Journal of the Operations Research Society of Japan
Vol. 59, No. 1, January 2016, pp. 35–71

TWO-SIDED MATCHING WITH EXTERNALITIES: A SURVEY

Keisuke Bando Ryo Kawasaki Shigeo Muto
Tokyo Institute of Technology

(Received September 1, 2015; Revised December 14, 2015)

Abstract The literature on two-sided matching markets with externalities has grown over the past several years, as it is now one of the primary topics of research in two-sided matching theory. A matching market with externalities is different from the classical matching market in that agents not only care about who they are matched with, but also care about whom other agents are matched to. In this survey, we start with two-sided matching markets with externalities for the one-to-one case and then focus on the many-to-one case. For many-to-one matching problems, these externalities often are present in two ways. First, the agents on the "many" side may care about who their colleagues are, that is, who else is matched to the same "one." Second, the "one" side may care about how the others are matched.

Keywords: Game theory, two-sided matching, externalities, colleagues

1. Matching without Externalities

1.1. Introduction

Two-sided matching problems in the realm of game theory originated with the seminal paper by Gale and Shapley [25]. Their paper, in particular, looked at two problems: the "marriage" problem and the "college admissions" problem.[1] In these problems, there are two distinct groups, and the objective is to form suitable pairs containing one from each group. In the case of the marriage problem, the two groups are men and women, and pairs consist of one man and one woman. Moreover, a man can be paired with at most one woman, and a woman can be paired with at most one man – hence the current name, *one-to-one matching problem*. For the college admissions problem, the two sets are students and colleges, and the objective is to form student-college pairs. In contrast to the marriage problem, while a student may be matched to at most one college, a college can be matched to more than one student, giving the problem its current name, the *many-to-one matching problem*. The collection of these pairs is called a *matching*.

Agents on each side have what are called *preferences* or rankings over those on the other set; for example, students may rank colleges not just based on prestige but also based on tuition or geographic location, while colleges may rank students based on test scores, ethnicity, and other background information. When individuals have such preferences, it is not sensible to completely ignore these preferences and decide upon a matching arbitrarily.

One criterion often imposed on matchings is pairwise stability or in many papers, simply called stability. To illustrate this condition, suppose that there is a man m and a woman w who want to be matched to each other instead of their respective partners. Then, m and w would find it worthwhile to dump their partners so that they elope with each other. When

[1] They also consider a third type of problem with pairs are to be formed from one group, which in today's language is called roommate problems.

this occurs, the original matching is dissolved, and we do not expect that matching to hold up. Roughly, a matching is said to be pairwise stable if there does not exist such a pair.

Pairwise stability is a nice concept and a relatively easy one to grasp. It is not too stringent either, as Gale and Shapley [25] not only prove the existence of a pairwise stable matching, but also provides an algorithm, called the deferred acceptance (DA) algorithm, to find one. Moreover, Roth [52] makes the observation that stability is an important factor in the success of matching mechanisms in the example of matching medical interns to hospitals. One surprising observation in that paper is that one of the mechanisms, the National Intern Matching Program (now the National Resident Matching Program), is mathematically equivalent to the DA algorithm of Gale and Shapley [25].

The theory of two-sided matching so far has many positive results. One of the key assumptions in the model is that each agent only cares about whom he/she is matched with. However, in some applications, agents may care about with whom other agents are matched. One example is envy. A student may envy another student because he/she may be matched to a prestigious school. In an economic setting, in a small industry, firms may care about who their rival firms hire. These examples all fall under *matching problems with externalities.* When these externalities are present, the basic parts of the theory of standard matching problems no longer hold. First, defining pairwise stability is not straightforward; in the example of the eloping pair, they would also need to take into account how other agents match with each other when they elope. Second, whether a pairwise stable matching exists or not depends on the particular definition of pairwise stability. In this survey, we present several lines of research that have tried to overcome these obstacles.

A model that is closely related to matching with externalities is the matching model with couples. One instance where couples have played a large role is the matching of medical interns to hospitals.[2] Each intern has preferences over the different hospitals, but among the interns there may be a couple. Besides the individual preferences of each member of the couple, they would also prefer to be assigned to hospitals that are close by. In this case, one member of the couple does take into account the matching assignment of the other, so that this would be an example of a matching problem with externalities. It is also known that a stable matching may not exist (Theorem 10 of Roth [52]), and several positive results from the original matching problem do not carry over (Aldershof and Carducci [2]). Several attempts have been made to obtain positive results, such as determining a preference domain for existence (Klaus and Klijn [33]) and considering a large market (Kojima et al. [36]). We refer the reader to Biró and Klijn [9] for a comprehensive survey on this topic. We remark here that since the model of matching with couples treats a couple as a single agent instead of two separate agents, there are subtle differences in the stability concepts of the model with couples with those in the model with externalities. Therefore, there is not a straightforward relationship between matching with couples and the models and results presented in this survey.

Before explaining the details of these research papers, in the following, we review some essential facts in the basic theory of two-sided matching without externalities to help give the theory with externalities some context. Currently, there are many surveys and books on matching problems. To our knowledge, the list includes books by Roth and Sotomayor [59] and Gusfield and Irving [26] and surveys by Roth [56] and Roth [54] for general two-sided matching problems; Biró and Klijn [9] for matching problems with couples; and in relation to market design, which uses among many theories the theory of two-sided matching to

[2] For details on this setting, see Roth [52] and Roth and Peranson [58].

solve real-life problems involving markets, there are Roth [55] and Roth [57]. In contrast, while the literature on matching with externalities has grown recently, to our knowledge, there is no survey that provides an overall perspective on the topic. Thus, our objective in this survey is to organize some apparently disparate theories in the theory of matching with externalities.

1.2. The basic one-to-one matching model and the deferred acceptance (DA) algorithm

Let the set of agents be partitioned into two finite disjoint sets M and W, which are commonly called men and women respectively. Each $m \in M$ has a strict linear ordering $\succ_m$ over $W \cup \{m\}$ where $w \succ_m w'$ denotes that m prefers w over w'.[3] This ordering $\succ_m$ is often called the strict preference relation of m.[4] In addition, each $w \in W$ also has an strict linear ordering $\succ_w$ over $M \cup \{w\}$ indicating the preferences of w. These components together define a one-to-one *(classical) matching market* $(M, W, (\succ_i)_{i \in M \cup W})$.

Note that the element m is included in this list as a possibility for m to remain unmatched to any $w \in W$; therefore, $w \succ_m m$ is interpreted as m preferring to be matched to w over being unmatched, while $m \succ_m w$ is interpreted as m preferring to be unmatched over being matched to w. A woman $w \in W$ is said to be *acceptable* to $m \in M$ if $w \succ_m m$. Likewise, $m \in M$ is said to be acceptable to $w \in W$ if $m \succ_w w$.

In addition, for each $i \in M \cup W$, denote by $\succeq_i$ the weak ordering of $\succ_i$. For example, for $m \in M$, $w \succeq_m w'$ if $w \succ_m w'$ or $w = w'$. This ordering is also used quite frequently in the literature.

A *matching* μ is a one-to-one mapping[5] from $M \cup W$ to $M \cup W$ such that

1. $\mu(m) \in W \cup \{m\}$ for all $m \in M$,
2. $\mu(w) \in M \cup \{w\}$ for all $w \in W$,
3. $\mu(\mu(i)) = i$ for all $i \in M \cup W$.

The mapping μ gives as its output the identity of the input's partner, so that $\mu(m)$ indicates m's partner, and $\mu(w)$ indicates w's partner. In order for μ to be a matching, aside from being one-to-one, each man m's partner must be someone in the opposite set W or m himself, while each woman w's partner must be someone in the opposite set M or w herself. The first two conditions state this property formally. Moreover, if w is m's partner, then m must be w's partner, which is the third condition. The objective in this matching market is to find a matching μ that satisfies some suitable property.

A matching μ is said to be *pairwise stable* if it satisfies the following two conditions:

- For each $i \in M \cup W$, $\mu(i) \succeq_i i$.
- There does not exist a pair $(m, w) \in M \times W$ with $\mu(m) \neq w$ such that $m \succ_w \mu(w)$ and $w \succ_m \mu(m)$.

The first condition is often called *individual rationality.* If $i \succ_i \mu(i)$, then i would prefer to be single, and because each agent has an option to be single, individual rationality is a minimal requirement for a matching to persist. The second condition states that no pair would find it worthwhile to dump their partners under μ to be matched to each other. Such a pair is typically called a *blocking pair* with respect to μ. If there is such a blocking pair, then the matching μ will likely be dissolved. Roth [52] finds that in the matching market of

[3] A linear ordering is such that every element is comparable, and that the ordering satisfies transitivity.

[4] The term "strict" is used to denote that given two different women w and w' in W, m can either rank one strictly lower or higher than the other. In other words, m is not faced with a situation involving two women that he equally likes. This same restriction is also imposed on the preferences of w.

[5] μ is one-to-one if $i \neq j$ implies $\mu(i) \neq \mu(j)$.

medical interns to hospitals, this stability property distinguishes those matching processes that last from those that are abandoned.

Note that any stable matching μ is *Pareto efficient*; that is, there is no matching ν such that $\nu(i) \succeq_i \mu(i)$ for all $i \in M \cup W$ and $\nu(j) \succ_j \mu(j)$ for some $j \in M \cup W$. To see this, suppose otherwise. Without loss of generality, we assume that $\nu(m) \succ_m \mu(m)$ for some $m \in M$. By individual rationality of μ, we have that $\nu(m) \in W$, which we denote by w. Then, we have that $m \succeq_w \mu(w)$. By $\nu(m) \neq \mu(m)$, we have that $\nu(w) \neq \mu(w)$, which implies $m \succ_w \mu(w)$ because $\succ_w$ is strict. This contradicts that μ is stable.

While pairwise stability seems to be a rather strong property, Gale and Shapley [25] show that for any matching market, there exists at least one pairwise stable matching. Moreover, they define an algorithm, called the deferred acceptance (DA) algorithm, that produces a pairwise stable matching. Below is a verbal description of the algorithm.

(Men-proposing) DA Algorithm (One-to-one Case)

(Step 1.a) Each man $m \in M$ proposes to $w \in W$ whom he likes the most among those acceptable to m. If no such $w \in W$ exists, m is matched to himself.

(Step 1.b) Each woman $w \in W$ chooses the most preferred $m \in M$ who proposed to w and rejects all other men who have proposed to her. In such a case, m and w are tentatively matched to each other.

$\cdots$

(Step $k.a$) Each man $m \in M$ proposes to $w \in W$ whom he likes the most among those acceptable to m and who has not rejected m at an earlier step. If no such $w \in W$ exists, m is matched to himself.

(Step $k.b$) Each woman $w \in W$ chooses the most preferred $m \in M$ who proposed to w and her tentative partner and rejects all other men. In such a case, m and w are tentatively matched to each other.

The algorithm ends when no man $m \in M$ is rejected, and the tentative matching becomes the final matching. Gale and Shapley [25] show that this procedure yields a stable matching. To understand the basic reasoning, consider $m \in M$, and let μ^M be the matching that is produced from the above DA algorithm.

To disrupt the stability of the matching, m needs to find a w with $w \succ_m \mu^M(m)$ and $m \succ_w \mu^M(w)$. When running this algorithm, m must have proposed to w at some point since w is ranked higher than $\mu^M(m)$. Because m is not ultimately matched to w, w must have rejected m at some point. The reason why w would reject m is that w rejected m's proposal to accept someone whom she prefers over m or she accepts m's proposal but accepts a proposal from a different man m' in favor of m. Because each woman in the procedure always accepts a proposal from someone whom she likes more, w's ultimate partner must be preferred over m by w.

The key argument is that when w rejects m, w never regrets letting m go, as once w rejects m, w is always tentatively matched to a man that she prefers over m. This is a result not just from the algorithm itself, but on the basic assumption of the preferences. In more complex models, more assumptions on preferences are made so that this "regret-free" property still holds. It should be noted here that with preferences of w of m over m' that are dependent on whom her rival, say w' is matched to, the previous argument may not hold, thereby implying some of the complexity behind matching problems with externalities. We will come back to this issue in Section 2.

In addition to the matching being pairwise stable, Gale and Shapley [25] show that it is best matching among pairwise stable matchings for the agents in M. Such a matching is said to be *M-optimal*.

Theorem 1.1 (Gale and Shapley [25]). *Let μ^M be the matching that is produced from the above DA algorithm with men-proposing and let μ be any pairwise stable matching. Then, μ^M is pairwise stable and $\mu^M(m) \succeq_m \mu(m)$ for all $m \in M$.*

A DA algorithm can also be defined with the roles of M and W reversed so that women propose to men. Then, the matching that results is pairwise stable and W-optimal. Moreover, it is shown that the W-optimal stable matching is M-worst in that all members in M are weakly worse off compared to any other pairwise stable matching. Similarly, the M-optimal matching μ^M is also W-worst.

The model outlined so far imposes the restriction of each m being matched to one w (or unmatched). However, in examples such as matching interns to hospitals or students to schools, the hospitals or schools often admit more than one intern or student. In the next section, we explain the extension of the model to fit these applications.

1.3. Extension to many-to-one matching markets

In this section, we outline the basic model of matching markets where one side can be matched to more than one agent on the other side. However, we maintain the restriction on the other side so that one student can only be admitted to at most one school. This model corresponds to the "college admissions" model in Gale and Shapley [25].

Instead of M and W being the two disjoint sets, in this model, we consider a matching market with disjoint sets S and C, where an element in S is called a student, and an element in C is called a college or school.[6] Each student can be matched to at most one school, while a school may admit more than one student.

As before, each $s \in S$ has a strict preference ordering over $C \cup \{\emptyset\}$, labeled by $\succ_s$, where $\emptyset$ denotes that s is not matched to a $c \in C$. Each college $c \in C$ has strict preferences $\succ_c$ over $2^S := \{S'|S' \subseteq S\}$. Given a set $S' \subseteq S$, denote by $Ch_c(S')$, called the choice set, the top group of students in S' for c. That is, $Ch_c(S')$ denotes the set $T \subseteq S'$ such that $T \succeq_c S''$, $\forall S'' \subseteq S'$. Because preferences are more general, choice sets give us a compact way of defining analogues to the concepts discussed in the previous simpler models.

A matching is a set-valued mapping μ from $C \cup S$ to itself, satisfying the following conditions.

1. For each $s \in S$, either $\mu(s) = \{c\}$ for some $c \in C$ or $\mu(s) = \emptyset$.
2. For each $c \in C$, $\mu(c) \subseteq S$
3. For $s \in S$ and $c \in C$, $\mu(s) = \{c\}$ if and only if $s \in \mu(c)$.

A matching is said to be *individually rational* if for each $s \in S$, $\mu(s) \succeq_s \emptyset$ and for each $c \in C$, $\mu(c) = Ch_c(\mu(c))$. The main difference is the condition for $c \in C$. Because each $c \in C$ can unilaterally reject students in $\mu(c)$, if $\mu(c) \neq Ch_c(\mu(c))$, then $Ch_c(\mu(c)) \subsetneq \mu(c)$ and $Ch_c(\mu(c)) \succ_c \mu(c)$, so that college c would benefit from rejecting those in $\mu(c)$ but not in $Ch_c(\mu(c))$.

A matching is *pairwise stable* if there does not exist a pair $(c, s) \in C \times S$ with $s \notin \mu(c)$ such that $c \succ_s \mu(s)$ and $s \in Ch_c(\mu(c) \cup \{s\})$. If such a pair were to exist, then with s preferring c to his/her assigned college $\mu(s)$, c would also prefer to admit s, possibly with the cost of rejecting some students in $\mu(c)$.

[6] While the notation for one-to-one matching problems is usually M and W or M and F (for female), the notation for the many-to-one model may differ depending on the context. In this survey, for the models discussed, we mostly adhere to the notation used in the original papers.

We introduce an assumption called substitutability that is widely used in the literature. Preferences $\succ_c$ is said to satisfy *substituability* if for all $T \subseteq S$ and $s, s' \in T$ with $s \neq s'$,

$$s \in Ch_c(T) \Rightarrow s \in Ch_c(T \setminus \{s'\}).$$

Substitutability essentially eliminates any complementarities between any two students s and s'. The unavailability of s' should not affect the desirability of college c's to include s. The condition in its form for two-sided matching problems originates from Roth [53] and is an adaptation of the gross substitutes condition defined in Kelso and Crawford [31]. These substitutability conditions have also been connected to the literature on discrete convexity. Fujishige and Yang [24] first connects gross substitutability of Kelso and Crawford [31] to M$^\natural$-convexity of Murota and Shioura [45], which is an equivalent variant of Murota [43, 44]. Subsequent matching models that have used discrete convexity include (and are not confined to) Danilov et al. [13], Fujishige and Tamura [23], Murota and Yokoi [46], and Kojima et al. [37]. For a comprehensive survey on this direction, we refer the reader to the survey by Shioura and Tamura [63].

A more useful equivalent expression of substitutability is given by

$$s \notin Ch_c(T \setminus \{s'\}) \Rightarrow s \notin Ch_c(T).$$

Repeated application of the above expression yields the following. Let $T \subseteq T' \subseteq S$. Then,

$$s \notin Ch_c(T) \Rightarrow s \notin Ch_c(T').$$

The above form is the one given in the literature of matching with contracts in Hatfield and Milgrom [28].

