1)$ and $q_{c_2} = (1, 1)$. Thus, the problem $\widetilde{G} = (S, C, (R_s)_{s\in S}, (\succeq_c)_{c\in C}, (\widetilde{q_c})_{c\in C})$ is that college c_1 implements the affirmative action policy. In the problem $\widetilde{G}$, there is a unique stable matching $\mu(\widetilde{G})$ given by

$$\mu(\widetilde{G}) = \begin{pmatrix} c_1 & c_2 & \emptyset \\ s_1, s_3 & s_2 & s_4 \end{pmatrix}$$

The minority student s_3 in college c_1 is strictly worse off under $\widetilde{G}$ than under G. Also, college c_1 is strictly better off under $\widetilde{G}$ than under G. Therefore, the stable matching is misused by c_1.

Q.E.D.

The result implies that a college has an incentive to misuse the affirmative action policy under any stable mechanism. However, we show another example such that stable mechanism may help the minority students without college misusing.

Example 1.

Let $G = (S, C, (R_s)_{s\in S}, (\succeq)_{c\in C}, (q_c)_{c\in C})$. The problem G is as follows. Let $S = \{s_1, s_2, s_3, s_4\}$, $C = \{c_1, c_2\}$, $S^M = \{s_1, s_2, s_3\}$, $S^m = \{s_4\}$. Student and college preferences are given by

$R_{s_1}: c_1,$
$R_{s_2}: c_1,$
$R_{s_3}: c_1, c_2,$

$R_{s_4} : c_2, c_1,$

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

There exists a unique stable assignment μ in this problem G given by

$$\mu(G) = \begin{pmatrix} c_1 & c_2 & \emptyset \\ s_1, s_4 & s_3 & s_2 \end{pmatrix}$$

Consider the case in which college c_1 implements the affirmative action policy, $(\widetilde{q_c})_{c \in C} = (\widetilde{q_{c_1}}, q_{c_2})$, where $\widetilde{q_{c_1}} = (2, 1)$ and $q_{c_2} = (1, 1)$. The problem $\widetilde{G} = (S, C, (R_s)_{s \in S}, (\succeq_c)_{c \in C}, (\widetilde{q_c})_{c \in C})$ as the stronger affirmative action policy than $G = (S, C, (R_s)_{s \in S}, (\succeq_c)_{c \in C}, (q_c)_{c \in C})$. In the problem $\widetilde{G}$, there is a unique stable matching $\mu(\widetilde{G})$ given by

$$\mu(\widetilde{G}) = \begin{pmatrix} c_1 & c_2 & \emptyset \\ s_1, s_4 & s_3 & s_2 \end{pmatrix}$$

The minority student in c_1, s_4 is not strictly worse off under $\widetilde{G}$ than under G. Also, c_1 is not better off under $\widetilde{G}$ than under G. Therefore c_1 does not misuse the affirmative action policy.

Next, consider the case in which college c_2 implements the affirmative action policy, $\widetilde{q} = (q_{c_1}, \widetilde{q_{c_2}})$, where $q_{c_1} = (2, 2)$ and $\widetilde{q_{c_2}} = (1, 0)$. In the problem G″, there is a unique stable assignment $\mu(G'')$ given by

$$\mu(G'') = \begin{pmatrix} c_1 & c_2 & \emptyset \\ s_1, s_2 & s_4 & s_3 \end{pmatrix}$$

The minority student in c_2, s_4 is strictly better off under G'' than under G. On the other hand, college c_2 is worse off under G'' *t*han under G. Therefore c_2 dose not misuse the affirmative action policy. All colleges do not misuse the affirmative action policy for any stable mechanism. Thus, this example is stable mechanism which is misuse-proof for the affirmative action policy.

So far, we have seen that there exist a college which misuses the affirmative action policy under any stable mechanism. Kojima (2012) defines an alternative affirmative action policy. We call it preference-based affirmative action policy: $\widetilde{G} = (S, C, (R_s)_{s \in S}, (\widetilde{\succcurlyeq})_{c \in C}, (q_c)_{c \in C})$ is said to have a stronger preference-based affirmative action policy than $G = (S, C, (R_s)_{s \in S}, (\succcurlyeq)_{c \in C}, (q_c)_{c \in C})$ if, for every $c \in C$ and $s, s' \in S$, $s \succcurlyeq_c s'$ and $s \in S^m$ imply $s \widetilde{\succcurlyeq}_c s'$.

We consider whether a college implements the preference-based affirmative action policy under any stable mechanism. Let $\widetilde{G}$ be $(S, C, (R_s)_{s \in S}, (\widetilde{\succcurlyeq}_c, \succcurlyeq_{-c}), (q_c)_{c \in C})$, where $\widetilde{\succcurlyeq}_c$ means that college c′s priority improves one minority student which is lower ranked than a majority student, while keeping the relative ranking of each student within her own group fixed. Note that $\succcurlyeq_{-c} = (\succcurlyeq_{c'})_{c' \in C \setminus \{c\}}$. We as before define misuse for the preference-based affirmative action policy by college.

If for any problem all colleges do not misuse the affirmative action policy under a mechanism, then we say the mechanism is *misuse-proof for preference-based the affirmative action policy*. The current definition similarly considers the strategy-proof for college preferences. However, our requirement is different from it because our condition is also imposed on what happen to a minority student's welfare in the college. However, the following result may be trivial from the lattice structure of stable matching.

Theorem 2.

