ORIGINAL PAPER

Hideo Konishi · M. Utku Ünver

Games of capacity manipulation in hospital-intern markets

Received: 17 January 2002 / Accepted: 9 June 2004 / Published online: 10 June 2006 © Springer-Verlag 2006

Abstract In this paper, we analyze capacity manipulation games in hospital-intern markets inspired by the real-life entry-level labor markets for young physicians who seek residencies at hospitals. In a hospital-intern market, the matching is determined by a centralized clearinghouse using the preferences revealed by interns and hospitals and the number of vacant positions revealed by hospitals. We consider a model in which preferences of hospitals and interns are common knowledge. Hospitals play a capacity-reporting game. We analyze the equilibria of the game-form under the two most widely used matching rules: hospital-optimal and intern-optimal stable rules. We show that (i) there may not be a pure strategy equilibrium in general; and (ii) when a pure strategy equilibrium exists, every hospital weakly prefers this equilibrium outcome to the outcome of any larger capacity profile. Finally, we present conditions on preferences to guarantee the existence of pure strategy equilibria.

1 Introduction

This paper examines capacity manipulation games in hospital-intern markets. Hospital-intern markets are entry-level labor markets, where new physicians

H. Konishi Department of Economics, Boston College, Chestnut Hill, MA 02467, USA E-mail: hideo.konishi@bc.edu http://fmwww.bc.edu/EC-V/Konishi.fac.html

M. U. Ünver (☒)
Department of Economics, The University of Pittsburgh, 4506 W. W Posvar Hall, 230 S. Bouquet St.,
Pittsburgh, PA 15260, USA
E-mail: uunver@pitt.edu
http://www.pitt.edu/~uunver/index.html

seek positions in hospitals as residents. Motivated by this real-life problem, Roth (1984, 1991) demonstrated that these markets can be modeled using the two-sided matching model of Gale and Shapley (1962).¹

A hospital-intern market consists of a finite set of hospitals each with a finite quota of positions, a finite set of interns, strict preferences of hospitals over groups of interns and strict preferences of interns over hospitals. Each intern can work at one hospital, while a hospital can hire as many interns as its capacity permits.² A *matching* assigns an intern at most to one hospital and assigns a hospital to a group of interns not larger than its quota. A *stable* matching is defined as a matching where (i) no hospital prefers keeping a position vacant to filling it with one of its assignments, (ii) no intern prefers being unemployed to her assignment, and (iii) there is no unmatched hospital-intern pair such that the intern prefers the hospital to her assignment and the hospital prefers the intern to one of its assignments or keeping a vacant position. A *matching rule* is a systematic procedure that assigns a matching to each hospital—intern market. In the game-theoretical model, hospitals report their preferences and capacities, and interns report their preferences to the authority. The authority matches interns to hospitals via a matching rule using the revealed information.

The two stable matching rules, known as *hospital-optimal* and *intern-optimal* stable matching rules, are being used in the United States and in different regions of the United Kingdom to match medical interns to hospitals (Roth 1984, 1991; Roth and Peranson 1999).³ Thus, the game-theoretical model is indeed a real-world practice. Hence, it is important to know the normative and strategic performances of these two matching rules.

Roth (1982) and Dubins and Freedman (1981) showed that truthful preference revelation is a dominant strategy for interns under the intern-optimal stable rule. This result is particularly important in applications where hospitals cannot misrepresent their preferences. The game-theoretical hospital-intern markets can be also used to model various real-life school admission markets. For example, in Turkish college admissions market, students take standardized exams on various subjects. In this market, preferences of colleges over students are determined using these test scores through a publicly known formula (Balinski and Sönmez 1999). Another

¹ See Roth and Sotomayor (1990) for an extensive game-theoretical treatment of two-sided matching problems.

² A hospital-intern market is also known as a *many-to-one two-sided matching* market or a *college admissions* market in the literature.

³ After each physician graduates from medical school, she is required to work in a residency position in a qualified hospital. From 1951 to 1997 the National Residency Matching Program (NRMP) used various mechanisms based on the hospital-optimal stable rule to match hospitals and interns in the United States (Roth 1984). Then, Roth and Peranson (1999) designed a mechanism based on the intern-optimal stable rule and Roth and Vande Vate (1990) study. This matching rule is now being used in the United States market (Roth and Peranson (1999)). Also, Roth (1991) observed that unstable rules failed in the field to obtain stable matches in Britain. Some were replaced by the hospital-optimal and intern-optimal rules. Matching rules based on the hospital-optimal and intern-optimal stable rules have been used in regions of England, Scotland, and Wales. For the centralized match, applicants submit an ordered list of hospital names and hospitals submit the number of vacant residencies and an ordered list of applicant names to the NRMP. Then, the NRMP uses this information to match hospitals to intern candidates using the Roth and Peranson algorithm. In our model, the Roth and Peranson matching rule is equivalent to the intern-optimal stable rule.