Substitutability ensures that a pairwise stable matching exists. A modification of the DA algorithm can be used to show this fact.

(Student-proposing) DA Algorithm (Many-to-one Case)

(Step 1.a) Each $s \in S$ applies to the top-ranked college $c \in C$ according to $\succ_s$.

(Step 1.b) For each $c \in C$, let A_c^1 be the set of students that apply to c. Then, $c \in C$ tentatively keeps $Ch_c(A_c^1)$ and rejects the rest. For notational ease, let $T_c^1 := Ch_c(A_c^1)$ be the set of students that are kept by c.

$\cdots$

(Step k.a) Each $s \in S$ applies to the top-ranked college $c \in C$ among those that have not rejected s.

(Step k.b) Let A_c^k be the set of students that newly apply to $c \in C$. Each $c \in C$ tentatively keeps $Ch_c(A_c^k \cup T_c^{k-1})$ where T_c^{k-1} is the set of students that $c \in C$ kept after Step $k-1$.b, and rejects the rest.

The algorithm ends when no students are rejected. This version of the DA algorithm also yields a pairwise stable matching. The argument behind this result is the same as in the one-to-one case. Each college when rejecting a student does not regret doing so. This fact is reflected in the substitutability assumption.

To illustrate the logic, suppose that s is rejected in Step 1.b, thereby implying that $s \notin Ch_c(A_c^1) = T_c^1$. Note that $Ch_c(T_c^1 \cup \{s\}) = Ch_c(T_c^1) = T_c^1$, since T_c^1 is the most preferred set when A_c^1 is the set of available students, then it must also be the most preferred when $T_c^1 \cup \{s\}$ is the set of available students.[7] Therefore, $s \notin Ch_c(T_c^1)$ must also hold. In the

[7] The first equality is a direct consequence of a property of choice sets called *consistency*, while the second equality uses *idempotence*. See, for example, Alkan [3].

next period, the set of students A_c^2 additionally apply to c, and c is faced with choosing a subset from $T_c^1 \cup A_c^2$. Consider the hypothetical situation in which s is available so that the set of alternatives is $T_c^1 \cup A_c^2 \cup \{s\}$. Then, by substitutability, $s \notin Ch_c(T_c^1 \cup \{s\} \cup A_c^2)$. Therefore, once s is rejected, there is no scenario in which c would want to add s in the subsequent steps.

Another line of research to establish existence along with the lattice structure of pairwise stable matchings is to use Tarski's fixed point theorem. This line of research has its origins in Adachi [1] for one-to-one matching markets and has been used extensively in more complex models, including those that are covered extensively in this survey. We relegate the discussion on this topic until Section 3.

1.4. Summary and overview of the contents of this survey

Thus far, we have provided an overview on the results of two-sided matching markets without externalities in the one-to-one case and in the many-to-one case. This is only a sliver of the numerous results in the literature, and for the interested reader on this topic, we refer the reader to some of the literature listed in at the end of Section 1.1.

An underlying assumption in these models made on the preferences of each individual, was that each $m \in M$, for example, had preferences over W that were independent of how other people were matched. As stated in the introduction, envious agents are examples in which preferences may be dependent on the matching. In Section 2, we first outline the results on the literature for one-to-one matching markets with externalities. These situations arise, for example, in collegiate sports in which these colleges compete with each other through sports leagues. These examples can also be extended to matching problems between firms and workers, sports teams and athletes, or research facilities and researchers. It should be noted that the regret-free property that made the DA algorithm work does not hold in the model with externalities. Because preferences can be contingent on the matching of other agents, w may reject m at some point, but would want m back at a later point when the matchings of other agents have changed.

The asymmetry in the many-to-one case introduces a categorization of externalities that can only appear in many-to-one matching problems. One example of such externality is the many-to-one matching problem with colleagues. In terms of the student-school example, a student not only cares about the school to which he/she is matched, but also takes into account the other students matched to the same school – that is, his/her *colleagues*. Schools are assumed to only care about which set of students to admit and does not care about how others schools are matched. This specific form of externality does not appear in the one-to-one model since there were no "other students" that were matched to the same school. However, if we decompose the slots of each school into individual agents and view this problem as an artificial one-to-one matching problems, this would be a case of a matching problem with externalities without any specific label. We review the literature on this model in Sections 3 and 4.

2. Basic Models of Matching with Externalities

In this section, we introduce a matching problem with externalities, in which each agent cares not only his/her match but also on the matching of the other agents. When such externalities exist, it is not straightforward to define a stability concept because a stability of a matching depends on how a deviating pair expects the reaction of the other agents. Since Sasaki and Toda [61, 62] initiates the study on the matching problem with externalities, various stability concepts are proposed and these properties are investigated. We summarize the studies on matching problems with externalities. We also briefly summarize the recent development on the literature.

2.1. The one-to-one model and stability concepts

Sasaki and Toda [62], along with its working paper version Sasaki and Toda [61], are to our knowledge the earliest papers that consider matching with externalities for one-to-one matching problems. In their papers, they define several stability concepts for this model.

Let M and W be disjoint finite sets. Recall that in the matching model without externalities, a matching μ is pairwise stable if there does not exist a pair $(m, w) \in M \times W$ such that $m \succ_w \mu(w)$ and $w \succ_m \mu(m)$. The implicit assumption is that m and w can together leave their respective partners and be matched with each other. Whether such a move benefits both m and w did not matter how the other agents were matched. It may be the case that m may prefer to be matched with w if his former partner $\mu(m)$ remained unmatched, but may find it loathsome if $\mu(m)$ were matched to his nemesis, say m'. This is just one example of a case in which there are externalities; other examples include matchings between professional sports athletes and teams and matchings between workers and firms that compete with each other.

To incorporate this possibility, for each $m \in M$, instead of $\succ_m$ being a preference ordering over $W \cup \{m\}$, now $\succ_m$ is a preference ordering over the set of all matchings, denoted $\mathcal{M}$. Similarly, for $w \in W$, $\succ_w$ is a preference ordering over $\mathcal{M}$ as well. Assume that all preferences are strict so that for each $i \in M \cup W$ and for all $\mu \neq \mu'$, either $\mu \prec_i \mu'$ or $\mu \succ_i \mu'$.

One complicating issue for matching with externalities involves deciding on a stability concept. In principle, the idea should be the same; a matching μ should be deemed unstable if there exists a pair (m, w), not matched with each other in μ, but would benefit from being matched with each other than with their partners $\mu(m)$ and $\mu(w)$ respectively. However, whether such a defection from μ is beneficial depends heavily on how the others are matched. Mathematically, for a pair (m, w) define $\mathcal{M}(m, w)$, the set of matchings in which m and w are matched, by

$$\mathcal{M}(m, w) = \{\mu \in \mathcal{M} | \mu(m) = w\}.$$

It could be the case that for some $\mu' \in \mathcal{M}(m, w)$, $\mu' \succ_m \mu$ and $\mu' \succ_w \mu$, while for some $\mu'' \in \mathcal{M}(m, w)$, it may be the case that $\mu'' \prec_m \mu$. Also, there is a matching $\bar{\mu} \in \mathcal{M}(m, w)$ such that $\bar{\mu}$ is the matching induced by the deviation of (m, w) that keeps all other agents' assignments "unchanged, if possible." To express the phrase in quotation marks formally, we say that $\bar{\mu}$ is induced from μ by the pair (m, w) if the following conditions are satisfied:

- $\bar{\mu} \in \mathcal{M}(m, w)$,
- $\bar{\mu}(i) = \mu(i)$ for all $i \notin \{m, w, \mu(m), \mu(w)\}$,
- $\bar{\mu}(j) = j$ for $j \in \{\mu(m), \mu(w)\}$.

We use the notation $\mu \rightarrow_{m,w} \bar{\mu}$ if $\bar{\mu}$ satisfies the conditions above.[8] Note that $\bar{\mu}$ is a particular matching in $\mathcal{M}(m, w)$ and so represents a particular expectation held by both m and w over $\mathcal{M}(m, w)$.

Sasaki and Toda [61] define three types of stability concepts based on the expectations of m and w over the matchings in $\mathcal{M}(m, w)$. Let $\mu \in \mathcal{M}$ and let m and w be such that

[8] The notation is borrowed from Chwe [11], who calls it an effectiveness relation, and is used prominently in the literature that considers farsighted agents – agents who can anticipate sequences of deviations and are only interested in the final outcome of those deviations. See Mauleon et al. [41] and Klaus et al. [34] for the use of this notation in matching markets. In contrast, agents in this survey are myopic – they only consider the immediate consequence of their deviations.

$\mu(m) \neq w$.

1. m and w *optimistically blocks* μ if for some $\mu' \in \mathcal{M}(m,w)$, $\mu' \succ_m \mu$ and $\mu' \succ_w \mu$.
2. m and w *conservatively blocks* μ if for all $\mu' \in \mathcal{M}(m,w)$, $\mu' \succ_m \mu$ and $\mu' \succ_w \mu$.
3. m and w *blocks* μ if for the matching $\bar{\mu}$ such that $\mu \rightarrow_{m,w} \bar{\mu}$, $\bar{\mu} \succ_m \mu$ and $\bar{\mu} \succ_w \mu$.

In the first definition, the pair (m,w) deviates from μ if in the best scenario, m and w are better off by doing so. This optimism is reflected in the condition "for some $\mu' \in \mathcal{M}(m,w)$." In the second definition, the pair (m,w) deviates from μ if even in the worst scenario, m and w are better off by doing so. This conservatism is reflected in the condition "for all $\mu' \in \mathcal{M}(m,w)$." In the third definition, the matching $\bar{\mu}$ reflects the matching in which m and w are matched with each other, those who were previously matched with m or w ($\mu(m)$ and $\mu(w)$) are single (that is, matched with himself or herself), and all others uninvolved remain matched to the same partners under $\bar{\mu}$.[9]

Recalling from the definition of stability in the case when there are no externalities, we also need to define concepts of individual rationality. To do so, we define an analogue of $\mathcal{M}(m,w)$ involving the set of matchings that result when an agent i breaks up with a partner to be alone. For $i \in M \cup W$, define

$$\mathcal{M}(i) = \{\mu \in \mathcal{M} | \mu(i) = i\}.$$

That is, $\mathcal{M}(i)$ represents all matchings in which i is single.

Then, we can also define a matching $\bar{\mu}$ to be induced from μ via a single agent $i \in M \cup W$, denoted by $\mu \rightarrow_i \bar{\mu}$ if the following conditions are satisfied.

- $\bar{\mu} \in \mathcal{M}(i)$,
- $\bar{\mu}(j) = j$ for $j = \mu(i)$,
- $\bar{\mu}(k) = \mu(k)$ for all $k \neq i$, $k \neq \mu(i)$.

Similarly, we can define three types of individual rationality conditions, which correspond to the different expectations that an agent $i \in M \cup W$ have over possible matchings that may result from i being single.

1. A matching μ is *optimistically individually rational (O-IR)* if there do not exist $i \in M \cup W$ and a matching $\mu' \in \mathcal{M}(i)$ such that $\mu' \succ_i \mu$.
2. A matching μ is *conservatively individually rational (C-IR)* if there does not exist $i \in M \cup W$ such that for all μ' with $\mu \in \mathcal{M}(i)$, $\mu' \succ_i \mu$.
3. A matching μ is *individually rational (IR)* if there does not exist $i \in M \cup W$ such that for the matching μ' such that $\mu \rightarrow_i \mu'$, $\mu' \succ_i \mu$.

The stability of a matching can be defined using each of the three definitions of blocking and their corresponding individual rationality conditions. The three different definitions of stability are given in the following. Let $\mu \in \mathcal{M}$.

1. μ is *optimistically stable (O-stable)* if it is O-IR and there do not exist a pair (m,w) with $\mu(m) \neq w$ such that (m,w) optimistically blocks μ.
2. μ is *conservatively stable (C-stable)* if it is C-IR and there do exist pair (m,w) with $\mu(m) \neq w$ such that (m,w) conservatively blocks μ.
3. μ is *pairwise stable (P-stable)* if it is IR and there do not exist pair (m,w) with $\mu(m) \neq w$ such that (m,w) blocks μ.[10]

[9] In most papers, when a pair (m,w) deviates, the matching $\bar{\mu}$ is the one that results. Details about how agents other than m and w are matched were not needed in defining stability concepts for matching without externalities. $\bar{\mu}$ is the matching with the minimum number of changes to have m and w be matched to each other.

[10] Sasaki and Toda [61] calls a conservatively stable matching a pessimistically stable matching, but reserves

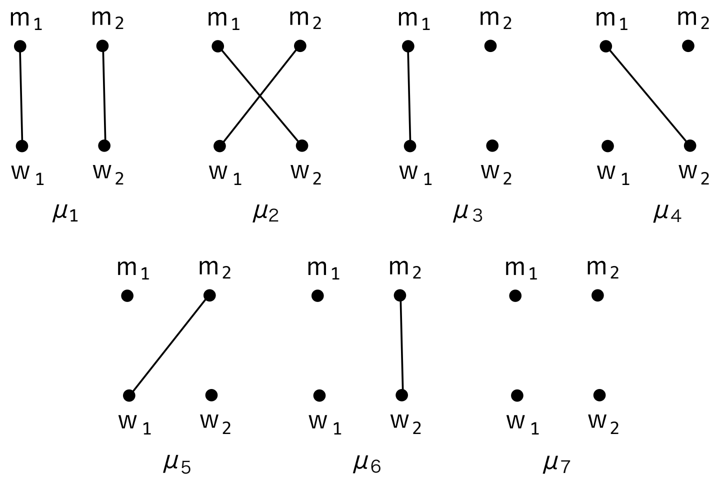

Figure 2.1: 7 matchings in Example 2.1

Now, let us introduce the following notation:

- S^O: the set of O-stable matchings
- S^C: the set of C-stable matchings
- S^P: the set of P-stable matchings

With externalities, three stability concepts are different (See Example 2.1). Therefore, a stability of a matching depends on how a deviating pair or a single agent anticipates the reaction of the other agents.

Sasaki and Toda [61] shows the following results.

Theorem 2.1 (Sasaki and Toda [61])**.** *Let S^O, S^C, and S^P be defined as above. Then, the following statements hold.*

1. *$S^O \subseteq S^P \subseteq S^C$ and $S^C \neq \emptyset$.*
2. *Every matching $\mu \in S^O$ is Pareto efficient. That is, there is no $\mu' \in \mathcal{M}$ such that for $\mu' \succeq_i \mu$ for all $i \in M \cup W$ and $\mu' \succ_j \mu$ for some $j \in M \cup W$.*
3. *There exists at least one matching $\mu \in S^C$ that is Pareto efficient.*

Theorem 2.1 highlights some strengths and weaknesses of S^C and S^O. S^C has the advantage that it is always nonempty, but not every C-stable matching is necessarily Pareto efficient. The following example illustrates this fact. This example also illustrates that $S^O \subsetneq S^P \subsetneq S^C$.

Example 2.1. Let $M = \{m_1, m_2\}$ and $W = \{w_1, w_2\}$. There are 7 matchings in this market which are denoted by Figure 2.1. m_1's preferences are given by:

$$\succ_{m_1}: \mu_2, \mu_5, \mu_7, \mu_1, \mu_6, \mu_3, \mu_4.$$

This means that m_1 ranks μ_2 first, μ_5 second, and so on. We assume that there are no externalities on m_2 in that there exists an ordering $\succ^*_{m_2}$ over $W \cup \{m_2\}$ such that for all $w, w' \in W \cup \{m_2\}$, $w \succ^*_{m_2} w'$ if and only if $\mu \succ_{m_2} \mu'$ for all $\mu \in \mathcal{M}(m_2, w)$ and all $\mu' \in \mathcal{M}(m_2, w')$. In this case, we can focus only on the associated ordering to check the blocking conditions and IR conditions. Specifically, m_2's preferences and the associated

the abbreviation P-stable for "pairwise stable." We have chosen the names here to avoid confusion in the abbreviation of the terms.

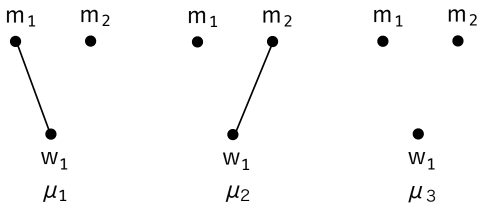

Figure 2.2: 3 matchings in Example 2.2

ordering are given by:

$$\succ_{m_2}: \mu_2, \mu_5, \mu_1, \mu_6, \mu_3, \mu_4, \mu_7 \ (\succ^*_{m_2}: w_1, w_2, m_2).$$

Similarly, we impose the same conditions on the women as on m_2. Their preferences and the associated orderings are given by:

$$\succ_{w_1}: \mu_1, \mu_3, \mu_2, \mu_5, \mu_6, \mu_4, \mu_7 \ (\succ^*_{w_1}: m_1, m_2, w_1)$$
$$\succ_{w_2}: \mu_1, \mu_6, \mu_2, \mu_4, \mu_5, \mu_3, \mu_7 \ (\succ^*_{w_2}: m_2, m_1, w_2).$$

In this market, we can show that $S^O \subsetneq S^P \subsetneq S^C$. To see this, we first consider μ_2. It is easy to see that μ_2 is O-IR. Moreover, each man ranks μ_2 first. This implies that μ_2 is O-stable.

