

# 論 文 要 旨

平成 28 年 3 月 4 日

専攻：グローバル経済	氏名：Yusuke Kasuya
論 文 題 名	Three Essays on Market Design and Its Application to Educational Economics

This thesis consists of three essays on market design with applications to educational economics. The first essay studies a new design of school choice mechanism. The implications of peer effects on school choice mechanism design are of important, yet underexplored educational concerns. For one thing, in most cases, analytical models have ruled out forms of peer effects in preferences by assuming that composition of classmates does not influence a student's evaluation of a match. For another thing, the standard welfare and fairness criteria do not incorporate a social concern for educational achievement that in turn is the product of peer effects among students. This essay addresses school bullying incidents as documented sources of significant negative peer effects among students, thereby dealing with the two concerns discussed above.

The four reasons that may legitimate my pursuit are; (i) there is considerable evidence suggesting that the experience of being a victim adversely affects academic achievement and school adjustment; (ii) a well-replicated finding in the longitudinal research is that the bully-victim relation is stable and persistent from elementary to middle school and even afterwards; (iii) the school choice program in Japan has been considered as a means to buffer victims against bullies; (iv) (anti-bullying) school choice may play a significant role in anti-bullying programs by crafting the mechanisms in which the assignment for a student becomes less preferred if she bullies other student.

Formally, I consider school choice that, based on the exogenously given set of “bullying-incidents”, requires that bullies and victims be assigned to different schools, and redefine the concepts of stability and efficiency

accordingly. I then show that variants of the Gale-Shapley (deferred-acceptance) mechanism and the Top Trading Cycles mechanism achieve these social goals respectively. Moreover, the mechanisms satisfy the desideratum that I described in the previous paragraph; i. e., they assign less preferred schools to bullies. Through this channel, the mechanisms can potentially help prevent bullying in, for example, elementary schools since they assign less preferred middle schools to bullies thereby punishing them.

The variant of the Gale-Shapley mechanism (called the two-round Gale-Shapley mechanism) is suited to “direct bullying” in which a victim confronts a bully face to face (examples cover physical aggression such as shoving, slapping, choking, and verbal abuse such as calling a derogatory nickname to make a victim feel mocked or humiliated) since it requires that the set of bullying-incidents satisfy the “regularity condition”. This condition ensures that victims never bully other students, which may be a good approximation of reality if only ongoing direct bullying incidents at the time a social planner decides on the assignment are considered. Yet, some forms of bullying clearly fall outside the scope set by direct bullying; examples include indirect bullying such as cyberbullying, bully/victims, and “ijime” (Japanese bullying). In this respect, the variant of the Top Trading Cycles mechanism dispenses with the regularity condition and hence it can deal with any form of bullying (that can be reasonably reduced to binary relations).

Finally, I discuss the complementary roles of anti-bullying programs and school choice. Anti-bullying programs can inform school choice through a channel of research on bullying since (i) longitudinal studies and other close examination of social relationships among students and teachers would tell us nuanced implications about the goodness and badness of student assignment which allows us to craft more complex mechanisms and go beyond further a simple utilitarian perspective; and (ii) the expertise of bullying researchers can clarify how and to what extent a social planner can extract a set of bullying incidents. On the other hand, school choice can inform anti-bullying programs since the desideratum on punishment which I described in the second paragraph is congruent with the premises of successful intervention, and thus the anti-bullying school choice mechanisms can help increase the effectiveness of a whole anti-bullying program in concert with other strategies such as restorative approaches, school council, and peer support.

The second essay studies stable and (one-sided) strategy-proof matching rules in many-to-one matching markets with contracts. Since its birth by Gale and Shapley (1962), the theory of two-sided matching markets has been centered around the deferred acceptance algorithm, which is known to satisfy a number of desiderata. Specifically, in the classic setup (including school choice), it is not only stable but also (one-sided) strategy-proof, and moreover, it is the unique such rule. More recently, Hatfield and Milgrom (2005) propose a more general model of matching with contracts and, among many other things, verify that the deferred acceptance is stable and strategy-proof under certain conditions, called substitutable contracts and the law of aggregated demand. With those conditions, Sakai (2011) also generalizes the uniqueness result and hence, the study of stable and strategy-proof rules is necessarily the study of the deferred acceptance. At the same time, however, the rapid developments of the matching with contracts literature has started to cover, in both theory and practice, the cases that violate those pre-known conditions for the deferred acceptance to be stable and/or strategy-proof (e.g., Dimakopoulos and Heller, 2014; Hatfield and Kojima, 2010; Kominers and Sonmez, 2013, 2014).

With those developments in mind, the purpose of this essay is to study stable and (one-sided) strategy-proof rules as generally as possible, and to disentangle some nature of such rules that arises purely from the two properties. Specifically, our results only require the choice functions on the hospital side to satisfy the irrelevance of rejected contracts (henceforth, IRC) condition, which is a very mild rationality requirement, but not any kind of “substitutes” condition.

Remark: note that in the anti-bullying school choice framework described in the first essay of this thesis, schools have choice functions (defined over the set of subsets of students) that naturally satisfy the IRC condition. Yet, what the two-round Gale-Shapley mechanism would achieve is slightly different from the standard stability notion. Hence, the current framework and the following results do not directly apply to anti-bullying school choice.