⁴ Hospitals represent schools and interns represent students in these applications.

example is elementary and secondary school choice in the USA. In many school districts, priority of each student at each school is determined by the school district (Abdulkadiroğlu and Sönmez 2003). Schools have no control over these priorities and they can only manipulate their capacities. Students, on the other hand, do not have any incentives to misreport their preferences under the intern-optimal stable matching rule.

In this paper, we consider strategic capacity manipulation by hospitals. Sönmez (1997b) analyzed hospitals' incentives to underreport their capacities in a hospital-intern market. He showed that there is no stable matching rule that is immune to manipulation via underreporting capacities under the standard responsive (Roth 1985) and strict preferences. Motivated by this observations, we analyze a capacity-reporting game where preferences of agents are common knowledge. In this game, hospitals are the active players and reveal their capacities. A matching rule is fixed and the outcome of the induced market is found using the revealed capacities. We analyze the properties of equilibria of this game under the hospital-optimal and intern-optimal stable matching rules. 6

Using responsive and strict preferences, we show that there may not be any pure strategy equilibrium of the capacity-reporting game. When a pure strategy equilibrium exists, we prove that every hospital weakly prefers this equilibrium outcome to the outcome of any capacity profile that is at least as large as the equilibrium in every component. Finally, we consider two preference restrictions each of which guarantees the existence of a pure strategy equilibrium. First, if every hospital always prefers a larger group of acceptable interns to a smaller group, then truthful capacity revelation is a weakly dominant strategy under the intern-optimal stable rule. Moreover, there exists a pure strategy equilibrium (not necessarily truthful capacity revelation) under the hospital-optimal rule. Second, if hospitals have exactly the same preferences over the set of individual interns or interns have exactly same preferences over the set of hospitals, then truthful capacity revelation becomes a weakly dominant strategy.

The paper is organized as follows. The rest of this section provides a brief discussion of the literature on two-sided matching games. In Sect. 2, we present the game-theoretical model for hospital-intern markets, introduce relevant concepts for our analysis, and define our notion of capacity-reporting games. In Sect. 3, we show that there may not be a pure strategy equilibrium in capacity-reporting games. In Sect. 4, we discuss the welfare implications of capacity manipulation on equilibria. We also present a useful example for understanding characteristics of the set of equilibria. In Sect. 5, we introduce two types of preference restrictions that guarantee nonemptiness of the set of pure strategy equilibria in capacity-reporting games. Section 6 concludes the paper.

⁵ It is informationally demanding to assume that every agent's preferences are common knowledge. One may want to analyze a capacity-revelation game where preferences are not common knowledge. However, there are multiple ways to formulate information structures over preferences. In this paper, we adopt a common knowledge information structure in order to conduct a benchmark study on games of capacity manipulation.

⁶ It is also worthwhile to know the strategic performance of the other commonly used matching rule, hospital-optimal stable rule, under capacity manipulation.

1.1 A brief literature review

There is an extensive literature on stable matching rules and their properties in two-sided matching markets. Strategic actions by agents are one of the central interests in the literature. Under what conditions and how can interns or hospitals benefit from choosing their actions strategically? Such questions have direct implications for real-life policy.

Incentives for preference manipulation have been widely studied in the literature. Dubins and Freedman (1981) show that the hospital-optimal stable matching rule is not immune to preference manipulation in the marriage framework (i.e., for the case when a hospital can hire only one intern). Roth (1982) shows that there is no stable matching rule immune to preference manipulation. Alcalde and Barberà (1994) generalize Roth's negative result to any individually rational and Paretoefficient matching rule. Roth and Rothblum (1999) study preference manipulation by interns with incomplete information about others' preferences under the hospital-optimal stable rule. Equilibrium analysis of preference manipulation games has been studied (generally in the marriage framework) as well. Gale and Sotomayor (1985a,b) study preference manipulation games under the hospital-optimal stable rule in marriage markets. Alcalde (1996), Ma (1995, 1997) and Shin and Suh (1996) characterize the equilibria of the preference reporting games induced by stable solutions in marriage markets. Sönmez (1997a) analyzes the equilibrium outcomes of the preference reporting games induced by Pareto-efficient and individually rational solutions in the context of marriage problems.⁸

Two additional strategic issues have been examined in hospital-intern markets. One question is whether a hospital and an intern can benefit from making an early contract prior to the centralized match. Sönmez (1999) shows that no stable matching rule is immune to manipulation through early contracting. This is in the same spirit as the model of unraveling in hospital-intern markets by Roth and Xing (1994). Another question is whether a hospital can benefit from underreporting its capacity. Sönmez (1997b) shows that there is no stable matching rule that is immune to manipulation via underreporting capacities.