Next, consider μ_1. Then, it is O-IR for m_2, w_1 and w_2. For m_1, μ_1 is IR, but not O-IR because $\mu_7 \succ_{m_1} \mu_1 \succ_{m_1} \mu_6$ holds. Moreover, μ_1 is ranked first for each woman, which implies that μ_1 is P-stable, but not O-stable.

Next, consider μ_5. This matching is IR. However, it is blocked by (m_1, w_2) and hence is not P-stable. On the other hand, the same pair cannot C-block μ_5 because m_1 anticipates that μ_4 occurs when he deviates with w_2 under the pessimistic expectation. Moreover, μ_5 is not C-blocked by (m_2, w_2) and (m_1, w_1). Therefore, μ_2 is C-stable, but not P-stable.

Note that μ_3, μ_4, and μ_7 are not C-stable because these are C-blocked by an unmatched pair. Moreover, μ_6 is not C-stable because it is C-blocked by (m_2, w_1).

From the above arguments, in this example, we have that

$$S^O = \{\mu_2\}, S^P = \{\mu_1, \mu_2\}, S^C = \{\mu_1, \mu_2, \mu_5\}.$$

In this market, C-stability does not imply Pareto efficiency. In fact, all agents strictly prefer μ_2 to μ_5, and hence μ_5 is not Pareto efficient. □

All matchings in S^O are at least Pareto efficient, but such matchings are not guaranteed to exist. Therefore, key properties of existence and efficiency that held for pairwise stable matchings for the case without externalities do not hold for each of the stable matchings considered thus far. The following example illustrates that S^O and S^P are empty.

Example 2.2. Let $M = \{m_1, m_2\}$ and $W = \{w_1\}$. In this market, there are 3 matchings which are denoted by Figure 2.2. Each agent's preferences are given by:

$$\succ_{m_1}: \mu_3, \mu_1, \mu_2 \qquad \succ_{m_2}: \mu_2, \mu_1, \mu_3 \qquad \succ_{m_3}: \mu_1, \mu_2, \mu_3$$

First, μ_1 is not IR because $\mu_3 \succ_{m_1} \mu_1$ holds. Next, μ_3 is blocked by (m_2, w_1) and the resulting matching is μ_2. However, μ_2 is blocked by (m_1, w_1). Therefore, S^P and S^C are empty. □

It should be noted that the DA algorithm cannot be used directly in the current setting. For $m \in M$, it may not be clear who in $w \in W$ is ranked the highest in m's list as m also may take into account how other agents are matched. This dependence also disrupts one of the key properties of the success of the DA algorithm. When $w \in W$ rejects $m \in M$, w does not regret doing so as she can hold on to a more favorable man, say m'. However, w's preference of m' over m may be reversed based on how other agents are matched throughout the process. In Section 2.3.3, we will give a formal discussion on the failure of the DA algorithm.

The nonemptiness of a C-stable matching, nonetheless, can be proved by artificially constructing a two-sided matching problem without externalities. Consider $m \in M$ and $w, w' \in W$. Let $\mu \in \mathcal{M}(m, w)$ be the least-preferred matching for m in the set $\mathcal{M}(m, w)$ – that is, $\mu \preceq_m \tilde{\mu}$ for all $\tilde{\mu} \in \mathcal{M}(m, w)$. Similarly, let $\mu' \in \mathcal{M}(m, w')$ be the least-preferred matching for m in the set $\mathcal{M}(m, w')$. Define the relation $\succ^*_m$ by

$$w \succ^*_m w' \Leftrightarrow \mu \succ_m \mu'.$$

Repeat for all pairs in $W \cup \{m\}$ to obtain binary comparisons of $\succ^*_m$. Because the original preferences $\succ_m$ satisfies transitivity, $\succ^*_m$ should yield a linear ranking over $W \cup \{m\}$. Define a similar preference relation $\succ^*_w$ for each $w \in W$ over $M \cup \{w\}$. Then, $(M, W, (\succ^*_i)_{i \in M \cup W})$ defines a two-sided matching market without externalities. Using the DA algorithm, there exists a pairwise stable matching μ^* in $(M, W, (\succ_i)_{i \in M \cup W})$, which is also a C-stable matching in the original market $(M, W, (\succ_i)_{i \in M \cup W})$ with externalities.

The following example illustrates the above construction. This example also illustrates that, in the associated one-to-one matching market $(M, W, (\succ^*_i)_{i \in M \cup W})$, while we can find at least one C-stable matching, we cannot find a Pareto efficient C-stable matching.

Example 2.3. Consider the same market as in Example 2.1. The associated one-to-one matching market without externalities is given by:

$$\succ^*_{m_1}: m_1, w_1, w_2 \quad \succ^*_{m_2}: w_1, w_2, m_2 \quad \succ^*_{w_1}: m_1, m_2, w_1 \quad \succ^*_{w_2}: m_2, m_1, w_2.$$

In this market, the unique stable matching is given by μ_5. However, the set of stable matchings in the associated market includes only the Pareto inefficient C-stable matching. Therefore, we cannot find a Pareto efficient C-stable matching unless focus on the original market. □

Example 2.3 shows that the associated one-to-one matching market may miss a Pareto efficient C-stable matching. Nevertheless, Theorem 2.1 states that there exists at least one Pareto efficient C-stable matching. To obtain this result, the key property is that if a C-stable matching μ is Pareto dominated by another matching ν, then ν is also C-stable. By repeatedly using this property, we can show the existence of a Pareto efficient C-stable matching.

While the focus has been on S^O and S^C, much of the recent literature focuses on S^P and variants of it due to the ease of interpretation of this stability concept. It is much in the spirit of Nash equilibrium in that when agents deviate, they assume that the situation involving other agents has not changed. However, Theorem 2.1 is silent about the existence of a P-stable matching. Mumcu and Saglam [42] derives a sufficient condition for the existence of a P-stable matching, based on a condition, called the top-coalition property, that is used for the existence of a core coalition structure for hedonic games introduced in Banerjee et al. [8].

Formally, let $P^i[k]$ be the kth best matching according to i's preferences. A set of matchings $\mathcal{V}$ is a *top matching collection* if $\mathcal{V} \neq \emptyset$ and for each $\mu \in \mathcal{V}$ and for each i, there exists $k \leq |\mathcal{V}|$ such that $\mu = P^i[k]$.

Informally, the top matchings collection is the set of matching consisting of k matchings such that each individual $i \in M \cup W$ can all agree to the fact that "the top k matches in terms of preferences are included in this set." Conversely, every individual would consider each matching outside this set to be ranked number $k+1$ or below.

We now present the result of Mumcu and Saglam [42], rewritten in terms of the language introduced in this survey.

Theorem 2.2 (Mumcu and Saglam [42])**.** *If a top matching collection $\mathcal{V}$ satisfies the following condition (*), then $\mathcal{V} \subseteq S^P$, and so $S^P \neq \emptyset$.*
(*) For all $\mu, \mu' \in \mathcal{V}$, either one of the following conditions hold: (i) if there exists a pair $\mu, \mu' \in \mathcal{V}$ such that $\mu \rightarrow_{m,w} \mu'$ for some (m, w), then either $\mu' \prec_m \mu$ or $\mu' \prec_w \mu$; (ii) $\mu \rightarrow_{i,j} \mu'$ does not hold for any $(i, j) \in M \times W$ or $i = j$ (this condition summarizes (ii) and (iii) of the original statement).

The conditions in Theorem 2.2 eliminate the cases in which one of the matchings in $\mathcal{V}$ acts as the matching $\bar{\mu}$ towards another matching also in $\mathcal{V}$, thereby making the latter matching not P-stable. Essentially, what the theorem states is that to check P-stability, first consider the matchings in $\mathcal{V}$ and then to check whether blocking can occur within just the set $\mathcal{V}$ itself.

Subsequent papers that extend analysis of matching with externalities to broader matching problems, such as many-to-one matching problems, focus primarily on P-stable matchings. These extensions will be summarized later in this survey. In what follows, we return to the paper of Sasaki and Toda [62], which define estimation functions, as one of the extensions that keep the one-to-one structure intact.

2.2. An extension of the one-to-one model using estimation functions

In the definitions of the stability concepts, it was assumed that when m and w consider deviating from a matching μ, they considered all or some of the possibilities from the set $\mathcal{M}(m, w)$. Instead, they may assume that some of those matchings in $\mathcal{M}(m, w)$ are not sensible for whatever reason. Sasaki and Toda [62] use what is called an estimation function to describe the set of matchings that m and w each thinks is plausible among those in $\mathcal{M}(m, w)$. In Sasaki and Toda [62], these estimations are defined *a priori* and are defined independently of the preferences of agents.

For expositional ease, consider the case when there is an equal number of men and women, and the set of feasible matchings are those in which each man m is matched to a woman w and vice-versa. Thus, there is no agent that is not matched.

Formally, let $m \in M$ and denote by $\varphi_m(w)$ the set of matchings μ that m perceives will occur if matched to w. Since m considers the possibilities of matchings that involve m matched with w, $\varphi_m(w) \subseteq \mathcal{M}(m, w)$ for all $w \in W$. This function φ_m defined on W constitutes the *estimation function* of $m \in M$. Similarly, for $w \in W$, we can define an estimation function φ_w.

Using the estimation functions, we can extend the definition of stability of matchings to this setting. Sasaki and Toda [62] only considers C-stable matchings for this setting, as one of the objectives is to investigate the extension of the existence result (Theorem 2.1).

Let estimation functions φ_i be defined for all $i \in M \cup W$. A matching μ is said to be *φ-stable* if there does not exist a pair (m, w) such that

$$\mu' \succ_m \mu, \ \forall \mu' \in \varphi_m(w)$$

and

$$\mu'' \succ_w \mu, \ \forall \mu'' \in \varphi_w(m)$$

That is, when considering the stability of a matching μ, m looks at all matchings that m expects to occur when matched to w, which is denoted by $\varphi_m(w)$, and w looks at all matchings that w expects to occur when matched to m, which is denoted by $\varphi_w(m)$. Note that it need not be the case that $\varphi_m(w) = \varphi_w(m)$ so that m and w need not agree on the estimation of the matchings that could be realized when matched to each other.

Sasaki and Toda [62] show a rather negative result regarding the existence of φ-stable matchings. Unless $\varphi_m(w) = \varphi_w(m) = \mathcal{M}(m, w)$, there exists an instance which does not admit a φ-stable matching. This result is a sharp contrast to the existence result stated earlier in Theorem 2.1. The formal statement is as follows.

Theorem 2.3 (Sasaki and Toda [62])**.** *Let M and W be two disjoint sets. Suppose that for some $m \in M$ and $w \in W$, $\varphi_m(w) \neq \mathcal{M}(m, w)$. Then, there exists a preference profile $(\succ_i)_{i \in M \cup W}$ such that the matching market $(M, W, (\succ_i)_{i \in M \cup W})$ does not admit a φ-stable matching.*

This negativity result is based on the fact that φ is fixed at the very beginning. In an extreme setting, these estimation functions can be arbitrary because they are defined independently from the preferences of the agents. As a modification to these estimation functions, Li [40] defines what are called *rational expectations* to make these estimation functions *consistent*.[11] When m and w deviate to form a pair, they should expect the set of agents $M \setminus \{m\}$ and $W \setminus \{w\}$ to be matched in a "stable" way. The formal definition starts with the usual definition of pairwise stability for a market with two agents each. Then, the definition proceeds for markets with more agents through a recursive formulation.

Hafalir [27] takes a different approach centered around the negativity result of Theorem 2.3. The trick behind the negativity result of Theorem 2.3 is that for a given profile of estimation functions for each $i \in M \cup W$, we can construct preferences such that φ-stable matchings do not exist without arguing whether estimations were sound based on the preferences at hand. Hafalir [27] first limits the estimation functions that can be construed based on the preferences of the agents. In this sense, estimation functions are endogenously given. Hafalir [27] then defines what are called sophisticated estimation functions and shows that these estimation functions can circumvent the negative result of Sasaki and Toda [62]. These estimation functions, like those of Li [40], are defined inductively. We refer the reader to the original paper for details.

2.3. Many-to-one matching markets with externalities among firms

In this section, we analyze the model of many-to-one matching with externalities among firms. That is, each firm's preferences depend not only on its workers, but also on the matching of its rival firms. Externalities do not exist among workers. This model is an extension of the model of Sasaki and Toda [61, 62] into a many-to-one matching market. In fact, we can directly apply the approach of Sasaki and Toda [61, 62] into this model. Bando [6, 7] takes a different approach from Sasaki and Toda [61, 62]. As we saw in Section 1.3, the choice function is a useful tool to analyze a classical model. Bando [6, 7] extends the choice function approach into the market with externalities. We summarize the results of Bando [6, 7].

2.3.1. Model

Let $F = \{f_1, \ldots, f_m\}$ be the set of m firms and $W = \{w_1, \ldots, w_n\}$ be the set of n workers. Each firm can hire multiple workers, but each worker is allowed to work at most one firm. A matching μ is a function from $F \cup W$ into $2^{F \cup W}$ such that (i) for each $w \in W$, either

[11] Technically, Li [40] is not a paper on matching theory but on principal-agent problems, which is closely related to organizational economics.

$\mu(w) = \{f\}$ for some $f \in F$ or $\mu(w) = \emptyset$, (ii) for each $f \in F$, $\mu(f) \subseteq W$, and (iii) for each $w \in W$ and $f \in F$, $\mu(w) = \{f\}$ if and only if $w \in \mu(f)$. Let $\mathcal{M}$ be the set of matchings. We often regard a matching μ as a tuple $(\mu(f_1), \dots, \mu(f_m))$. For $\mu \in \mathcal{M}$, $f \in F$ and $C \subseteq W$, define a matching $\mu^{f,C}$ as follows:

$$\begin{aligned}
&\text{if } w \in C, \text{ then } \mu^{f,C}(w) = \{f\},\\
&\text{if } w \ \in \mu(f) \setminus C, \text{ then } \mu^{f,C}(w) = \emptyset,\\
&\text{if } w \notin C \cup \mu(f), \text{ then } \mu^{f,C}(w) = \mu(w).
\end{aligned}$$

In other words, $\mu^{f,C}$ is a matching that is obtained from μ by satisfying f and C, keeping the matching of the other agents fixed. For $\mu \in \mathcal{M}$ and $f \in F$, we write $\cup_{f' \in F \setminus \{f\}} \mu(f')$ by $\mu(-f)$. That is, $\mu(-f)$ is the set of workers who are hired by f's rival firms in μ.

Each firm f has a strict preference ordering $\succ_f$ over the set of matchings $\mathcal{M}$. Each worker w has a strict preference ordering $\succ_w$ over $F \cup \{\emptyset\}$, where "$\emptyset$" represents being unmatched. We say that f is acceptable for w when $f \succ_w \emptyset$.

We next define choice functions for firms. In a classical model, for any subset of workers C, firm f's choice function $Ch_f(C)$ is f's most preferred workers among C. In our model, f's choice function may depend on the matching of the other firms. To define a choice function, we need to introduce additional notations. For $f \in F$ and $C \subseteq W$, define

$$R(f, C) = \{\mu \in \mathcal{M} | \mu(f) = \emptyset \text{ and } \mu(w) = \emptyset \text{ for all } w \in C\},$$

which is the set of matchings such that all workers in C are unemployed and firm f does not hire any workers. Then, f's choice function $Ch_f(C|\mu)$ is defined as firm f's most preferred subset of C given $\mu \in R(f, C)$. That is, $Ch_f(C|\mu)$ is the set such that (i) $Ch_f(C|\mu) \subseteq C$ and (ii) $\mu^{f, Ch_f(C|\mu)} \succeq_f \mu^{f,C'}$ for all $C' \subseteq C$.

Throughout, we assume that each firm's choice function depends on μ only through the workers hired by rival firms.[12] That is, for any $f \in F$, any $C \subseteq W$ and any $\mu, \mu' \in R(f, C)$, $\mu(-f) = \mu'(-f)$ implies $Ch_f(C|\mu) = Ch_f(C|\mu')$. Under this assumption, we can regard a choice function as a function from $\triangle$ to 2^W, where

$$\triangle := \{(C_1, C_2) | C_1, C_2 \subseteq W \text{ with } C_1 \cap C_2 = \emptyset\}.$$

We next define the stability concept provided by Bando [6]. As we saw in Section 2.1, when externalities exist, the stability of a matching depends on how a deviating pair expects the reaction of the other agents. We assume that a firm and workers deviate from a matching, assuming that the matching of the other agents is unchanged. This is the same assumption as that behind the P-stability provided in Section 2.1.