We first show that, under the IRC condition, the number of stable and strategy-proof rules is generically at most one. This result generalizes the existing results on the uniqueness of a stable and strategy-proof rule by Alcalde and Barbera (1994, Theorem 3) and Sakai (2011, Theorem 1). Second, we show that the doctor-optimal stable rule (i.e., the rule always choosing the “doctor-optimal” ( “student-optimal” ) stable matching), whenever it exists,

is the unique candidate for a stable and strategy-proof rule. In other words, if the doctor-optimal stable rule is well-defined and there is a stable and strategy-proof rule, that rule must be the doctor-optimal stable rule.

Compared to the existing uniqueness results, the aforementioned two results are technically novel for two related reasons. First, our proof of the first result requires no dominance relation between two hypothetical stable and strategy-proof rules. Consequently, it is applicable even when the doctor-optimal stable allocation does not always exist. In contrast, the uniqueness results by Alcalde and Barbera (1994, Theorem 3) and Sakai (2011, Theorem 1) are established by showing any stable rule that is strictly dominated by the doctor-optimal stable rule cannot be strategy-proof and hence, the existence of the doctor-optimal stable rule is critical in their proofs. Second, our proofs do not call for the rural hospital theorem either, which states that every agent (i.e., every doctor and every hospital) signs the same number of non-null contracts (i.e., assignment that is different from outside option) across all stable allocations. Instead we utilize Lemma 2 which states that under the IRC condition, given two distinct stable allocations, there is at least one doctor (student) who is assigned distinct non-null contracts by them, which could be seen as a weaker version of the rural hospital theorem but holds true without any restrictions on the hospitals' choice functions other than the IRC condition. It is this distinction that makes the proof of the second result non-trivial, although its statement might look very close to the previous results.

Third, we show that a stable and strategy-proof rule, when exists, is second-best optimal for doctor welfare, in the sense that no individually-rational and strategy-proof rule can dominate it. This result is further generalized to non-wasteful and strategy-proof rules. This extends the existing results in the school choice literature that the student-optimal stable rule is second-best optimal among strategy-proof rules (Abdulkadiroglu et al., 2009; Kesten, 2010; Kesten and Kurino, 2013).

Finally, we also show by example that the outcomes of a stable and strategy-proof rule do not always coincide with those of the cumulative offer process that is a generalization of the deferred-acceptance algorithm. Hence, the above three results hold not because the cumulative offer process is the only candidate for stable and strategy-proof rules. In other words, this example justifies our “algorithm-free” approach to stable and strategy-proof rules.

The third essay studies some pitfalls that can arise when we implement (one-sided) strategy-proof (two-sided) matching mechanisms. In markets such as school choice and the residency match, the whole agenda for the researchers is to achieve strategy-proofness that implies that telling the truth is a dominant strategy for every participant, together with other desiderata such as efficiency and stability/fairness. This state of affairs is borne out by a large body of evidence suggesting that when participants can beneficially tell a lie, a system would be harmed and eventually abandoned.

The purpose of this essay is to point out that despite the evidence in favor of strategy-proof matching mechanisms, some of them may be significantly harmed in actuality by either preference changes or “mismanipulations” ; in either case, reported preferences of students/doctors may differ from their true preferences.

Preferences may change for various reasons. One obvious reason is that students/doctors may acquire new information about the institutions after they reported their preferences, which eventually updates their evaluations. For New York City, the first deadline of high school application in the academic year 2012–2013 is December 10th in 2012. For Boston high school match, the first deadline of application is February 1st in 2013. In either case, students are given enormous amounts of time to collect the relevant information and reevaluate the schools.

On the other hand, in the matching literature, experiments have repeatedly confirmed that people misrepresent their preferences even if the underlying mechanisms are strategy-proof. For one instance, even if a central system tells participants to truthfully report their preferences, they may simply ignore or overlook the instructions. For one thing, the system may not clearly explain its algorithm to the participants and thus may not provide any grounds for believing its instructions. Indeed, in numerous school districts, their adopted mechanisms are not strategy-proof despite the fact that they are advocated to be so (Pathak and Sonmez 2013). For another thing, the standard matching theory may overlook some important aspects of real life such as peer effects. In such cases, strategy-proofness may lose its appeal.

When preferences may change, or preference manipulations prevail, (in)stability or (in)efficiency of a matching should be evaluated in terms of

the true/ex-post preferences. Based on this motivation, we show theoretically and by computer simulations that when these types of preference distortions prevail, instability that goes along with the Gale-Shapley (student-optimal stable matching) mechanism is still insignificant while inefficiency for the Top Trading Cycles mechanism is severely large. More specifically, (i) in a centralized matching market that mimics the Boston public high school choice system, regardless of whether the preference distortions prevail, the outcome of the Gale-Shapley mechanism is significantly inefficient. Similarly, the outcome of the Top Trading Cycles mechanism is significantly. Thus, as the extant matching literature has suggested, stability and efficiency are indeed incompatible. (ii) In the same market, when preferences change after students truthfully reported their preferences, the outcome of the Gale-Shapley mechanism is fairly stable while that of the Top Trading Cycles mechanism is significantly inefficient. (iii) In the same market, when preference mismanipulations prevail, the outcome of the Top Trading Cycles mechanism is significantly inefficient. Under the same condition, the outcome of the Gale-Shapley mechanism may be slightly instable, but that instability may indeed be insignificant since it is shown that anyone who constitutes some blocking pair must be one of those who mismanipulated in the first place.

Above all, our findings offer new rationales for adopting the Gale-Shapley mechanism rather than the Top Trading Cycles mechanism in real life, especially in school choice.