In this paper, we further pursue Sönmez (1997b) research program on capacity manipulation. We inspect the pure strategy equilibria of the capacity-reporting games in the spirit of Gale and Sotomayor (1985a,b).

2 Hospital-intern markets and capacity-reporting games

First, we define hospital-intern markets (Gale and Shapley 1962). A hospital-intern market is a quadruple (H, I, q, R) where

- 1. $H = \{h_1, h_2, \dots, h_m\}$ is a set of hospitals, 2. $I = \{i_1, i_2, \dots, i_n\}$ is a set of interns,
 - ⁷ See Roth and Sotomayor (1990) to have an account of the literature prior to 1990.

⁸ Kara and Sönmez (1996, 1997) analyze implementable matching rules in the marriage and hospital-intern frameworks, respectively. Alcalde and Romero-Medina (2000) analyze simple mechanisms that implement the core of hospital-intern markets.

⁹ This is in the same spirit as Postlewaite (1979) which studies manipulations via endowments in exchange economies. See also Sertel (1994) and Thomson (1987a,b, 1995).

- 3. $q = (q_{h_1}, \dots, q_{h_m})$ is a list of hospital capacities where q_h is the capacity of hospital h,
- 4. $R = (R_{h_1}, \ldots, R_{h_m}, R_{i_1}, \ldots, R_{i_n})$ is a profile of preference relations where R_h is the preference relation of hospital h and R_i is the preference relation of intern i.

For any $h \in H$, R_h is a binary preference relation that is a *linear order* (or strict preference)¹⁰ on groups of interns and staying unmatched. Formally, R_h is defined on the members of the set $X_h^I \equiv X^I = 2^I$. The preferences of hospitals are *responsive* (Roth 1985). For any $h \in H$, its preference relation R_h is responsive if and only if (i) for any $i, j \in I$ and any $J \subseteq I \setminus \{i, j\}$, we have $(J \cup \{i\}) R_h (J \cup \{j\}) \iff \{i\} R_h \{j\}$, and (ii) for any $i \in I$ and $J \subseteq I \setminus \{i\}$, we have $(J \cup \{i\}) R_h J \iff \{i\} R_h \emptyset$. Let \mathcal{R}_h be the class of all such responsive preference relations.

For any $i \in I$, R_i is a binary preference relation that is a *linear order* on individual hospitals and staying unmatched. Formally, R_i is defined on the members of set $X_i^H \equiv X^H = \{\{h_1\}, \{h_2\}, \dots, \{h_m\}, \emptyset\}$. Let \mathcal{R}_i be the class of all such preference relations. Let $\mathcal{R} = \prod_{h \in H} \mathcal{R}_h \times \prod_{i \in I} \mathcal{R}_i$.

Let P_h be the strict relation induced by R_h for any $h \in H$. Since R_h is a linear order, $LP_hL' \iff LR_hL'$ and $L \neq L'$ for any $L, L' \subseteq I$. Similarly, let P_i denote the strict relation induced by R_i for any $i \in I$. Since R_i is a linear order, $\{h\}P_i\{h'\} \iff \{h\}R_i\{h'\}$ and $\{h\} \neq \{h'\}$ for any $h, h' \in H$.

Let $N = H \cup I$ be the set of agents. Each hospital h can hire at most q_h interns and each intern i can work at most for one hospital. Hospital h is acceptable for intern i if and only if $\{h\}P_i\emptyset$, i.e., intern i prefers working in the hospital h to staying unmatched. An intern $i \in I$ is acceptable for hospital h if and only if $\{i\}P_h\emptyset$.

Note that for any agent $v \in N$, $\mu_v = \emptyset$ means that this agent is unmatched by matching μ . Let $\mathcal{M}(H, I, q, R)$ be the set of matchings of market (H, I, q, R).