We first recall the definition of P-stable matching. A matching μ is individually rational if $\mu(w) \succeq_w \emptyset$ for all $w \in W$ and $Ch_f(\mu(f))|\mu(-f)) = \mu(f)$ for all $f \in F$. A pair $(f, C) \in F \times 2^W$ blocks a matching μ if $C \setminus \mu(f) \neq \emptyset$, $\mu^{f,C} \succ_f \mu$, and $f \succeq_w \mu(w)$ for all $w \in C$. We say that a matching μ is P-stable when it is individually rational, and is not blocked. We denote by S^P the set of P-stable matchings. In this model, S^P may also be empty. The following is the same example as in Example 2.2.

$$\succ_{f_1}: (\emptyset, \emptyset), (\{w_1\}, \emptyset), (\emptyset, \{w_1\}), \quad \succ_{f_2}: (\emptyset, \{w_1\}), (\emptyset, \emptyset), (\{w_1\}, \emptyset), \quad \succ_{w_1}: f_1, f_2, \emptyset.$$

[12] Without this assumption, Bando [6] shows that a stable matching does not exist.

Note that $(\{w_1\}, \emptyset)$ is not individually rational for f_1, $(\emptyset, \emptyset)$ is blocked by $(f_2, \{w_1\})$ and $(\emptyset, \{w_1\})$ is blocked by $(f_1, \{w_1\})$.

The definition of blocking admits incredible deviations. To see this, consider the above example. The matching $(\emptyset, \{w_1\})$ is blocked by $(f_1, \{w_1\})$. As a result of deviation, $(\{w_1\}, \emptyset)$ is formed. However, f_1 has an incentive to fire w_1 in $(\{w_1\}, \emptyset)$. Therefore, $(\{w_1\}, \emptyset)$ is not likely to be their final matching and this deviation may not be credible for w_1.

To rule out such incredible deviations, Bando [6] introduces the notion of strongly blocking and weak stability: A pair $(f, C) \in F \times 2^W$ *strongly blocks* a matching μ if it blocks μ, and $Ch_f(\mu(f) \cup C | \mu(-f) \setminus C) = C$. A matching μ is *weakly stable* if it is not strongly blocked. The additional condition guarantees that the matching obtained from satisfying a blocking pair is persistent. For example, $(f_1, \{w_1\})$ can block $(\emptyset, \{w_1\})$, but $Ch_{f_1}(\{w_1\}|\emptyset) \neq \{w_1\}$. So, $(f_1, \{w_1\})$ cannot strongly block $(\emptyset, \{w_1\})$. In the above market, $(\emptyset, \{w_1\})$ is a unique weakly stable matching. We denote by S^W the set of weakly stable matchings. By definition, we have $S^P \subseteq S^W$.

The following result shows that P-stability and weak stability are equivalent under positive externalities in that each firm is better off when its rival firms additionally hire new workers.

Theorem 2.4 (Bando [6]). *Suppose that for all $f \in F$ and $\mu, \mu' \in \mathcal{M}$, if $\mu(f) = \mu'(f)$ and $\mu(f') \subseteq \mu'(f')$ for all $f' \in F \setminus \{f\}$, then $\mu \preceq_f \mu'$. Then, $S^P = S^W$.*

We also define the notion of quasi-stability, which is a more mathematically tractable notion than the weak stability. A matching μ is *quasi blocked* by $(f, C) \in F \times 2^W$ if (i) $C \setminus \mu(f) \neq \emptyset$, (ii) $Ch_f(\mu(f) \cup C | \mu(-f) \setminus C) = C$ and (iii) $f \succeq_w \mu(w)$ for all $w \in C$, and is *quasi stable* if it is individually rational and is not quasi blocked. Note that the quasi blocking does not require $\mu^{f,C} \succ_f \mu$. By definition, we have that $S^Q \subseteq S^W$.[13] Therefore, to prove the existence of a weakly stable matching, it is sufficient to show the existence of a quasi stable matching. We will provide a sufficient condition for the existence in Section 2.3.2. Moreover, under the same assumption, worker-optimal and worker-worst quasi stable matchings exist. We will see this point in Section 2.3.3.

2.3.2. Restriction to preferences of firms

In this section, we introduce preference restrictions for firms to guarantee the existence of a quasi stable matching:

- Firm f's preferences satisfy *substitutability* (SUB) if for any $(C_1, C_2) \in \triangle$, $w, w' \in Ch_f(C_1|C_2)$ and $w \neq w'$ imply $w \in Ch_f(C_1 \setminus \{w'\}|C_2)$.
- Firm f's preferences satisfy *increasing choice* (IC) if for any $(C_1, C_2), (C_1', C_2') \in \triangle$, $C_1 = C_1'$ and $C_2 \subseteq C_2'$ imply $Ch_f(C_1|C_2) \subseteq Ch_f(C_1'|C_2')$.
- Firm f's preferences satisfy *no external effect by unchosen workers* (NEUW) if for any $(C_1, C_2) \in \triangle$, $w \notin Ch_f(C_1|C_2)$ and $w \in C_1$ imply $Ch_f(C_1 \setminus \{w\}|C_2) = Ch_f(C_1 \setminus \{w\}|C_2 \cup \{w\})$.

As we saw in Section 1.3, SUB is a sufficient condition for the existence of a stable matching in a many-to-one matching market without externalities. IC requires that the choice function of a firm expands when the set of workers hired by its rival firms expands. NEUW means that if firm f does not choose worker w from a subset of workers, then firm f's choice from another subset of workers in which worker w is excluded is not affected by a rival firm additionally hiring worker w. Intuitively, NEUW means that the external effect to firm f's choice function is caused only by an important worker for firm f.

[13] When externalities do not exist, $S^P = S^W = S^Q$ holds.

NEUW is related to a condition called irrelevance of rejected contracts (IRC).[14] In a classical model, IRC means that

$$Ch_f(C_1) \subseteq C_1' \subseteq C_1 \Rightarrow Ch_f(C_1) = Ch_f(C_1').$$

This condition is automatically satisfied when choice functions are constructed from preferences. It is well-known that when each firm's choice function satisfies SUB and IRC (even if each firm does not have preferences), there exists a stable matching and both conditions are crucial for the existence result. On the other hand, NEUW implies that letting $\hat{Ch}_f(C_1|C_2) := Ch_f(C_1|C_2 \setminus C_1)$ for each $C_1, C_2 \subseteq W$,

$$\hat{Ch}_f(C_1|C_2) \subseteq C_1' \subseteq C_1 \Rightarrow \hat{Ch}_f(C_1|C_2) = \hat{Ch}_f(C_1'|C_2).$$

Thus, we can regard NEUW as an extension of IRC into the model of externalities.

Bando [6] shows that under SUB, IC and NEUW, a quasi stable matching exists, and a quasi stable matching may not exist without one of these conditions. Moreover, under the same assumption, Bando [7] shows that worker-optimal and worker-worst quasi stable matchings exist. We will introduce the main result of Bando [7] in the next section.

2.3.3. Modified deferred acceptance

In this section, we introduce a modified deferred acceptance algorithm to find the worker-optimal quasi stable matching.

We say that a quasi stable matching μ is *worker-optimal* if for all quasi stable matching ν, $\mu(w) \succeq_w \nu(w)$ for all $w \in W$. When externalities exist, finding the worker-optimal quasi stable matching is not so straightforward. To see this, we first apply a sequential version of the DA algorithm in which (i) an unmatched worker proposes to his/her most preferred firm that has not rejected him and (ii) the proposed firm chooses the acceptable workers based on the current matching. The formal description is as follows:

- Step 0: The initial matching is given by $\mu_0(i) = \emptyset$ for all $i \in F \cup W$. Pick any worker w_0. w_0 proposes to his/her best firm, say f_0. If $w_0 \in Ch_{f_0}(\{w_0\}|\emptyset)$, then set $\mu_1 = \mu_0^{f_0,\{w_0\}}$. Otherwise, f_0 rejects w_0 and set $\mu_1 = \mu_0$.
- Step $k(\geq 1)$: Pick any unmatched worker w_k in μ_k. w_k proposes to his/her most preferred firm that has not rejected w_k at an earlier step, say f_k. In the case with $w_k \in Ch_{f_k}(\mu_k(f_k) \cup \{w_k\}|\mu_k(-f_k))$, consider $C^* = Ch_{f_k}(\mu_k(f_k) \cup \{w_k\}|\mu_k(-f_k))$. Then, f_k accepts C^* and rejects workers in $\mu_k(f_k) \setminus C^*$, and set $\mu_{k+1} = \mu_k^{f,C^*}$. If $w_k \notin Ch_{f_k}(\mu_k(f_k) \cup \{w_k\}|\mu_k(-f_k))$, f_k rejects w_k and define $\mu_{k+1} = \mu_k$.

This algorithm terminates when every worker is hired or every unmatched worker proposes all acceptable firms. Without externalities, this algorithm finds the worker-optimal stable matching. With externalities, it may not find the worker-optimal quasi stable matching. The following example illustrates this fact.

Example 2.4. Let $F = \{f_1, f_2, f_3\}$ and $W = \{w_1, w_2, w_3\}$. We assume that f_1 always wants to hire workers w_3 and w_4. Therefore, f_1's choice function satisfies:

$$Ch_{f_1}(W|\emptyset) = \{w_3, w_4\}.$$

f_1's choice function also satisfies:

$$Ch_{f_1}(\{w_1, w_2\}|\emptyset) = \emptyset, Ch_{f_1}(\{w_1, w_2\}|\{w_4\}) = \{w_1\},$$
$$Ch_{f_1}(\{w_1, w_2\}|\{w_3\}) = \{w_2\}, Ch_{f_1}(\{w_1, w_2\}|\{w_3, w_4\}) = \{w_1, w_2\}.$$

[14] See Ayğun and Sönmez [4]) for more details.

This means that f_1 does not want to hire workers w_1 and w_2 when its rival firms does not hire w_3 or w_4. However, f_1 wants to hire w_1 (w_2) when w_4 (w_3) is hired by a rival firm.

f_2 never wants to hire workers w_2 and w_4. For w_1 and w_3, f_2's choice function satisfies:

$$Ch_{f_2}(W|\emptyset) = \{w_1\}, Ch_{f_2}(\{w_3\}|\emptyset) = \emptyset, Ch_{f_2}(\{w_3\}|\{w_1\}) = \{w_3\}.$$

This means that f_2 always wants to hire w_1, and wants to hire w_3 if and only if w_1 is hired by a rival firm.

f_3 never wants to hire workers w_2 and w_4. For w_1 and w_3, f_3's choice function satisfies:

$$Ch_{f_3}(W|\emptyset) = \{w_3\}, Ch_{f_3}(\{w_1\}|\emptyset) = \emptyset, Ch_{f_3}(\{w_1\}|\{w_3\}) = \{w_1\}.$$

This means that f_3 always wants to hire w_3, and wants to hire w_1 if and only if w_3 is hired by a rival firm.

Each worker's preferences are given by:

$$\succ_{w_1}: f_1, f_3, \emptyset \quad \succ_{w_2}: f_1, \emptyset \quad \succ_{w_3}: f_2, \emptyset \quad \succ_{w_4}: f_3, \emptyset.$$

We apply the sequential DA algorithm to this market. Suppose that w_1 proposes to f_1 at step 0. f_1 rejects this offer because w_4 is not hired by f_2 or f_3. At step 1, suppose that w_1 proposes to f_3. This offer is also rejected because w_3 is not hired by f_1 or f_2. At step 3, suppose that w_2 proposes to f_1. This offer is also rejected because w_3 is not hired by f_2 or f_3. By repeating this argument, the seaquential DA algorithm yields the empty matching $(\emptyset, \emptyset, \emptyset)$. Note that for any ordering proposals, the seaquential DA algorithm yields the empty matching $(\emptyset, \emptyset, \emptyset)$ in this example. It is easy to see $(\emptyset, \emptyset, \emptyset)$ is a quasi stable matching. However, it is not worker-optimal because $(\{w_2\}, \{w_3\}, \{w_1\})$ is also a quasi stable matching □

In the above example, the sequential DA algorithm finds a quasi stable matching. However, in general, Bando [7] shows that the sequential DA algorithm may not find even a quasi stable matching. The reason for the failure of the algorithm is that each firm chooses its workers based on the current matching. To find the worker-optimal quasi stable matching, Bando [7] provides a modified DA algorithm, where each firm chooses workers, assuming that the other workers proposing its rival firms are hired. The formal description of this algorithm is defined as follows:

- Step 0: Each worker simultaneously proposes to his/her best firm. Then, we can define

  $$P(0) = \{w \in W | w \text{ proposes to some firm at step } 0\} = W,$$
  $$P_f(0) = \{w \in W | w \text{ proposes to } f \text{ at step } 0\} \text{ for each } f \in F.$$

  Each firm f accepts workers in $Ch_f(P_f(0)|P(0) \setminus P_f(0))$ and rejects workers in $P_f(0) \setminus Ch_f(P_f(0)|P(0) \setminus P_f(0))$. If some worker $w \in P(0)$ is rejected, then proceed to the next step. Otherwise, the algorithm terminates.

- Step $k(\geq 1)$: If a worker has been rejected by all acceptable firms in step $k-1$, then he cannot propose to any firm in this step. The other workers simultaneously propose to his/her best firm f that has never rejected him. Then, we can define

  $$P(k) = \{w \in W | w \text{ proposes to some firm at step } k\},$$
  $$P_f(k) = \{w \in W | w \text{ has proposed to } f \text{ at some step } j \text{ with } j \leq k\} \text{ for each } f \in F.$$

  Each firm f accepts workers in $Ch_f(P_f(k)|P(k) \setminus P_f(k))$ and rejects workers in $P_f(k) \setminus Ch_f(P_f(k)|P(k) \setminus P_f(k))$. If some worker $w \in P(k)$ is rejected by all firms that he has proposed to by step k; that is, $w \notin Ch_f(P_f(k)|P(k) \setminus P_f(k))$ for all $f \in F$ with $w \in P_f(k)$, then proceed to the next step. Otherwise, the algorithm terminates.

This algorithm terminates in finite steps, because at least one worker is rejected in every step and there exists only a finite number of firms. Note that for each step k, (i) $P(k)$ is the set of workers who propose to some firm at step k, (ii) $P_f(k)$ is the set of workers who have proposed to f at some step j with $j \leq k$, and (iii) $P(k) \setminus P_f(k)$ is the set of workers who have never proposed to f up until k but have proposed to some rival firm of f at step k. The key feature of this algorithm is that at each step k, each firm chooses its acceptable workers from $P_f(k)$, assuming that workers in $P(k) \setminus P_f(k)$ are accepted. By the definition, we have the following property: $P_f(k) \subseteq P_f(k+1)$, $P(k+1) \subseteq P(k)$ and hence $P(k+1) \setminus P_f(k+1) \subseteq P(k) \setminus P_f(k)$ for each $k \geq 0$. This monotone property, IC and SUB guarantee that each firm has no incentive to rehire workers whom it rejects earlier.

While the assumption that workers to be accepted by rival firms guarantees the monotone property of the algorithm, it may be inconsistent with its rival firms' actual choices (Example 2.5 illustrates this fact). However, the termination of the algorithm guarantees that the expectation is consistent with its rival firms' actual choices, because all workers are not rejected at the final step.

Let k^* be the termination of the algorithm. Define $\mu_{k^*}(f) = Ch_f(P_f(k^*)|P(k^*))$ for all $f \in F$. By the monotone property of the algorithm and the consistency at the termination, Bando [7] prove that μ_{k^*} is the worker-optimal quasi stable matching. As a summary, we have the following result.

Theorem 2.5 (Bando [7]). *If each firm's preferences satisfy SUB, IC and NEUW, then the modified DA algorithm converges to the worker-optimal quasi stable matching.*

The following example illustrates how the modified DA algorithm works.

Example 2.5. Consider the same market as in Example 2.4

In step 1, w_1 and w_2 propose to f_1, w_3 to f_2 and w_4 to f_3. $P(0)$ is given by W. For each $f \in F$, $(P_f(0), P(0) \setminus P_f(0))$ is defined as follows:

$$f_1 : (\{w_1, w_2\}, \{w_3, w_4\}) \quad f_2 : (\{w_3\}, \{w_1, w_2, w_4\}) \quad f_3 : (\{w_4\}, \{w_1, w_2, w_3\}).$$

Each firm f chooses from $P_f(0)$ assuming that all workers in $P(0) \setminus P_f(0)$ are hired by its rival firms. Therefore, f_1 accepts w_1 and w_2, because it assumes that w_3 and w_4 are hired. f_2 accepts w_3, because it assumes that w_1, w_2 and w_4 (especially, w_1) are hired. f_3 rejects w_4. Hence, the intermediate matching $\mu_0 = (\{w_1, w_2\}, \{w_3\}, \emptyset)$ is produced. Notice that while f_1 assumes that w_4 is accepted by f_3, w_4 is rejected by f_3. So, f_1's expectation is inconsistent with f_3's actual choices. In fact, f_1 has an incentive to fire w_1 at μ_0. However, f_1 does not reject w_1 in this step. Note that w_4 is rejected by all acceptable firms in this step and he cannot propose to any firms hereafter.