There exists no stable mechanism which is misuse for the priority-based affirmative action policy.

Proof

Let $G=(S, C, (R_s)_{s\in S}, (\succcurlyeq)_{c\in C}, (q_c)_{c\in C})$. 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

$R_{s_1}: c_2\ c_1,$
$R_{s_2}: c_1\ c_2,$
$R_{s_3}: c_2, c_1,$

College preferences are as follows.

$\succeq_{c_1}: s_1\ s_2\ s_3 \qquad q_{c_1}=(q_{c_1}, q^M_{c_1})=(1, 1),$
$\succeq_{c_2}: s_2\ s_1\ c_3 \qquad q_{c_2}=(q_{c_2}, q^M_{c_2})=(1, 1).$

In this problem, there are two stable matchings μ and μ' given by

$$\mu(G)=\begin{pmatrix} c_1 & c_2 & \emptyset \\ s_2 & s_1 & s_3 \end{pmatrix}$$

and

$$\mu'(G)=\begin{pmatrix} c_1 & c_2 & \emptyset \\ s_1 & s_2 & s_3 \end{pmatrix}$$

We consider the following cases.

(1) Suppose that $\phi(G)=\mu$. In the case, consider college c_2 implements the preference-based affirmative action policy as follows:

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

In problem $\widetilde{G}$, a unique stable matching $\widetilde{\mu}$ given by

$$\widetilde{\mu}(\widetilde{G})=\begin{pmatrix} c_1 & c_2 & \emptyset \\ s_1 & s_2 & s_3 \end{pmatrix}$$

College c_2 is better off under $\widetilde{G}$ than under G, but the minority student s_2 is worse off under $\widetilde{G}$ than under G.

Thus, $\widetilde{\mu}(\widetilde{G})$ is not misuse-proof for the preference-based affirmative action policy.

(2) Suppose that $\phi(G) = \mu'$. In that case, consider $\widetilde{G}'$ which changes c_2's preference as follows:

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

In problem $\widetilde{G}'$, a unique stable matching $\widetilde{\mu}'$ given by

$$\widetilde{\mu}'(\widetilde{G}') = \begin{pmatrix} c_1 & c_2 & \emptyset \\ s_2 & s_1 & s_3 \end{pmatrix}$$

College c_2 implements the preference-based affirmative action policy, when changing $\succcurlyeq_{c_2}$ from $\widetilde{\succcurlyeq}'_{c_2}$. Then, college c_2 is better off under $\widetilde{G}'$ than under G, but the minority student s_2 is worse off under $\widetilde{G}'$ than under G.

Thus, $\widetilde{\mu}'$ is not misuse-proof for the preference-based affirmative action policy.

Q.E.D.

The following example show that there exists a stable mechanism which is misuse-proof for the preference-based affirmative action policy.

Example 2.

Let $G = (S, C, (R_s)_{s \in S}, (\succcurlyeq)_{c \in C}, (q_c)_{c \in C})$. 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

$$R_{s_1} : c_2\ c_1,$$
$$R_{s_2} : c_1\ c_2,$$
$$R_{s_3} : c_2,\ c_1,$$

College preferences is as follows.

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

In this problem, there is a unique stable matching μ given by

$$\mu(G) = \begin{pmatrix} c_1 & c_2 & \emptyset \\ s_2 & s_1 & s_3 \end{pmatrix}$$

We consider the case in which a college c_1 implements the preference-based affirmative action policy as follows,

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

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

In problem $\widetilde{G}$, there is a unique stable matching μ' given by

$$\mu'(\widetilde{G}) = \begin{pmatrix} c_1 & c_2 & \emptyset \\ s_2 & s_1 & s_3 \end{pmatrix}$$

The minority student in college c_1 does not strictly better off under $\widetilde{G}$ than under G.

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

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

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

In problem $\widetilde{G'}$, there is a unique stable matching μ'' given by

$$\mu''(\widetilde{G'}) = \begin{pmatrix} c_1 & c_2 & \emptyset \\ s_2 & s_3 & s_1 \end{pmatrix}$$

The minority s_3 in collage c_3 strictly better off under μ''than under μ.

Even if any collage implements the preference-based affirmative action policy, the minority student in the college is not worse off. Thus, in this problem stable mechanism is misuse-proof for the preference-based affirmative action policy.

4. Conclusions

This paper investigated whether college misuse affirmative action policies in the contest of college admission problem. Unfortunately, our main results say that there exists no stable mechanism which is misuse-proof for affirmative action policies.

This result implies that a college to achieve affirmative action policies does not necessarily promote equal opportunity for the minority students.

Kojima (2012) also showed the negative results with regard to affirmative action policies. As sating in our introduction, he demonstrates that there are environments in which affirmative action policy inevitably hurt every minority student under any stable mechanism. Our results may imply that his negative results are caused by a college misusing affirmative action policies.

The results of this paper, as 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, the policy can be misused by a college.

References

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

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

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

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

Hafalir, Isa E., M. Bumin Yenmez and Muhammed A. Yildirim (2013) “Effective affirmative action in school choice”, *Theoretical Economics*, 8(2), 325–363.

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

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

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

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

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

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

Sönmez, Tayfun and Utku Unver (2009) “Matching, Allocation, and Exchange of Discrete Re-sources”, *Handbook of Social Economics*, eds. Jess Benhabib, Alberto Bisin, and Matthew Jackson, Elsevier.

(Assistant Professor, Faculty of Economics, Chuo University, Ph. D. (Economics))