A matching $\mu \in \mathcal{M}(H, I, q, R)$ is *stable* if and only if (i) there exists no hospital h and intern i such that for some group of interns $J \subseteq \mu_h$ with $|J| < q_h$, we have $(J \cup \{i\}) P_h \mu_h$ and $\{h\} P_i \mu_i$ (pair (h, i) is said to be a "blocking pair"), and (ii) there exists no agent v such that for some agent $v \in \mu_v$ we have $\emptyset P_v v$ (agent v is said to be a "blocking agent"). Let S(H, I, q, R) denote the set of stable matchings of market (H, I, q, R). This set is non-empty (Gale and Shapley 1962). In many cases it is multivalued.

¹⁰ A binary relation R_v on the set X is a *linear order* if

^{1.} For every $x \in X$, we have $x R_v x$.

^{2.} For every $x, x', x'' \in X$ such that $x R_v x'$ and $x' R_v x''$, we have $x R_v x''$.

^{3.} For every $x, x' \in X$ such that $x \neq x'$, we have $x R_v x'$ or $x' R_v x$ but not both.

¹¹ For any agent $v \in N$, $\mu(v)$ refers to the allocation of agent v in matching μ . For purposes of notation, we will denote $\mu(v)$ by μ_v .

We will sometimes say that "agent v prefers matching μ to matching μ " to describe $\mu_v P_v \mu_v'$.

2.1 Hospital-optimal and Intern-optimal stable matchings

In each market (H, I, q, R), there exists a stable matching that is weakly preferred to any other stable matching by every hospital (Gale and Shapley 1962). We refer to this matching as the *hospital-optimal stable matching* and denote it by $\mu^{\rm H}(H, I, q, R)$. Every intern weakly prefers any stable matching to the hospital-optimal stable matching (Roth and Sotomayor 1990). Similarly, there exists a stable matching that is weakly preferred to any other stable matching by every intern. We refer to this matching as the *intern-optimal stable matching* and denote it by $\mu^{\rm I}(H, I, q, R)$. Every hospital weakly prefers any stable matching to the intern-optimal stable matching.

The hospital-proposing deferred acceptance algorithm (Gale and Shapley 1962) can be used to find the hospital-optimal stable matching. This algorithm can be outlined as follows, when hospitals report capacity profile q':

Step 1 Each hospital h makes as many as q'_h offers to its best acceptable interns in R_h . Each intern i who receives at least one offer holds the best acceptable offer and rejects the rest. The algorithm terminates if no offer is rejected. Otherwise, hospitals skip to the next step.

:

Step t Each hospital h that has fewer than q'_h held offers, and still has acceptable interns to whom it has not proposed yet, makes offers to as many acceptable interns as possible in order to complete the number of offers to q'_h . Each intern i holds the best acceptable offer among the ones she receives at this step and the one she was holding from the previous step. She rejects the rest. The algorithm terminates if no offer is rejected by any intern. Otherwise, hospitals skip to the next step.

When the algorithm terminates, the tentatively held offers are realized as assignments.

The intern-proposing deferred acceptance algorithm of Gale and Shapley can be used to find the intern-optimal stable matching. This algorithm can be outlined, when hospitals report capacity q':

Step 1 Each intern i makes an offer to her best acceptable hospital in R_i . Each hospital h that receives one or more offers holds as many as q'_h best acceptable offers and rejects the rest. The algorithm terminates if no offer is rejected. Otherwise, interns skip to the next step.

: :

Step t Each intern i whose offer was rejected at step t-1 proposes to the best acceptable hospital to which she has not proposed before. Each hospital h holds as many as q'_h best acceptable offers among the ones it receives at this step and the ones it was holding from the previous step. It rejects the rest. The algorithm terminates if no offer is rejected by any hospital. Otherwise, interns skip to the next step.

When the algorithm terminates, the tentatively held offers are realized as assignments.

2.2 Capacity-reporting games

Let (H, I, q, R) be a hospital-intern market. Since we are interested in capacity manipulation, we will fix H, I, and R throughout the paper. We need to define a capacity-reporting game. A *matching rule* φ is a systematic procedure that assigns a matching for each hospital-intern market. We consider only different capacity possibilities. An admissible capacity of each hospital h is a nonnegative integer that is no greater than the true capacity. Thus, h can report $q'_h \in Q_h = \{0, 1, \ldots, q_h\}$. Define the set of admissible capacity profiles as $Q = \Pi_{h \in H} Q_h$. Let φ^H be the matching rule such that for each $q' \in Q$, $\varphi^H(q') = \mu^H(H, I, q', R)$. We refer to this rule as the *hospital-optimal stable matching rule*. Let φ^I be the matching rule such that for each $q' \in Q$, $\varphi^I(q') = \mu^I(H, I, q', R)$. We refer to this rule as the *intern-optimal stable matching rule*. As Gale and Sotomayor (1985b) focused in a preference manipulation game, we focus on these two matching rules. These rules have been widely used in real markets. φ^{I}