In step 1, w_1 and w_2 propose to f_1, w_3 to f_2. Therefore, $P(1)$ is given by $\{w_1, w_2, w_3\}$. For each $f \in F$, $(P_f(1), P(1) \setminus P_f(1))$ is defined as follows:

$$f_1 : (\{w_1, w_2\}, \{w_3\}) \quad f_2 : (\{w_3\}, \{w_1, w_2\}) \quad f_3 : (\{w_4\}, \{w_1, w_2, w_3\}).$$

Then, f_1 accepts w_2 and rejects w_1, assuming that w_3 is hired. f_2 accepts w_3, assuming that w_1 and w_2 (especially, w_1) are hired. Hence, the intermediate matching is given by $\mu_1 = (\{w_2\}, \{w_3\}, \emptyset)$. Note that in this step, f_2's expectation is inconsistent with f_1's actual choice.

In step 2, w_1 proposes to f_3, w_2 to f_1 and w_3 to f_2. Therefore, $P(2)$ is given by $\{w_1, w_2, w_3\}$. For each $f \in F$, $(P_f(2), P(2) \setminus P_f(2))$ is defined as follows:

$$f_1 : (\{w_1, w_2\}, \{w_3\}) \quad f_2 : (\{w_3\}, \{w_1, w_2\}) \quad f_3 : (\{w_1, w_4\}, \{w_2, w_3\}).$$

Then, f_1 accepts w_2, because it assumes that w_3 is hired. f_2 accepts w_3, because it assumes that w_1 and w_2 (especially w_1) are hired. f_3 accepts w_1, because it assumes that w_2 and w_3 (especially w_3) are hired. The intermediate matching is given by $\mu_2 = (\{w_2\}, \{w_3\}, \{w_1\})$. In this step, no rejections are issued and each firm's expectation is consistent with its rival firms' actual choice. The algorithm terminates at this step and μ_2 is the worker-optimal quasi stable matching. □

The modified DA algorithm can be generalized into a fixed point algorithm similar to those defined by Adachi [1] and Echenique and Oviedo [17]. By using the fixed point approach, we can also find the worker-worst stable matching. We refer the reader to the original paper for details.

2.4. Recent developments

Pycia and Yenmez [50] present the model of many-to-many matching with contracts that incorporates externalities. They analyze the model based on choice functions of agents that may depend on the matching of the other agents. However, they take a different approach from Bando [6, 7] in that they introduce the notion of consistent preorder over the set of matchings for each side of the market. Roughly speaking, when matching μ' is greater than another matching μ in a consistent preorder of firms (workers), μ' has a better market condition than μ for firms (workers). By using the consistent preorder, they extend substitutability into the market with externalities. It requires that when each firm's available set to choose expands and each firm faces a better market condition, it rejects more workers. They also extend the irrelevance of rejected contracts into the market with externalities. Then, they show that if each agent's choice function satisfies substitutability and irrelevance of rejected contracts, a stable matching exists. Theoretically, it is remarkable that in their setting, we cannot apply the Tarski's fixed point theorem.

Fisher and Hafalir [20] present the model of one-to-one matching with aggregate externalities. In their model, there exists an externality function that assigns, for each matching, a level of externalities measured by a real number, for example, amount of pollution. Each agent has a utility function that depends on his/her match and externality levels. Under the assumption that individuals have only a small effect on the externality level, they provide a sufficient condition for the existence of a stable matching. It is notable that their sufficient condition holds for specific economic models that include knowledge spillover and public goods models.

3. Many-to-one Matching with Preferences over Colleagues

In this section, we introduce many-to-one matching problems with preferences over colleagues in which each student cares about not only his/her match but also his/her colleagues. In this model, it is straightforward to define a stability of a matching, because the incentive that a college and students deviate from a matching does not depend on the reaction of the other agents. In Section 3.1, we introduce a basic model, and summarize the results of Dutta and Masso [15], which is the first study on the model of preferences over colleagues. In contrast to a classical problem, a stable matching may not exist even if we impose a strong assumption. In Section 3.2, we introduce an algorithm that finds all stable matchings or report that a stable matching does not exist, which is provided by Echenique and Yenmez [19]. In Section 3.3, we consider the relationship between the model of preferences over colleagues and the model without externalities. Specifically, the results of Kominers [38] and Flanagan [21] are summarized.

3.1. Model

Let C be a set of colleges and S be a set of students. As in the classical model, each college c has a strict preference ordering $\succ_c$ over 2^S. Each student cares not only about his/her match, but also about the other students in the same college. Therefore, each student s has a strict preference ordering $\succ_s$ over $(C \times S_s) \cup \{(\emptyset, \emptyset)\}$, where $S_s := \{S' \in 2^S | s \in S'\}$ is the set of all subsets of students that contain s and "$(\emptyset, \emptyset)$" denotes being unmatched. For example,

$$\succ_{s_1}: (c_2, \{s_1\}), (c_1, \{s_1, s_2\}), (\emptyset, \emptyset)$$

means that s_1 ranks c_2 with no colleagues first, c_1 with s_2 second, and only the $(c_2, \{s_1\})$ and $(c_1, \{s_1, s_2\})$ are acceptable for s_1.[15]

In the model with preferences over colleagues, it is useful to define a matching as a function that specifies to each student not only his/her college but also his/her colleagues. Formally, a matching μ is a mapping defined on $C \cup S$ that satisfies (i) $\mu(c) \in 2^S$ for all $c \in C$, (ii) $\mu(s) \in (C \times S_s) \cup \{(\emptyset, \emptyset)\}$ for all $s \in S$, (iii) $s \in \mu(c)$ implies $\mu(s) = (c, \mu(c))$ and (iv) when $\mu(s) = (c, S') \in C \times S_s$, $\mu(c) = S'$.

We define a standard stability concept used in the literature. A matching μ is *individually rational for students* if $\mu(s) \succeq_s (\emptyset, \emptyset)$ for all $s \in S$, and *for colleges* if $\mu(c) \succeq_c \emptyset$ for all $c \in C$. A matching μ is *individually rational* if it is individually rational for all agents. We say that a pair $(c, B) \in C \times 2^S$ *blocks* a matching μ if $B \neq \emptyset$ and $B \succ_c \mu(c)$ and $(c, B) \succeq_s \mu(s)$ for all $s \in B$, and *strongly blocks* μ if strict inequality holds for all students. Note that in the model of preferences over colleagues, a pair blocks a matching μ if and only if it strongly blocks μ because each agent's preferences are strict.[16] A matching μ is *stable* if it is individually rational and not blocked.

The definition of individual rationality for colleges are different from the classical model. In the classical model, the individual rationality of colleges requires that each college c prefers $\mu(c)$ to all subsets of $\mu(c)$. When this condition fails, a college can be better off by dumping some of its students without affecting the remaining students. In the model of preferences over colleagues, the remaining students are affected when some of their colleagues leaves. The stability concept defined here implicitly assumes that a college can dump some of its students only if the remaining students are not worse off (The definition of "block" includes such deviations). We also note that stable matchings are equivalent to core-stable matchings which take the possibility of "blocking" by larger coalitions into account (See Remark 3.1).

In contrast to the classical problem, a stable matching may not exist even if each college's preferences satisfy substitutability. The following example simply illustrates this fact.

Example 3.1. Let $C = \{c_1, c_2\}$ and $S = \{s_1, s_2\}$. Each college's and student's preferences are given by:

$$\succ_{c_1}: \{s_1, s_2\}, \{s_2\}, \{s_1\}, \emptyset \qquad \succ_{c_2}: \{s_2\}, \{s_1\}, \emptyset$$
$$\succ_{s_1}: (c_2, \{s_1\}), (c_1, \{s_1, s_2\}), (\emptyset, \emptyset)$$
$$\succ_{s_2}: (c_1, \{s_1, s_2\}), (c_2, \{s_2\}), (\emptyset, \emptyset).$$

In this example, matchings μ such that $|\mu(c_1)| = 1$ are not individually rational for students because c_1 is acceptable for $s_1(s_2)$ only if s_2 (s_1) matches with c_1. Hence, the following 4

[15] The ranking of unacceptable pairs of a college and colleagues are omitted.

[16] This is not true in a classical model.

matchings are individually rational:

$$\mu_1 : \begin{pmatrix} c_1 & c_2 \\ s_1, s_2 & \emptyset \end{pmatrix} \quad \mu_2 : \begin{pmatrix} c_1 & c_2 \\ \emptyset & s_1 \end{pmatrix} \quad \mu_3 : \begin{pmatrix} c_1 & c_2 \\ \emptyset & s_2 \end{pmatrix} \quad \mu_4 : \begin{pmatrix} c_1 & c_2 \\ \emptyset & \emptyset \end{pmatrix}.$$

However, μ_1 is blocked by $(c_2, \{s_1\})$, μ_2 is blocked by $(c_2, \{s_2\})$, and μ_3 and μ_4 are blocked by $(c_1, \{s_1, s_2\})$. Therefore, a stable matching does not exist. □

Remark 3.1. The core stability is a standard solution concept in cooperative game theory. In a core-stable matching, all coalitions of agents cannot deviate profitably. Formally, a set of colleges and students $C' \cup S'$ ($C' \subseteq C$ and $S' \subseteq S$) *blocks* a matching μ if $C' \cup S' \neq \emptyset$, and there exists a matching ν between C' and S' such that (i) $\nu(i) \succeq_i \mu(i)$ for all $i \in C' \cup S'$ and (ii) $\nu(i) \succ_i \mu(i)$ for some $i \in C' \cup S'$. We say that $C' \cup S'$ *strongly blocks* μ when strict inequality holds for all agents in the above definition. A matching is *core-stable (strict core-stable)* if it is strongly blocked (blocked). We denote by $Core$ ($Core^s$) the set of core-stable (strict core-stable) matchings, and by S^P the set of stable matchings. In the model of preferences over colleagues where all agents have strict preferences, it is straightforward to show that $S^P = Core^s = Core$.[17]

Dutta and Masso [15] search for domains of preferences that guarantee the existence of stable matchings. One of the domains of preferences that guarantees the existence of a stable matching is college-lexicographic preferences, in which the students first care about the college, then about their colleagues.[18] Formally, $\succ_s$ is *college-lexicographic* if there exists a strict preference ordering $\succ_s^*$ on $C \cup \{\emptyset\}$ such that for any $c, c' \in C$ with $c \neq c'$, $c \succ_s^* c'$ if and only if $(c, S_1) \succ_s (c', S_1')$ for all $S_1, S_1' \in S_s$, and $c \succ_s^* \emptyset$ if and only if $(c, S_1) \succ_s (\emptyset, \emptyset)$ for all $S_1 \in S_s$. When all students have college-lexicographic preferences, the matching market involving S and C with preferences given by $(\succ_S^*, \succ_C)$ constitutes a classical matching market. Then, (i) if a matching μ is individually rational in $(\succ_S^*, \succ_C)$, it is also individually rational in $(\succ_S, \succ_C)$, and (ii) if a pair (c, B) can block a matching μ in $(\succ_S, \succ_C)$, the same pair can also block a matching μ in $(\succ_S^*, \succ_C)$. Therefore, we have the following result.

Theorem 3.1 (Dutta and Masso [15])**.** *Suppose that all students have college-lexicographic preferences. Then, any stable matching in the associated classical market* $(\succ_S^*, \succ_C)$ *is stable in the original market* $(\succ_S, \succ_C)$.

In the classical model, a stable matching exists under substitutability. So, Theorem 3.1 implies that when all students have college-lexicographic preferences and each college's preferences satisfy substitutability, there exists a stable matching.

The college-lexicographic preferences are restrictive. Dutta and Masso [15] also analyze the existence of a stable matching on other domains of preferences. However, they obtain negative results. Specifically, they consider the domain of colleague-lexicographic preferences, in which the students first care about their colleagues then about their colleges. They further impose additional restrictions called *unanimous ranking according desirability (URD)* and *separability.* URD requires that there exists a common ranking over students such that all students prefer to join the set of students that contains higher-ranked students. Separability requires that (i) each student divides the set of students into a set of good students and bad students, and (ii) each student is always better off by joining a good student for him and worse off by joining a bad student for him. Dutta and Masso [15] shows that

[17] In a classical model, $S^P = C^s$ always holds, but C may be strictly larger than C^s.

[18] Dutta and Masso [15] consider a many-to-one matching with married couples as a special case of the model of preference over colleges. In this setting, they provide a sufficient condition for the existence of a stable matching. Revilla [51] extends this result.

even if each student's preferences satisfy URD or separability, there may not exist a stable matching.

One might expect that the converse of Theorem 3.1 holds. That is, any stable matching in the original market is stable in the associated classical market, and hence the model of preferences over colleagues with college-lexicographic preferences is exactly equivalent to the classical model. This is not true (See Example 3.2). We will discuss the relationship between the model of preferences over colleagues and the classical model in more details in Section 3.3.

Example 3.2. Consider the following market with one college and two students:

$$\begin{aligned}
&\succ_{c_1}: \{s_1, s_2\}, \{s_1\}, \{s_2\}, \emptyset\\
&\succ_{s_1}: (c_1, \{s_1\}), (c_1, \{s_1, s_2\}), (\emptyset, \emptyset)\\
&\succ_{s_2}: (c_1, \{s_2\}), (c_1, \{s_1, s_2\}), (\emptyset, \emptyset).
\end{aligned}$$

In this example, all students have college-lexicographic preferences. The associated classical market is given by:

$$\begin{aligned}
&\succ_{c_1}: \{s_1, s_2\}, \{s_1\}, \{s_2\}, \emptyset\\
&\succ^*_{s_1}: c_1, \emptyset\\
&\succ^*_{s_2}: c_1, \emptyset.
\end{aligned}$$

Clearly, the unique stable matching in this classical market is given by:

$$\mu_1 : \begin{pmatrix} c_1 \\ s_1, s_2 \end{pmatrix}$$

Hence, μ_1 is stable in the original market. Consider the following matching:

$$\mu_2 : \begin{pmatrix} c_1 & \emptyset \\ s_1 & s_2 \end{pmatrix}.$$

In the associated classical market, only $(c_1, \{s_1, s_2\})$ is the blocking pair for μ_2. However, $(c_1, \{s_1, s_2\})$ cannot block μ_2 in the original market because s_1 is worse off by joining with s_2. Therefore, μ_2 is also stable in the original market. □

3.2. T^2-Algorithm

As we saw in Example 3.1, a stable matching may not exist in the model of preferences over colleagues. Echenique and Yenmez [19] provide an algorithm that finds all stable matchings if any. We introduce their algorithm and extensions of their method to other models.

The algorithm is based on characterizing the stable matchings as the fixed points of a certain mapping. This characterization is a standard technique in the literature of matching problems (See Adachi [1], Fleiner [22], Echenique and Oviedo [17, 18], Hatfield and Milgrom [28], Ostrovsky [47]).

The mapping will be defined on the set of *individually rational prematchings.* A *prematching* is a mapping defined on $C \cup S$ that satisfies (i) $\mu(c) \in 2^S$ for all $c \in C$, and (ii) $\mu(s) \in (C \times S_s) \cup \{(\emptyset, \emptyset)\}$ for all $s \in S$. A prematching only specifies each college's students, and each student's college and his/her colleagues, and may not be a matching. We say that a prematching ν is *individually rational (IR)* if $\nu(s) \succeq_s (\emptyset, \emptyset)$ for all $s \in S$ and $\nu(c) \succeq_c \emptyset$ for all $c \in C$. We denote by Φ the set of all IR prematchings.

For each $\nu \in \Phi$, $c \in C$ and $s \in S$, define

$$U(c,\nu) := \{S' \subseteq S | (c,S') \succeq_s \nu(s) \text{ for all } s \in S'\},$$
$$V(s,\nu) := \{(c,S') \in C \times S_s | S' \succeq_c \nu(c) \text{ and } (c,S') \succeq_{s'} \nu(s') \text{ for all } s' \in S' \setminus \{s\}\} \cup \{(\emptyset,\emptyset)\}$$

In words, $U(c,\nu)$ is the collection of sets of students S' such that all of the students in S' weakly prefer (c,S') to their match at ν and $V(s,\nu)$ is the set of (c,S') such that all students in $S' \setminus \{s\}$ weakly prefer (c,S') to their match in ν and c weakly prefers S' to its students in ν. Note that when μ is a matching, $U(c,\mu)$ is the set of available stundets to form a blocking pair with c in μ, and $V(s,\mu)$ is the set of available students and colleges to form a blocking pair with s in μ. So, we say that $U(c,\nu)$ is the available sets for c in ν and $V(c,\nu)$ is the available sets for s in ν.

Now, define a mapping $T : \Phi \to \Phi$ as follows: for each $c \in C$ and $s \in S$,

$$(T\nu)(c) = \max_{\succ_c} U(c,\nu),$$

$$(T\nu)(s) = \max_{\succ_s} V(s,\nu).$$

That is, for each agent $i \in C \cup S$, $(T\nu)(i)$ is a maximal one with respect to i's preferences among the available sets in ν. Note that $T\nu$ is IR-prematching because $\emptyset \in U(c,\nu)$ and $(\emptyset,\emptyset) \in V(s,\nu)$ always hold.