In the 'capacity-reporting game,' each hospital h reports an admissible capacity q_h' , and a matching rule φ is used to find a matching in the induced matching market. The preferences of agents are a common knowledge. Interns are passive players in this setting. Hospitals are active players, and hospital h's strategy space is Q_h , that is, the space of admissible capacities. Hospital h's preferences over reported capacities are represented as a binary relation \succsim_h^φ over Q such that $q' \succsim_h^\varphi q''$ if and only if $\varphi_h(q')R_h\varphi_h(q'')$. A capacity-reporting game under matching rule φ is described by a strategic form game $(H, (Q_h, \succsim_h^\varphi)_{h \in H})$. A pure strategy (Nash) equilibrium of $(H, (Q_h, \succsim_h^\varphi)_{h \in H})$ is a strategy profile $q' \in Q$ such that $q' \succsim_h^\varphi (q_h'', q_{-h}')$ for any $h \in H$ and any $q_h'' \in Q_h$. Note that $(q_h'', q_{-h}') \in Q$ is a strategy profile obtained by replacing component q_h' of profile q' by q_h'' . ¹⁵

3 Nonexistence of pure strategy equilibria

We will analyze pure strategy equilibria of the capacity-reporting game under the hospital-optimal and intern-optimal stable matching rules, φ^H and φ^I , respectively.

Note that we could permit upward manipulation of capacities as well. In that case, we should define the preferences of hospitals responsive relative to their capacity. Since hospitals cannot honor extra jobs over their capacity, we can simply assume that for any hospital being matched to any group of interns that is larger than its true capacity is worse than being unmatched (this preference restriction can be weakened much further). Under such a preference restriction, it can be shown that upward manipulation would be weakly dominated by truthful capacity revelation under the intern-optimal and hospital-optimal stable rules.

¹⁴ In the British hospital-intern markets, algorithms based on both intern-optimal and hospital-optimal stable rules have been used to match interns to hospitals in different regions (Roth 1991). In the United States hospital-intern market, an algorithm based on the hospital-optimal stable rule have been used until recently (Roth 1984). They started to use an algorithm based on the intern-optimal stable rule in 1998 (Roth and Peranson 1999). In Turkish college admissions market, the hospital-optimal stable rule is used to allocate students in college departments (Balinski and Sönmez 1999). Recently, many entry-level professional specialty markets started using variants of the intern-optimal stable rule for their centralized match. These markets include clinical psychology internships in the USA and Canada, article positions with law firms in Canada, hospital-intern markets in Scotland (Irving 1998), and various medical specialty markets.

Similarly, we will use the notation (q'_G, q_{-G}) to denote the capacity profile $(q'_G, q_{H\backslash G})$ for any $G \subset H$ such that |G| > 1.

The first two theorems show that pure strategy equilibria may not exist. The proofs are based on two examples in which no pure strategy profile is an equilibrium of the capacity-reporting games. For notational convenience, define hospital h's best-response correspondence under matching rule φ by β_h^{φ} : $\Pi_{h' \in H \setminus \{h\}} Q_{h'} \to Q_h$ such that $\beta_h^{\varphi}(q'_{-h}) = \{\tilde{q}_h \in Q_h : (\tilde{q}_h, q'_{-h}) \succsim_h^{\varphi} (q''_h, q'_{-h}) \text{ for any } q''_h \in Q_h\}.$

Theorem 1 The capacity-reporting game under φ^H may not have a pure strategy equilibrium.

Proof We prove the proposition with an example. Consider the following market (H, I, q, R) with $H = \{h_1, h_2\}$, $I = \{i_1, i_2, i_3, i_4, i_5\}$, $q_1 = 3$, $q_2 = 3$, hospitals' preferences satisfying 16

$$\{i_1\}P_{h_1}\{i_2\}P_{h_1}\{i_4\}P_{h_1}\{i_3\}P_{h_1}\{i_5\}P_{h_1}\emptyset$$
 with $\{i_3,i_4\}P_{h_1}\{i_1\}$
 $\{i_4\}P_{h_2}\{i_5\}P_{h_2}\{i_1\}P_{h_2}\{i_3\}P_{h_2}\{i_2\}P_{h_2}\emptyset$ with $\{i_4\}P_{h_2}\{i_1,i_3\}$ and $\{i_2,i_3\}P_{h_2}\{i_5\}$

and interns' preferences are given by

$$\begin{aligned} \{h_2\} P_{i_1}\{h_1\} P_{i_1} \emptyset \\ \{h_2\} P_{i_2}\{h_1\} P_{i_2} \emptyset \\ \{h_2\} P_{i_3}\{h_1\} P_{i_3} \emptyset \\ \{h_1\} P_{i_4}\{h_2\} P_{i_4} \emptyset \\ \{h_1\} P_{i_5}\{h_2\} P_{i_5} \emptyset \end{aligned}$$

We give the outcome allocations of the game for each pure strategy $q' \in Q$:

According to the specifications of hospitals' preferences, their best-response correspondences are:

$$\beta_1^{\varphi^{H}}(1) = \{3\}, \quad \beta_1^{\varphi^{H}}(2) = \{3\}, \quad \beta_1^{\varphi^{H}}(3) = \{2\};$$

 $\beta_2^{\varphi^{H}}(1) = \{3\}, \quad \beta_2^{\varphi^{H}}(2) = \{2\}, \quad \beta_2^{\varphi^{H}}(3) = \{3\}.$

Therefore, this game has no pure strategy equilibrium.

¹⁶ We also impose responsiveness on hospitals' preferences. It is easy to see that we can construct such preferences. The same comment applies to the example in the proof of Theorem 2 and Example 1.

We discuss the above example in detail in order to understand why no pure strategy equilibrium exists. These observations will help us to proceed especially in identifying preference restrictions that guarantee the existence of a pure strategy equilibrium (in Sect. 5).

First, we describe why and how a capacity manipulation improves a deviator's payoff. At $(q'_1, q'_2) = (3, 3)$, hospitals h_1 and h_2 get assignments $\{i_4, i_5\}$ and $\{i_1, i_2, i_3\}$, respectively. However, both hospitals are matched with relatively undesirable interns, although each intern is matched with her favorite hospital under this matching. If interns i_1 or i_2 could be swapped with interns i_4 or i_5 , then both hospitals would be better off. Capacity manipulation by hospital h_1 makes such a swap feasible. To see this, we can describe how the hospital-optimal stable rule makes its assignment for $(q'_1, q'_2) = (3, 3)$ by using the hospital-proposing deferred acceptance algorithm.

Step 1 Hospitals h_1 and h_2 make offers to $\{i_1, i_2, i_4\}$ and $\{i_4, i_5, i_1\}$, respectively. Since i_1 and i_4 got two offers, they choose h_2 and h_1 , respectively. Thus, the tentative allocation is

$$\begin{pmatrix} h_1 & h_2 \\ \{i_2, i_4\} & \{i_5, i_1\} \end{pmatrix}$$
.

Step 2 Hospitals h_1 and h_2 make offers to i_3 . Then i_3 chooses h_2 . Thus, the tentative allocation is

$$\begin{pmatrix} h_1 & h_2 \\ \{i_2, i_4\} & \{i_5, i_1, i_3\} \end{pmatrix}$$
.

Step 3 Hospital h_1 makes an offer to i_5 , and she accepts h_1 by rejecting h_2 .

$$\begin{pmatrix} h_1 & h_2 \\ \{i_2, i_4, i_5\} & \{i_1, i_3\} \end{pmatrix}$$
.

Step 4 Hospital h_2 makes an offer to i_2 , and she accepts h_2 by rejecting h_1 .

$$\begin{pmatrix} h_1 & h_2 \\ \{i_4, i_5\} & \{i_1, i_3, i_2\} \end{pmatrix}$$
.

Step 5 Hospital h_1 has no other intern to make an offer to, so the algorithm terminates.

At steps 3 and 4, having a vacancy, hospital h_1 makes an offer to i_5 and ends up stealing her from h_2 . Because of this newly created vacancy, hospital h_2 makes an offer to i_2 and ends up stealing her from hospital h_1 . Unfortunately, hospital h_1 prefers i_2 to i_5 , yet hospital h_2 prefers i_5 to i_2 . If hospital h_1 had not got an extra vacancy at the tentative matching at the end of step 2, then a welfare decreasing swapping (for hospitals) would not have happened. This is exactly why h_1 has an incentive to reduce q_1' to 2 at $(q_1', q_2') = (3, 3)$. Having larger capacities means that they have the chance to make offers to many interns, which implies that interns

¹⁷ Note that interns i_2 and i_5 are better off by this swapping.

have more choices. This makes interns better off, but not hospitals. ¹⁸ By reducing capacities, hospitals may be able to keep their favorite interns who prefer other hospitals. Whenever a hospital h_i has an incentive to reduce its capacity, h_i is making a deal with another hospital h_j : by not stealing a favorite intern from h_j , h_i closes off h_i 's opportunity to steal h_i 's favorite intern.