It is easy to see that provided μ is a individually rational matching, $(T\mu)(c) = \mu(c)$ if and only if there exists no blocking pair that includes c in μ and, $(T\mu)(s) = \mu(s)$ if and only if there exist no blocking pair that includes s. This observation directly implies that any stable matching is a fixed point of T. Surprisingly, the converse holds. That is, any fixed point of T is a stable matching. Echenique and Yenmez [19] prove this fact by showing that when a prematching is a fixed point of T, it must be a matching. The proof technique used in their paper is the same as in Adachi [1]. As a summary, we have the following theorem.

Theorem 3.2 (Echenique and Yenmez [19])**.** *The set of stable matchings is equivalent to the set of fixed point of T.*

We next see that T is monotone decreasing by appropriately providing a partial order over the set of IR prematchings. Define an order $\geq$ over Φ by $\nu' \geq \nu$ if and only if $\nu'(i) \succeq_i \nu(i)$ for all $i \in S \cup C$. Note that (i) $\geq$ is a partial order because all agents' preferences are strict, and (ii) $(\geq, \Phi)$ is a finite lattice.[19]

Then, T is monotone decreasing in $\geq$. To see this, consider $\nu, \nu' \in \Phi$ with $\nu' \geq \nu$. By the definition, we have that for all $c \in C$ and $s \in S$, $U(c,\nu') \subseteq U(c,\nu)$ and $V(s,\nu') \subseteq V(s,\nu)$, which implies $\max_{\succ_c} U(c,\nu) \succeq_c \max_{\succ_c} U(c,\nu')$ and $\max_{\succ_s} V(s,\nu) \succeq_s \max_{\succ_s} V(s,\nu')$. Therefore, we can get $T\nu \geq T\nu'$.

Theorem 3.3 (Echenique and Yenmez [19])**.** *T is monotone decreasing in $\geq$.*

We next introduce T^2 mapping that is a key notion to define the algorithm. Define $T^2 : \Phi \to \Phi$ by $T^2\nu = T(T\nu)$ for all $\nu \in \Phi$. The above theorem directly implies the following result:

Theorem 3.4 (Echenique and Yenmez [19])**.** *T^2 is monotone increasing in $\geq$.*

[19] Let X be a finite set endowed with a partial order $\geq$. $(X, \geq)$ is a finite lattice if for any $x, y \in X$, the greatest lower bound on $\{x,y\}$ and the least upper bound on $\{x,y\}$ exist. Consider $(\Phi, \geq)$. For each $\nu, \nu' \in \Phi$, define $\nu \vee \nu' \in \Phi$ by $\nu \vee \nu'(i) := \max_{\succ_i}\{\nu(i), \nu'(i)\}$ for all $i \in C \cup S$, and $\nu \wedge \nu' \in \Phi$ by $\nu \wedge \nu'(i) := \min_{\succ_i}\{\nu(i), \nu'(i)\}$ for all $i \in C \cup S$. Then, $\nu \vee \nu'$ is the greatest lower bound on $\{\nu, \nu'\}$ and $\nu \wedge \nu' \in \Phi$ is the least upper bound on $\{\nu, \nu'\}$.

We denote by $\mathcal{E}(T^2)$ the set of fixed points of T^2. By Tarski's fixed point theorem, we have the following result:[20]

Theorem 3.5 (Echenique and Yenmez [19])**.** *$\mathcal{E}(T^2)$ is a nonempty lattice. In particular, the greatest fixed point of T^2 and least fixed point of T^2 exist; that is, there exist $\bar{\nu}, \underline{\nu} \in \epsilon(T^2)$ such that $\bar{\nu} \geq \nu \geq \underline{\nu}$ for all $\nu \in \mathcal{E}(T^2)$*

We can find, in finite steps, the greatest fixed point (least fixed point) of T^2 by the iterated application of T^2 starting from the greatest (least) element in Φ.[21] To see this, letting ν^0 be the greatest element in Φ, define $\nu^1 = T^2\nu^0$ and $\nu^k = T^2\nu^{k-1}$ for $k \geq 2$. Since ν^0 is the greatest element in Φ, we have $\nu^0 \geq \nu^1$. By monotonicity of T^2, $T^2\nu^0 \geq T^2\nu^1$ and hence $\nu^1 \geq \nu^2$. By repeating the same argument, we can get $\nu^k \geq \nu^{k+1}$. Since Φ is finite, there exists k^* such that $\nu^{k^*+1} = \nu^{k^*}$ and ν^{k*} is a fixed point of T^2. Thus, the iterative application of T^2 from ν^0 finds a fixed point of T^2 in finite steps. We next show that ν^{k^*} is the greatest fixed point of T^2. Pick any fixed point ν of T^2. Since ν^0 is the greatest point in Φ, we have $\nu^0 \geq \nu$. By monotonicity of T^2, we have $T^2\nu^0 \geq T^2\nu$ and hence $\nu^1 \geq \nu$. By repeating the same argument, we can get $\nu^k \geq \nu$ for each k. This implies $\nu^{k^*} \geq \nu$.

We now define T^2-algorithm that finds all stable matchings. This algorithm iteratively finds the greatest fixed points of T^2, while updating each agent's preferences. To define this algorithm, we need to introduce some additional notations. For each prematching ν and $i \in C \cup S$, let $\succ_i^\nu$ denote a strict preference ordering constructed from $\succ_i$ by making any one that i strictly prefers to $\nu(i)$ unacceptable without changing the ordering of the remaining ones. We denote by $(\succ_S, \succ_C)_\nu$ the market where each agent i's preferences are given $\succ_i^\nu$. Then, T^2-algorithms is defined as follows:

- Step 0: Define $\mathcal{E}^0 = \emptyset$. Find the least fixed point $\underline{\nu}$ of T^2 in $(\succ_S, \succ_C)$.
- Step 1: Find the greatest fixed point $\bar{\nu}$ of T^2 in $(\succ_S, \succ_C)$. If $T\bar{\nu} = \bar{\nu}$, define $\mathcal{E}^1 := \{\bar{\nu}\}$. Otherwise, define $\mathcal{E}^1 := \mathcal{E}^0$. If $\bar{\nu} = \underline{\nu}$, the algorithm terminates at this step. If $\bar{\nu} > \underline{\nu}$, let $Q^1 := \{\bar{\nu}\}$ and proceed to the next step.
- Step k ($k \geq 2$): For each $\nu \in Q^{k-1}$ and each $i \in C \cup S$ with $\nu(i) \succ_i \underline{\nu}(i)$, find $\nu[i]$ by the following procedure:
    - Let $n(i, \nu(i)) := \max_{\succ_i}\{j | \nu(i) \succ_i j\}$. That is, i ranks $n(i, \nu(i))$ next to $\nu(i)$.
    - Define a prematching ν' by $\nu'(i) = n(i, \nu(i))$ and $\nu'(j) = \nu(j)$ for all $j \in C \cup S$ with $j \neq i$.
    - Find the greatest fixed point $\bar{\nu}$ of T^2 in $(\succ_S, \succ_C)_{\nu'}$ and let $\nu[i] := \bar{\nu}$.

  Let $\bar{Q}$ be the set of all prematchings that are found in the above procedure; that is,

  $$\bar{Q} := \{\nu' | \nu' = \nu[i] \text{ for some } \nu \in Q^{k-1} \text{ and some } i \in C \cup S\}.$$

  Define $\bar{Q}_1$ and $\bar{Q}_2$ by

  $$\bar{Q}_1 := \{\nu' \in \bar{Q} | T\nu' = \nu' \text{ in } (\succ_C, \succ_S)\},$$

  $$\bar{Q}_2 := \{\nu' \in \bar{Q} | \nu' > \underline{\nu}\} \setminus \bar{Q}_1.$$

  Let $\mathcal{E}^k := \mathcal{E}^{k-1} \cup \bar{Q}_1$. If $\bar{Q}_2 = \emptyset$, then the algorithm terminates at this step. Otherwise, let $Q^k := \bar{Q}_2$ and proceed to the next step.

[20] Tarski's fixed point theorem states that if f is increasing function from a finite lattice into itself, then the set of fixed points is a nonempty lattice.

[21] The greatest IR prematching is given by ν^0 such that $\nu^0(s) = \max_{\succ_s}(C \times S_s) \cup (\emptyset, \emptyset)$ for all $s \in S$ and $\nu^0(c) = \max_{\succ_c} 2^S$ for all $c \in C$. The least IR prematching is given by $\nu^\emptyset$ such that $\nu^\emptyset(s) = (\emptyset, \emptyset)$ for all $s \in S$ and $\nu^\emptyset(c) = \emptyset$ for all $c \in C$.

Echenique and Yenmez [19] prove that T^2-algorithm terminates in finite steps. Let k^* be the termination of the algorithm. Any matching $\nu \in \mathcal{E}^{k^*}$ is stable because $T\nu = \nu$ holds by the definition. The converse holds. That is, any stable matching μ is in $\mathcal{E}^{k^*}$ if any and hence T^2-algorithm never misses all stable matchings. To show this result, Echenique and Yenmez [19] prove that for any step $k \geq 1$, $\mu \notin \mathcal{E}^k$ implies that there exists $\nu \in Q^k$ such that $\nu \geq \mu$, and hence proceed to the next step. This property implies that for some step $k \geq 0$, $\mu \in \mathcal{E}^k$ because the algorithm terminates in finite steps. As a summary, we have the following theorem.

Theorem 3.6 (Echenique and Yenmez [19]). *Let k^* be the termination of T^2-algorithm. Then, $\mathcal{E}^{k^*}$ coincides with the set of all stable matchings. Therefore, T^2-algorithm finds all stable matchings or reports that a stable matching does not exist.*

The following example illustrates how T^2-algorithm works.

Example 3.3. Consider the same market as in Example 3.2:

$$\begin{aligned}
&\succ_{c_1}: \{s_1, s_2\}, \{s_1\}, \{s_2\}, \emptyset \\
&\succ_{s_1}: (c_1, \{s_1\}), (c_1, \{s_1, s_2\}), (\emptyset, \emptyset) \\
&\succ_{s_2}: (c_1, \{s_2\}), (c_1, \{s_1, s_2\}), (\emptyset, \emptyset).
\end{aligned}$$

We apply T^2-algorithm to this market. Throughout this example, we denote by $[\nu(s_1), \nu(s_2), \nu(c_1)]$ a prematching ν.

Step 0: We first calculate the least fixed point of T^2. Table 1 shows the iteration of T starting from the empty matching denoted by $\nu^\emptyset$.

	s_1	s_2	c_1
$\nu^\emptyset$	$(\emptyset, \emptyset)$	$(\emptyset, \emptyset)$	$\emptyset$
$T\nu^\emptyset$	$(c_1, \{s_1\})$	$(c_1, \{s_2\})$	$\{s_1, s_2\}$
$T^2\nu^\emptyset$	$(\emptyset, \emptyset)$	$(\emptyset, \emptyset)$	$\{s_1\}$
$T^3\nu^\emptyset$	$(c_1, \{s_1\})$	$(c_1, \{s_1, s_2\})$	$\{s_1, s_2\}$
$T^4\nu^\emptyset$	$(c_1, \{s_1, s_2\})$	$(\emptyset, \emptyset)$	$\{s_1\}$
$T^5\nu^\emptyset$	$(c_1, \{s_1\})$	$(c_1, \{s_1, s_2\})$	$\{s_1, s_2\}$
$T^6\nu^\emptyset$	$(c_1, \{s_1, s_2\})$	$(\emptyset, \emptyset)$	$\{s_1\}$

Table 1: The iteration of T from $\nu^\emptyset$ in $(\succ_S, \succ_C)$

Table 1 reveals that $T^4\nu^\emptyset = T^6\nu^\emptyset$ and hence $\underline{\nu} := [(c_1, \{s_1, s_2\}), (\emptyset, \emptyset), \{s_1\}]$ is the least fixed point of T^2. Note that $\underline{\nu}$ is not a matching because $\underline{\nu}(s_1) = (c_1, \{s_1, s_2\})$ but $s_2 \notin \underline{\nu}(c_1)$.

Step 1: Calculate the greatest fixed point of T^2 in the original market. Note that $T\nu^\emptyset$ is the greatest IR prematching. Table 1 reveals that $\bar{\nu} := [(c_1, \{s_1\}), (c_1, \{s_1, s_2\}), \{s_1, s_2\}]$ is the greatest fixed point of T^2. By $T\bar{\nu} \neq \bar{\nu}$ and $\bar{\nu} > \underline{\nu}$, set $\mathcal{E}^1 = \emptyset$ and $Q^1 = \{\bar{\nu}\}$, and proceed to the next step.

Step 2: We denote $\bar{\nu}$ by overlines and $\underline{\nu}$ by underlines as follows:

$$\begin{aligned}
&\succ_{c_1}: \overline{\{s_1, s_2\}}, \underline{\{s_1\}}, \{s_2\}, \emptyset \\
&\succ_{s_1}: \overline{(c_1, \{s_1\})}, \underline{(c_1, \{s_1, s_2\})}, (\emptyset, \emptyset) \\
&\succ_{s_2}: (c_1, \{s_2\}), \overline{(c_1, \{s_1, s_2\})}, \underline{(\emptyset, \emptyset)}.
\end{aligned}$$

Clearly, all agents strictly prefer $\bar{\nu}$ to $\underline{\nu}$. Therefore, in step 2, we need to calculate $\bar{\nu}[c_1], \bar{\nu}[s_1]$ and $\bar{\nu}[s_2]$.

To obtain $\bar{\nu}[c_1]$, we need to consider the preference profile $(\succeq_S, \succeq_C)_{\nu'}$ where ν' is the prematching defined by $[\{s_1\}, (c_1, \{s_1\}), (c_1, \{s_1, s_2\})]$, which is given by:

$$\begin{aligned}
&\succ_{c_1}\colon \{s_1\}, \{s_2\}, \emptyset\\
&\succ_{s_1}\colon (c_1, \{s_1\}), (c_1, \{s_1, s_2\}), (\emptyset, \emptyset)\\
&\succ_{s_2}\colon (c_1, \{s_1, s_2\}), (\emptyset, \emptyset).
\end{aligned}$$

We can obtain $\bar{\nu}[c_1]$ by calculating the greatest fixed point of T^2 in $(\succ_S, \succ_C)_{\nu'}$. Table 2 shows that the iteration of T starting from ν' in $(\succ_S, \succ_C)_{\nu'}$. This implies that $[(c_1, \{s_1\}), (\emptyset, \emptyset), \{s_1\}]$ is the greatest fixed of T^2 in $(\succ_S, \succ_C)_{\nu'}$. Therefore, $\bar{\nu}[c_1]$ is given by $[(c_1, \{s_1\}), (\emptyset, \emptyset), \{s_1\}]$.

	s_1	s_2	c_1
ν'	$(c_1, \{s_1\})$	$(c_1, \{s_2, s_2\})$	$\{s_1\}$
$T\nu'$	$(c_1, \{s_1\})$	$(\emptyset, \emptyset)$	$\{s_1\}$
$T^2\nu'$	$(c_1, \{s_1\})$	$(\emptyset, \emptyset)$	$\{s_1\}$

Table 2: The iteration of T from ν' in $(\succ_S, \succ_C)_{\nu'}$

To obtain $\bar{\nu}[s_1]$, we need to calculate the fixed point of T^2 in the following market:

$$\begin{aligned}
&\succ_{c_1}\colon \{s_1, s_2\}.\{s_1\}, \{s_2\}, \emptyset\\
&\succ_{s_1}\colon (c_1, \{s_1, s_2\}), (\emptyset, \emptyset)\\
&\succ_{s_2}\colon (c_1, \{s_1, s_2\}), (\emptyset, \emptyset).
\end{aligned}$$

It is easy to see that $[(c_1, \{s_2, s_2\}), (c_1, \{s_1, s_2\}), \{s_1, s_2\}]$ is the greatest fixed point of T^2. Therefore, $\bar{\nu}[s_1] = [(c_1, \{s_2, s_2\}), (c_1, \{s_1, s_2\}), \{s_1, s_2\}]$.

To obtain $\bar{\nu}[s_2]$, we need to calculate the fixed point of T^2 in the following market:

$$\begin{aligned}
&\succ_{c_1}\colon \{s_1, s_2\}, \{s_1\}, \{s_2\}, \emptyset\\
&\succ_{s_1}\colon (c_1, \{s_1\}), (c_1, \{s_1, s_2\}), (\emptyset, \emptyset)\\
&\succ_{s_2}\colon (\emptyset, \emptyset).
\end{aligned}$$

It is easy to see that $[(c_1, \{s_1\}), (\emptyset, \emptyset), \{s_1\}]$ is the greatest fixed point of T^2. Therefore, $\bar{\nu}[s_2] = [(c_1, \{s_1\}), (\emptyset, \emptyset), \{s_1\}]$.