Second, note that there is a best response cycle among strategy profiles: $(3, 3) \rightarrow (2, 3) \rightarrow (2, 2) \rightarrow (3, 2) \rightarrow (3, 3) \rightarrow \dots$ As we have seen in the previous paragraph, at $(q_1', q_2') = (3, 3)$, hospital h_1 has an incentive to reduce q_1' to 2. Given this, hospital h_2 now has an incentive to reduce q_2' to 2. By so doing, hospital h_2 loses interns i_1 and i_3 , but it can get its most preferable intern i_4 . As a result, at $(q_1', q_2') = (2, 2)$, both hospitals get the two most preferable interns each. However, once q_2' is reduced to 2, hospital h_1 has an incentive to increase its capacity to 3. Hospital h_1 did not mind reducing its capacity from 3 to 2 at $(q_1', q_2') = (3, 3)$, only because filling all three positions was never possible as long as $q_2' = 3$. Finally, once q_1' is raised to 3, hospital h_2 also has an incentive to raise q_2' to 3. Hospital h_2 reduced its capacity from 3 to 2 at $(q_1', q_2') = (2, 3)$, only because it wanted to get intern i_4 . However, after q_1' is raised to 3, it became impossible for hospital h_2 to get intern i_4 anyway. This is why we have a cycle à la matching pennies game.

Third, related to the second point, hospitals have non-monotonic best-response correspondences.¹⁹ This creates a best response cycle. Such a non-monotonicity is not special to this example. With larger reported capacities, hospitals tend to get less preferable interns unless hospitals' preferences are positively correlated with interns' preferences. Thus, hospitals may reduce their reported capacities in order to prevent welfare decreasing swapping for themselves. However, under smaller reported capacities, a hospital may have an incentive to raise its reported capacity, since there may be acceptable unmatched interns left in the market.

We can also observe similar properties in games under the intern-optimal stable matching rule. As a result, we have a similar proposition for the capacity-reporting game under φ^{I} .

Theorem 2 The capacity-reporting game under φ^I may not have a pure strategy equilibrium.

Proof We prove the proposition with an example. Consider the following market (H, I, q, R) with $H = \{h_1, h_2\}$, $I = \{i_1, i_2, i_3, i_4, i_5\}$, $q_1 = 3$, $q_2 = 2$, hospitals' preferences satisfying

$$\{i_1\}P_{h_1}\{i_2\}P_{h_1}\{i_3\}P_{h_1}\{i_4\}P_{h_1}\{i_5\}P_{h_1}\emptyset$$
 with $\{i_2\}P_{h_1}\{i_4,i_5\}$ and $\{i_3,i_4\}P_{h_1}\{i_1\}$
 $\{i_3\}P_{h_2}\{i_4\}P_{h_2}\{i_1\}P_{h_2}\{i_5\}P_{h_2}\{i_2\}P_{h_2}\emptyset$ with $\{i_3\}P_{h_2}\{i_1,i_4\}$,

and interns' preferences are given by

¹⁸ If hospitals have no capacity limitation, then there is a unique stable matching. This is simply because every intern receives offers from all hospitals that find her acceptable. Obviously, she chooses her favorite hospital. See Sönmez (1996).

¹⁹ If such monotonicity is shown for all hospitals, then the game becomes an ordinary supermodular game (see Milgrom and Shannon 1994), and it has a pure strategy equilibrium. Unfortunately, it is hard to satisfy such a property in our game.

$$\{h_2\}P_{i_1}\{h_1\}P_{i_1}\emptyset$$

 $\{h_2\}P_{i_2}\{h_1\}P_{i_2}\emptyset$
 $\{h_1\}P_{i_3}\{h_2\}P_{i_3}\emptyset$
 $\{h_1\}P_{i_4}\{h_2\}P_{i_4}\emptyset$
 $\{h_1\}P_{i_5}\{h_2\}P_{i_5}\emptyset$.