In this step, we have that

$$\bar{Q} = \{\bar{\nu}[c_1], \bar{\nu}[s_1], \bar{\nu}[s_2]\} = \{[(c_1, \{s_1\}), (\emptyset, \emptyset), \{s_1\})], [(c_1, \{s_2, s_2\}), (c_1, \{s_1, s_2\}), \{s_1, s_2\}]\}.$$

We can confirm that for all $\nu \in \bar{Q}$, ν is the fixed point of T in the original market. Therefore, $\bar{Q}^1 = \bar{Q}$, $\bar{Q}^2 = \emptyset$ and $\mathcal{E}^2 = \bar{Q}^1$. From $\bar{Q}^2 = \emptyset$, the algorithm terminates at this step. The set of all stable matchings is given by:

$$\mathcal{E}^2 = \{[(c_1, \{s_1\}), (\emptyset, \emptyset), \{s_1\})], [(c_1, \{s_2, s_2\}), (c_1, \{s_1, s_2\}), \{s_1, s_2\}]\}.$$

□

Interestingly, T^2-algorithm was first introduced for the model with preferences over colleagues and had no counterpart in the market without externalities, when it was provided.[22]

[22] T^2-algorithm is motivated by the algorithm for noncooperative games presented by Echenique [16], which finds all Nash equilibria in games with strategic complementarities.

However, T^2-algorithm also works in other models. Flanagan [21] shows that T^2-algorithm works for a many-to-one matching model without externalities. Kojima [35] shows that T^2-algorithm works for a two-sided matching with married couples.

İnal [29] extends the method of Echenique and Yenmez [19] to a coalition formation model that is introduced by Banerjee et al. [8] and Bogomolnaia and Jackson [10]. In a coalition formation model, each payer has preferences over coalitions that contains him. The outcome is given by a partition of the agents. A partition of the agents is core-stable if there is no coalition of agents such that each agents in the coalition are strictly better off compared with the coalition that contains him in the partition. A many-to-one matching can be seen as a partition of colleges and students (See Section 4.1 in details). Moreover, the set of stable matchings in the model of preferences over colleagues is equivalent to core-stable matchings (See Remark 3.1). Therefore, a coalition formation model includes the model of preferences over colleagues. İnal [29] characterizes the set of core-stable partitions by the set of fixed points of a mapping that is similar to T, provided that all agents have strict preferences. Based on this result, the algorithm to find all core-stable partitions if any is proposed.

Note that T^2-algorithm may search all possible matchings and may perform very slowly in the worst case. However, Echenique and Yenmez [19] expect that the algorithm is fast in many applications. It would be interesting to analyze the algorithm performance by a computer simulation.

3.3. Embedding

In this subsection, we consider the relationship between the model of preferences over colleagues and the classical model. Kominers [38] shows that for any classical model $(\succ_S^*, \succ_C)$, there exists a model of preferences over colleagues $(\bar{\succ}_S, \bar{\succ}_C)$ such that the sets of stable matchings in $(\succ_S^*, \succ_C)$ and $(\bar{\succ}_S, \bar{\succ}_C)$ coincide. In other words, any classical many-to-one matching model can be embedded into the model with preferences over colleagues. We introduce this result and its extension.

As we saw in Example 3.2, given a classical model, if we generate a model of preferences over colleagues in which each student has a college-lexicographic preferences, the sets of stable matchings in two-models may be different. Kominers [38] shows that by appropriately providing lexicographic preferences, the set of stable matchings of a classical model can be represented by the model of preferences over colleagues. To see this, consider a classical model $(\succ_S^*, \succ_C)$. Define the associated model of preferences over colleagues $(\bar{\succ}_S, \bar{\succ}_C)$ as below:

- For all $c \in C$, $\bar{\succ}_c$ coincides with $\succ_c$.
- Student s's preferences $\bar{\succ}_s$ are defined as follows. We represent $\succ_s^*$ by:

$$\succ_s^*\colon c_1, c_2, c_3, \cdots, c_k, \emptyset, c_{k+1}, \cdots, c_m.$$

For each $c \in C$, we represent the restriction of $\succeq_c$ into $S_s := \{S' \subseteq S | s \in S'\}$ by:

$$\succ_c^*\colon S_1^c, S_2^c, \cdots, S_{|S_s|}^c.$$

Define college lexicographic preferences $\bar{\succ}_s$ by:

$$\begin{aligned} &\bar{\succ}_s : (c_1, S_1^{c_1}), (c_1, S_2^{c_1}), \cdots, (c_1, S_{|S_s|}^{c_1}), (c_2, S_1^{c_2}), (c_2, S_2^{c_2}), \cdots, (c_2, S_{|S_s|}^{c_2}), \cdots, \\ &(c_k, S_1^{c_k}), (c_k, S_2^{c_k}), \cdots, (c_k, S_{|S_s|}^{c_k}), (\emptyset, \emptyset), (c_{k+1}, S_1^{c_{k+1}}), (c_{k+1}, S_2^{c_{k+1}}), \cdots, (c_{k+1}, S_{|S_s|}^{c_{k+1}}), \cdots . \end{aligned}$$

The following example illustrates the above construction.

Example 3.4. Consider the same market $(\succ_S^*, \succ_C)$ as in Example 3.2:

$$\begin{aligned}
&\succ_{c_1}: \{s_1, s_2\}, \{s_1\}, \{s_2\}, \emptyset \\
&\succ_{s_1}^*: c_1, \emptyset \\
&\succ_{s_2}^*: c_1, \emptyset.
\end{aligned}$$

Note that in this market, the unique stable matching in this classical market is given by:

$$\mu_1 : \begin{pmatrix} c_1 \\ s_1, s_2 \end{pmatrix}$$

The associated market is given by $(\bar{\succ}_S, \bar{\succ}_C)$:

$$\begin{aligned}
&\succ_{c_1}: \{s_1, s_2\}, \{s_1\}, \{s_2\}, \emptyset \\
&\bar{\succ}_{s_1} : (c_1, \{s_1, s_2\}), (c_1, \{s_1\}), \emptyset \\
&\bar{\succ}_{s_2} : (c_1, \{s_1, s_2\}), (c_1, \{s_2\}), \emptyset.
\end{aligned}$$

It is easy to see that μ_1 is also a unique stable matching in the associated market $(\bar{\succ}_S, \bar{\succ}_C)$.

Kominers [38] show that the set of stable matchings in $(\succ_S^*, \succ_C)$ and $(\bar{\succ}_S, \bar{\succ}_C)$ are equivalent.

Theorem 3.7 (Kominers [38])**.** *A matching μ is stable in $(\succ_S^*, \succ_C)$ if and only if it is stable in $(\bar{\succ}_S, \bar{\succ}_C)$. Therefore, any classical model can be embedded into the model with preferences over colleagues.*

Note that the model of preferences over colleagues is strictly larger than the classical model. For instance, the set of stable matchings given in Example 3.2 cannot be represented by any classical model.

Flanagan [21] extends the result of Kominers [38] by examining the relationship between the model of "matching with contracts" and the model of preferences over colleagues. Here, we briefly explain the model of matching with contracts proposed by Hatfield and Milgrom [28], which is a generalization of the classical many-to-one matching problem. In a classical model, each student-college pair can be matched or unmatched. In the model of matching with contracts, for each student-college pair, there are finitely many contracts, and they can match through a contract. For example, a firm and a worker can match with "high salary" or "low salary". The outcome of this model is given by a set of contracts in which each student signs at most one contract. The stability concept can be extended into the model of matching with contracts. Note that the model of matching with contracts does not take externalities into account in that each agent does not care about the matching of the other agents.

Flanagan [21] shows that the model of matching with contracts can be embedded into the model with preferences over colleagues. This result generalizes Theorem 3.7. Moreover, he shows that the converse holds; that is, the model with preferences over colleagues can be embedded into the model of matching with contracts. Therefore, two models are equivalent when we focus on the set of stable matchings.

4. Matching and Coalition Formation

As we discussed in Section 3.2, the model of many-to-one matching with preferences over colleagues is related to a coalition formation model. Pycia [49] presents a model that includes an unrestricted coalition formation and a many-to-one matching of preferences over colleagues. He provides a domain of preferences in which a stable coalition structure always exists. The key assumption is a condition called pairwise alignment, which means that any two agents have the same preferences over proper coalitions that contain both of them. Typically, the literature on matching problems focuses on the restriction to individual preferences to guarantee the existence of a stable matching (for example, substitutability). The pairwise alignment condition restricts preferences between two agents rather than individual preferences. Therefore, the proof technique used in his paper is substantially different from the standard matching problems. In this section, we summarize Pycia [49]'s result.

4.1. General results

Let A be a finite set of agents. We assume that A is partitioned into a set of firms F and workers W. We say that a nonempty subset of A is a coalition. Coalition T is proper if $T \neq A$. Let $\mathcal{C} \subseteq 2^A$ be the set of possible coalitions that agents can form. We say that $\mathcal{C}$ is an *unrestricted coalition formation* when $\mathcal{C} = 2^A \setminus \{\emptyset\}$, and $\mathcal{C}$ is an *exact many-to-one matching* if each firm f has a capacity M_f such that $2 \leq M_f < |W|$ and $\mathcal{C}$ is given by:

$$\{\{f\} \cup S | f \in F, S \subseteq W \text{ with } |S| \leq M_f\} \cup \{\{w\} | w \in W\}.$$

Pycia [49] considers a more general model that includes these two models. For simplicity, we assume that $\mathcal{C}$ is unrestricted coalition formation or exact many-to-one matching throughout this section.

A coalition structure $\mu = \{T_1, T_2, \cdots, T_n\}$ is a partition of A such that $T_i \in \mathcal{C}$ for all $i = 1, 2, \cdots, n$. For example, letting $F = \{f_1, f_2\}$ and $W = \{w_1, w_2, w_3\}$, $\mu = \{\{f_1, w_1, w_2\}, \{f_2\}, \{w_3\}\}$ is a coalition structure in which f_1 hires workers w_1 and w_2, f_2 and w_3 are being unmatched. For each coalition structure μ and $a \in A$, let $\mu(a)$ denote the coalition in μ that contains a.

Each agent $a \in A$ has a complete and transitive preference relation $\succeq_a$ on $\mathcal{C}_a := \{T \in \mathcal{C} | a \in T\}$, which is the set of possible coalitions that contain a. Note that $\succeq_a$ may not be strict, and $T' \succeq_a T$ means that agent a prefers at least T' to T. We denote $T' \sim_a T$ when $T' \succeq_a T$ and $T \succeq_a T'$ hold. Let $\mathcal{R}_a$ denote the set of all preferences over $\mathcal{C}_a$, and $\mathcal{R} := \times_{a \in A} \mathcal{R}_a$ denote the set of all preference profiles. We say that $R \subseteq \mathcal{R}$ is a *domain.* A domain R represents a set of possible preference profiles that agents can have. We assume that R includes a rich variety of preference profiles. Formally, a domain R is *rich* if it satisfies the following conditions $R1$ and $R2$:

$R1$. For any $\succeq_A \in R$, $a \in A$ and any three different coalitions T_0, T and T_1, if $T_1 \succeq_a T_0$, then there exists $\succeq'_A \in R$ such that (i) $T_1 \succeq'_a T \succeq'_a T_0$, and (ii) for all $a' \in A$ and all $T'_1, T'_0 \in \mathcal{C}_{a'} \setminus \{T\}$,

$$T'_1 \succeq'_{a'} T'_0 \Leftrightarrow T'_1 \succeq_{a'} T'_0.$$

$R2$. (a) For any $\succeq_A \in R$ and any two different coalitions T and T_1, there exists $\succeq'_A \in R$ such that (i) $T_1 \succ'_a T$ for all $a \in T \cap T_1$, and (ii) for all $a' \in A$ and all $T'_1, T'_0 \in \mathcal{C}_{a'} \setminus \{T\}$,

$$T'_1 \succeq'_{a'} T'_0 \Leftrightarrow T'_1 \succeq_{a'} T'_0.$$

(b) For any $\succeq_A \in R$, any two agents $a, b \in A$ and any three different coalitions T_0, T and T_1, if $T_1 \sim_b T \succ_a T_2$, then then there exists $\succeq'_A \in R$ such that (i) $T_1 \succ'_b T \succ'_a T_2$, and (ii) for all $a' \in A$ and all $T'_1, T'_0 \in \mathcal{C}_{a'} \setminus \{T\}$,

$$T'_1 \succeq'_{a'} T'_0 \Leftrightarrow T'_1 \succeq_{a'} T'_0.$$

We next define stability for a coalition structure. Given a preference profile $\succeq_A \in R$, a coalition structure μ is blocked by a coalition $T \in \mathcal{C}$ if $T \succ_a \mu(a)$ for all $a \in T$, and is stable if it is not blocked. Note that in the model of preferences over colleagues where all agents have strict preferences, stable coalition structures are equivalent to stable matchings.

Pycia [49] characterizes a rich domain in which a stable coalition structure exists for all preference profiles. We introduce a key concept for this result. A preference profile $\succeq_A \in R$ is *pairwise-aligned* if for any two agents $a, b \in A$, and any proper coalitions T and T' that contain both a and b, we have

$$T' \succeq_a T \Leftrightarrow T' \succeq_b T.$$

We say that a domain R is pairwise-aligned if $\succeq_A$ is pairwise-aligned for every $\succeq_A \in R$.

The pairwise alignment assumption means that any two agents have the same preferences over proper coalitions that contain both of them. This assumption is restrictive. For example, in an exact many-to-one matching model without externalities, the pairwise alignment implies that each worker has the same ordering over firms. We will provide an example that satisfies the pairwise alignment and richness assumptions in the next section.

Pycia [49] shows that under a rich domain, there is a stable coalition structure for all preference profiles if and only if all preference profiles are pairwise-aligned.

Theorem 4.1 (Pycia [49]).*(1) Let R be a domain that satisfies $R1$. If R is pairwise-aligned, then there is a stable coalition structure for all preference profiles in R.*

(2) Let R be a domain that satisfies $R1$ and $R2$. If there is a stable coalition structure for all preference profiles in R, then R is pairwise-aligned.

From (1) and (2), we have the following result.

(3) Let R be a rich domain. There is a stable coalition structure for all preference profiles in R if and only if R is pairwise-aligned.

One might expect that if a preference profile $\succeq_A$ is pairwise-aligned, there is a stable coalition structure in $\succeq_A$; that is, $R1$ can be dropped. This is not true. In fact, Pycia [49] shows that even if a preference profile is pairwise-aligned, a stable coalition structure may not exist in an exact many-to-one matching. Therefore, $R1$ is crucial in obtaining Theorem 4.1-(1).

Pycia [49] gives a nonconstructive proof for the existence of a stable coalition structure. Here, we briefly explain the proof of Theorem 4.1-(1) given by Pycia [49]. He defines n-cycle ($n \geq 3$) to be a sequence of n proper coalitions $T_1, T_2, \cdots, T_n$ and n agents $a_1, a_2, \cdots, a_n$ such that

$$T_1 \preceq_{a_1} T_2 \preceq_{a_2} \cdots T_i \preceq_{a_i} T_{i+1} \cdots \preceq_{a_{n-1}} T_n \preceq_{a_n} T_1, \text{ and } T_j \prec_{a_j} T_{j+1} \text{ for some } j.$$

In the first step, Pycia [49] shows that given a preference profile, if there are no n-cycles for all $n \geq 3$, there exists a stable coalition structure. Next, by using the pairwise alignment assumption and $R1$, it can be shown that there are no 3-cycles for each preference profile in R, which in turn implies that there are no n-cycles for all $n \geq 3$. Therefore, under the pairwise alignment assumption and $R1$, a stable coalition structure exists for all preference profiles.

We also note that the structure of $\mathcal{C}$ is crucial to obtain Theorem 4.1-(1). For example, consider a roommate problem model which is defined as $\mathcal{C} := \{T \subseteq A | |T| \leq 2\}$. Then, every preference profile satisfies the pairwise alignment assumption, because for any two agents $a, b \in A$, $\{a, b\}$ is the only coalition that contains both a and b. Therefore, the set of all preference profiles satisfies the richness and pairwise alignment assumptions. However, it is a well-known fact that a stable coalition structure (stable matching) may not exist in a roommate problem.

To obtain Theorem 4.1-(2), we need to assume $R1$ and $R2$. Pycia [49] first shows that under $R2$, if there is a stable coalition structure for each preference profile in R, then there are no 3-cycles such that $T_1 \cap T_2 = \{a_1\}, T_2 \cap T_3 = \{a_2\}$ and $T_3 \cap T_1 = \{a_3\}$ for each preference profile in R. Next, under R_1, if there are no 3-cycles with the above property for each preference profile in R, then R is pairwise-aligned. Therefore, we can obtain Theorem 4.1-(2).

Teytelboym [64] extends the result of Pycia [49] into multilateral matching markets. His model is constituted from a finite set of agents and a finite set of contracts. Each contract involves at least two agents in a multilateral matching model while it involves exactly two agents in a usual matching market. Each agent has a preferences over sets of contracts involving him/her. In this setting, Teytelboym [64] shows that under a rich domain of preference profiles, the pairwise alignment condition is a necessary and sufficient condition for the existence of stable contract allocations in every preference profile.

4.2. Applications

In this section, we provide an example of a domain that satisfies the richness and pairwise alignment given by Pycia [49]. Here, we consider a situation where agents form coalitions to produce output (money) and then share the output within members of a coalition. We analyze sharing rules under which a stable coalition structure always exists. The formal description of the model is as bellow.

We say that $\Omega := \{(y_T)_{T\in\mathcal{C}} | y_T \in (0,\infty) \text{ for all } T \in \mathcal{C}\}$ is the set of possible output profiles. Given an output profile $(y_T)_{T\in\mathcal{C}}$, when coalition T is formed, T can earn y_T. Each agent a has a utility function $U_a : \mathbb{R} \to \mathbb{R}$ defined over money so that when agent a gets money s, a's utility is given by $U_a(s)$. We assume that U_a is strictly increasing and $U_a(0) = 0$ for all $a \in A$.

For each $a \in A$ and $T \in \mathcal{C}_a$, we say that $D_{a,T} : (0,\infty) \to \mathbb{R}$ is a sharing function. This means that if coalition T is formed and earns y, agent a's share ($a \in T$) is given by $D_{a,C}(y)$. Therefore, $\sum_{a\in T} D_{a,T}(y) \leq y$ is assumed. We say that a collection of sharing functions $D := (D_{a,T})_{a\in A, T\in\mathcal{C}}$ is a *sharing rule*. Then, for a given sharing rule D, we consider the following game.

- An output profile $(y_T)_{T\in\mathcal{C}} \in \Omega$ is realized and each agent learns the output profile.
- Agents form coalitions and get money according to the sharing rule D.

Note that given a sharing rule D, an output profile $(y_T)_{T\in\mathcal{C}}$ generates a preference profile $\succeq_A \in \mathcal{R}$ so that for each $a \in A$ and $T', T \in \mathcal{C}_a$,

$$T' \succeq_a T \Leftrightarrow U_a((D_{a,T'}(y_{T'})) \geq U_a((D_{a,T}(y_T)).$$

Let R^D denote the set of all preference profiles that are generated from all output profiles given a sharing rule D. The problem is to find a sharing rule D under which a stable coalition structure exists for all preference profiles in R^D. From Theorem 4.1, it is sufficient to find a sharing rule D in which R^D is rich and pairwise-aligned.

We say that a sharing rule D is *rich* if for all $a \in A$ and $T \in \mathcal{C}_a$, (i) $D_{a,T}$ is strictly increasing and continuous, and (ii) $\lim_{y\to\infty} D_{a,T}(y) = \infty$. Then, we can show that when D is rich, R^D is rich. We also say that a sharing rule D is *pairwise-aligned* if R^D is pairwise-aligned. Then, Theorem 4.1 directly implies the following result.

Theorem 4.2 (Pycia [49])**.** *Let D be a rich sharing rule. Then, there is a stable coalition structure for all preference profiles in R^D if and only if D is pairwise-aligned.*

We next provide examples of sharing rules that are rich and pairwise-aligned.

- *Equal sharing rule*: for all $a \in A$ and $T \in \mathcal{C}$, $D_{a,T}(y) = \frac{y}{|T|}$.
- *Nash bargaining rule*: Assume that U_a is twice differentiable and $U_a'' < 0$ for all $a \in A$. Then, any given $T \in \mathcal{C}$ and $y > 0$, the following maximization problem has a unique solution:

$$\max_{(s_a)_{a\in T}} \Pi_{a\in T} U_a(s_a) \text{ subject to } s_a \geq 0 \;\; \forall a \in T \text{ and } \sum_{a\in T} s_a \leq y. \qquad (1)$$

Let $(s_a^*)_{a\in C}$ be the solution of (1). Then, for each $a \in T$, $D_{a,T}(y)$ is given by s_a^*.

It is straightforward to see that the equal sharing rule and Nash bargaining rule satisfy richness. Moreover, these two rules satisfy the pairwise alignment condition. It is easy to see that the equal sharing rule satisfies the pairwise alignment. To see that Nash bargaining rule is pairwise-aligned, consider two agents $a, b \in A$ and coalitions T and T' that contain a and b. Let $(s_a^T)_{a\in T}$ and $(s_a^{T'})_{a\in T}$ be the solutions of (1). To obtain that Nash bargaining rule is pairwise-aligned, it is sufficient to show that $s_a^{T'} \geq s_a^T$ implies that $s_b^{T'} \geq s_b^T$. Define $x_a(s) =: \frac{U_a(s)}{U_a'(s)}$ and $x_b(s) =: \frac{U_b(s)}{U_b'(s)}$. Then, the first order condition of (1) implies that $x_a(s_a^T) = x_b(s_b^T)$ and $x_a(s_a^{T'}) = x_b(s_b^{T'})$. We also have that x_i is strictly increasing for $i = 1, 2$. Therefore, $s_a^{T'} \geq s_a^T$ implies that $x_a(s_a^{T'}) \geq x_a(s_a^T)$. This implies that $x_b(s_b^{T'}) \geq x_b(s_b^T)$ and hence $s_b^{T'} \geq s_b^T$. As a summary, we obtain the following result.

Theorem 4.3 (Pycia [49])**.** *Suppose that D is equal sharing rule or Nash bargaining rule. Then, there is a stable coalition structure for all preference profiles in R^D.*

Note that there are rich sharing rules that do not satisfy the pairwise alignment assumption. In fact, Pycia [49] shows that a sharing rule called Kalai-Smorodinsky rule does not satisfy the pairwise alignment assumption, while it satisfies the richness. So, under the Kalai-Smorodinsky rule, there does not exist a stable coalition structure for some preference profiles. Pycia [49] characterizes a class of rich sharing rules that satisfies the pairwise alignment assumption (See Pycia [49] for more details). By using Theorem 4.2, Teytelboym [64] analyzes a sharing rules on networks and characterizes a class of sharing rules under which strongly stable networks always exists.

5. Concluding Discussion

In this survey, we have summarized some of the main findings on the theory of matching markets with externalities. At first glance, this branch of two-sided matching theory seems very narrow with very few papers. We hope that by introducing several papers, that this is not the case. The theory in many-to-one matching markets has grown relative to the one-to-one matching markets, as there are types of externalities that can be distinguished in the many-to-one case that could not have been in the one-to-one case. Sections 3 and 4 summarized the findings for externalities among colleagues, while in Section 2 summarized findings regarding general externalities that can arise.

Another goal that we had set forth in this survey was to show that the theory of matching markets with externalities can give us some new insight into several existing models. In Section 3, the part regarding embedding is an example of such and potentially gives a connection to the model of Hatfield and Milgrom [28], which is widely popular.

Another topic in matching theory which is related to matching with externalities is dynamic matching. In a dynamic matching problem, a matching assignment needs to be determined for each time period $t = 1, 2, \cdots$. Even though at each time period t there are no externalities in the matching market, when considering the matching markets of all periods as a whole, a player $m \in M$ at time period t depends on the well-being of himself at a later date t' so that the player in time t' affects the same agent at an earlier time t. Thus, a dynamic matching problem can be seen as a special instance of a matching problem with externalities.

Damiano and Lam [12] is one of the earlier papers to explicitly define a model of dynamic matchings. Their model is essentially a two-sided matching market that is repeated many times. Kurino [39] modifies the model by allowing a matching of a period to depend on how players in the previous periods were matched. The papers that have followed have been based on real-life applications such as Pereyra [48], which is based on matching between teachers and schools in Mexico, and Kennes et al. [32], which is inspired by a day-care assignment mechanism in Denmark. The model by Kadam and Kotowski [30] is the closest theoretical paper to Kennes et al. [32] and also presents as an application an insight into the unraveling of matching markets as documented by Roth and Xing [60]. Doval [14] introduces stochastic processes in which men and women enter the market and proposes a stability concept in that model.

While there is a large number of applications of dynamic matching, applications of the theory of matching with externalities to real-life matching markets still is an open area of research despite the number of theoretical papers. To our knowledge, the paper by Baccara et al. [5], which empirically analyzes the assignment of rooms to faculty members with network externalities, comes closest to applications of some of the concepts in this survey, but their paper is technically built off a one-sided matching model in which one side does not have preferences. In matching problems with relatively small number of agents, we can expect that peer effects in some form may play a role in determining a suitable matching that could not have been taken into account under a model without externalities. At the present moment, we do not know of any papers that explicitly use the theory of two-sided matching with externalities outlined in this survey in analyzing real-life matching markets with external effects. However, in the near future we hope that some of the theoretical foundations outlined in this survey be used in the design of markets where externalities play a large role.

References

[1] H. Adachi: On a characterization of stable matchings. *Economics Letters*, **68** (2000), 43–49.

[2] B. Aldershof and O. M. Carducci: Stable matchings with couples. *Discrete Applied Mathematics*, **68** (1996), 203–207.

[3] A. Alkan: A class of multipartner matching markets with a strong lattice structure. *Economic Theory*, **19** (2002), 737–746.

[4] O. Ayğun and T. Sönmez: Matching with contracts: Comment. *American Economic Review*, **103(5)** (2013), 2050–2051.

[5] M. Baccara, A. İmorohoroğlu, A. J. Wilson, and L. Yariv: A field study on matching with network externalities. *American Economic Review*, **102** (2012), 1773–1804.

[6] K. Bando: Many-to-one matching markets with externalities among firms. *Journal of Mathematical Economics*, **48** (2012), 14–20.

[7] K. Bando: A modified deferred acceptance algorithm for many-to-one matching markets with externalities among firms. *Journal of Mathematical Economics*, **52** (2014), 173–181.

[8] S. Banerjee, H. Konishi, and T. Sönmez: Core in a simple coalition formation game. *Social Choice and Welfare*, **18** (2001), 135–153.

[9] P. Biró and F. Klijn: Matching with couples: A multidisciplinary survey. *International Game Theory Review*, **15** (2013), 1–18.

[10] A. Bogomolnaia and M. O. Jackson: The stability of hedonic coalition structures. *Games and Economic Behavior*, **38** (2002), 201–230.

[11] M. S.-Y. Chwe: Farsighted coalitional stability. *Journal of Economic Theory*, **63** (1994), 299–325.

[12] E. Damiano and R. Lam: Stability in dynamic matching markets. *Games and Economic Behavior*, **52** (2005), 34–53.

[13] V. Danilov, G. Koshevoy, and K. Murota: Discrete convexity and equilibria in economies with indivisible goods and money. *Mathematical Social Sciences*, **41** (2001), 251–273.

[14] L. Doval: A theory of stability in dynamic matching markets. working paper (2014).

[15] B. Dutta and J. Masso: Stability of matchings when individuals have preferences over colleagues. *Journal of Economic Theory*, **75** (1997), 464–475.

[16] F. Echenique: Finding all equilibria in games of strategic complements. *Journal of Economic Theory*, **135** (2007), 514–532.

[17] F. Echenique and J. Oviedo: Core many-to-one matchings by fixed-point methods. *Journal of Economic Theory*, **115** (2004), 358–376.

[18] F. Echenique and J. Oviedo: A theory of stability in many-to-many matching markets. *Theoretical Economics*, **1** (2006), 233–273.

[19] F. Echenique and M. B. Yenmez: A solution to matching with preferences over colleagues. *Games and Economic Behavior*, **59** (2007), 46–71.

[20] J. C. Fisher and I. E. Hafalir: Matching with aggregate externalities. mimeo (2015).

[21] F. Flanagan: Contracts vs. preferences over colleagues in matching. *International Journal of Game Theory*, **44** (2015), 209–223.

[22] T. Fleiner: A fixed-point approach to stable matchings and some applications. *Mathematics of Operations Research*, **28** (2003), 103–126.

[23] S. Fujishige and A. Tamura: A two-sided discrete-convex market with possibily bounded side payments: An approach by discrete convex analysis. *Mathematics of Operations Research*, **32** (2007), 136–155.

[24] S. Fujishige and Z. Yang: A note on Kelso and Crawford's gross substitutes condition. *Mathematics of Operations Research*, **28** (2003), 463–469.

[25] D. Gale and L. S. Shapley: College admissions and the stability of marriage. *American Mathematics Monthly*, **69** (1962), 9–15.

[26] D. Gusfield and R. W. Irving: *The Stable Marriage Problem: Structure and Algorithms* (The MIT Press, 1989).

[27] I. Hafalir: Stability of marriage with externalities. *International Journal of Game Theory*, **37** (2008), 353–369.

[28] J. W. Hatfield and P. Milgrom: Matching with contracts. *American Economic Review*, **95** (2005), 913–935.

[29] H. İnal: Core of coalition formation games and fixed-point methods. *Social Choice and Welfare*. forthcoming (2015).

[30] S. V. Kadam and M. H. Kotowski: Multi-period matching. working paper (2015).

[31] A. S. Kelso and V. P. Crawford: Job matching, coalition formation, and gross substitutes. *Econometrica*, **50** (1982), 1483–1504.

[32] J. Kennes, D. Monte, and N. Tumennasan: The day care assignment: A dynamic matching problem. *American Economic Journal: Microeconomics*, **6** (2014), 362–406.

[33] B. Klaus and F. Klijn: Stable matchings and preferences of couples. *Journal of Economic Theory*, **121** (2005), 75–106.

[34] B. Klaus, F. Klijn, and M. Walzl: Farsighted stability for roommate markets. *Journal of Public Economic Theory*, **13** (2011), 921–933.

[35] F. Kojima: Finding all stable matchings with couples. *Journal of Dynamics and Games*. forthcoming (2015).

[36] F. Kojima, P. A. Pathak, and A. E. Roth: Matching with couples: Stability and incentives in large markets. *Quarterly Journal of Economics*, **128** (2013), 1585–1632.

[37] F. Kojima, A. Tamura, and M. Yokoo: Designing matching mechanisms under constraints: An approach from discrete convex analysis. mimeo (2015).

[38] S. D. Kominers: Matching with preferences over colleagues solves classical matching. *Games and Economic Behavior*, **68** (2010), 773–780.

[39] M. Kurino: Credibility, efficiency, and stability: A theory of dynamic matching markets. mimeo (2009).

[40] S. Li: Competitive matching equilibrium and multiple principal-agent models. mimeo (1993).

[41] A. Mauleon, V. Vannetelbosch, and W. Vergote: Von Neumann - Morgenstern farsightedly stable sets in two-sided matching. *Theoretical Economics*, **6** (2011), 499–521.

[42] A. Mumcu and I. Saglam: Stable one-to-one matchings with externalities. *Mathematical Social Sciences*, **60** (2010), 154–159.

[43] K. Murota: Convexity and Stenitz's exchange property. *Advances in Mathematics*, **124** (1996), 272–311.

[44] K. Murota: Discrete convex analysis. *Mathematical Programming*, **83** (1998), 313–371.

[45] K. Murota and A. Shioura: M-convex function on generalized polymatroid. *Mathematics of Operations Research*, **24** (1999), 95–105.

[46] K. Murota and Y. Yokoi: On the lattice structure of stable allocations in a two-sided discrete-concave market. *Mathematics of Operations Research*, **40** (2015), 460–473.

[47] M. Ostrovsky: Stability in supply chain networks. *American Economic Review*, **98** (2008), 897–923.

[48] J. S. Pereyra: A dynamic school choice model. *Games and Economic Behavior*, **80** (2013), 100–114.

[49] M. Pycia: Stability and preference alignment in matching and coalition formation. *Econometrica*, **80** (2012), 323–363.

[50] M. Pycia and M. B. Yenmez: Matching with externalities. mimeo (2015).

[51] P. Revilla: Many-to-one matching when colleagues matter. mimeo (2007).

[52] A. E. Roth: The evolution of the labor market for medical interns and residents: A case study in game theory. *Journal of Political Economy*, **72** (1984a), 991–1016.

[53] A. E. Roth: Stability and polarization of interests in job matching. *Econometrica*, **52** (1984b), 47–58.

[54] A. E. Roth: The economist as engineer: Game theory, experimentation, and computation as tools for design economics. *Econometrica*, **70** (2002), 1341–1378.

[55] A. E. Roth: Repugnance as a constraint on markets. *Journal of Economic Perspectives*, **21** (2007), 37–58.

[56] A. E. Roth: Deferred acceptance algorithms: history, theory, practice, and open questions. *International Journal of Game Theory*, **36** (2008a), 537–569.

[57] A. E. Roth: What we have learned from market design. *The Economic Journal*, **118** (2008b), 285–310.

[58] A. E. Roth and E. Peranson The redesign of the matching market for american physicians: Some engineering aspects of economic design. *American Economic Review*, **89** (1999), 748–780.

[59] A. E. Roth and M. O. Sotomayor: *Two-sided Matching: A Study in Game-theoretic Modeling and Analysis* (Cambridge University Press, 1990).

[60] A. E. Roth and X. Xing: Jumping the gun: Imperfections and institutions related to the timing of transactions. *American Economic Review*, **84** (1994), 992–1044.

[61] H. Sasaki and M. Toda: Marriage problem reconsidered – externalities and stability. (University of Rochester, Department of Economics, 1986).

[62] H. Sasaki and M. Toda: Two-sided matching problems with externalities. *Journal of Economic Theory*, **70** (1996), 93–108.

[63] A. Shioura and A. Tamura: Gross substitutes condition and discrete concavity for multi-unit valuations: A survey. *Journal of the Operations Research Society of Japan*, **58** (2015), 61–103.

[64] A. Teytelboym: Strong stability in networks and matching markets with contracts. mimeo (2013).

Keisuke Bando
Department of Social Engineering
Graduate School of Decision Science and Technology
Tokyo Institute of Technology
2-12-1 Ookayama, Meguro-ku
Tokyo 152-8552, Japan
E-mail: `bando.k.aa@m.titech.ac.jp`