We give the outcome allocations of the game for each pure strategy $q' \in Q$:

		q_2'	2
	1	$\begin{pmatrix} h_1 & h_2 & \emptyset \\ \{i_1\} & \{i_3\} & i_2, i_4, i_5 \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 & \emptyset \\ \{i_1\} & \{i_3, i_4\} & i_2, i_5 \end{pmatrix}$
q_1'	2	$\begin{pmatrix} h_1 & h_2 & \emptyset \\ \{i_1, i_2\} & \{i_3\} & i_4, i_5 \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 & \emptyset \\ \{i_2, i_3\} & \{i_1, i_4\} & i_5 \end{pmatrix}$
	3	$\begin{pmatrix} h_1 & h_2 & \emptyset \\ \{i_2, i_3, i_4\} & \{i_1\} & i_5 \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 \\ \{i_3, i_4, i_5\} & \{i_1, i_2\} \end{pmatrix}$

According to the specification of hospitals' preferences, their best-response correspondences are:²⁰

$$\begin{split} \beta_1^{\varphi^{\rm I}}(1) &= \{3\}, \quad \beta_1^{\varphi^{\rm I}}(2) = \{2\}; \\ \beta_2^{\varphi^{\rm I}}(1) &= \{2\}, \quad \beta_2^{\varphi^{\rm I}}(2) = \{1\}, \quad \beta_2^{\varphi^{\rm I}}(3) = \{2\}. \end{split}$$

Therefore, this game has no pure strategy equilibrium.

As before, similar observations can be made in this example. Thus, we will make only a brief remark on the first point: an incentive to reduce capacity. In the example above, it is easy to see that if q' = (3, 2), then every intern is matched with her most preferable hospital: in the intern-proposing deferred acceptance algorithm, interns offer to their most preferable hospitals at step 1, and having enough capacity, hospitals accept all interns. However, such a matching is not good for hospitals. Hospitals are getting their least preferable acceptable interns. Hospitals can do better by swapping interns. Hospital h_1 can obtain intern h_2 by cutting its capacity to 2, because h_2 is a more preferable intern than h_2 for hospital h_2 . By rejecting h_2 through reducing its capacity, hospital h_1 gets a chance to obtain h_2 .

4 Results on welfare properties of equilibria

In both examples in the previous section, we observe that when a hospital reduces its capacity, the other hospital is better off. Although interns are not the active players in our games, their welfare is affected by capacity manipulation of hospitals. When all the examples given above are inspected, it is straightforward to see that at equilibrium, no intern ever benefits by capacity misreports.

This is not a coincidence. The next lemma shows that if some hospital decreases its quota under the hospital-optimal and intern-optimal stable rules, no other hospital will be worse off and no intern will be better off. This lemma plays an important role in proving our main result on welfare properties of the equilibria (when pure

²⁰ Responsiveness together with $\{i_3, i_4\}P_{h_1}\{i_1\}$ implies $\{i_2, i_3\}P_{h_1}\{i_1\}$.

strategy equilibria exist). This lemma will be referred to as the "Capacity Lemma." Its proof depends on a similar result proven for one-to-one matching (or marriage) markets.

Lemma 1 (Capacity Lemma) In capacity-reporting games under φ^H and φ^I , a hospital's capacity underreport makes all other hospitals weakly better off and all interns weakly worse off.

Proof Let (H, I, q, R) be a hospital-intern market. We define the *corresponding* one-to-one two-sided matching market (H^c, I, q^c, R^c) . In the corresponding hospital-intern market, every position of any hospital h is itself an agent. The set of positions is denoted by H^c . Positions of any $h \in H$ are indexed from 1 to q_h and each of them has the same preferences as hospital h over the members of $X_h^I = \{\{i_1\}, \{i_2\}, \ldots, \{i_n\}, \emptyset\}$. Formally, we denote kth position of hospital h by h^k . For any index $k \in \{1, 2, \ldots, q_h\}$ and $i, i' \in I$, (i) $\{i\}P_{h^k}^c\{i'\} \iff \{i\}P_h\{i'\}$, and (ii) $\{i\}P_{h^k}^c\}\emptyset \iff \{i\}P_h\emptyset$.

By definition, each position has capacity 1: $q_{h^k}^c = 1$ for all $h^k \in H^c$.

The interns are still the same agents. The preference relation of intern i, R_i^c is defined on the members of $X_i^{H^c} = \{\{h_1^1\}, \ldots, \{h_1^{q_1}\}, \ldots, \{h_m^1\}, \ldots, \{h_m^{q_m}\}, \emptyset\}$. For any $i \in I$ and any $h, h' \in H$, (i) $\{h^k\}P_i^c\{h'^\ell\}$ for any $k \leq q_h$ and $\ell \leq q_{h'} \iff \{h\}P_i\{h'\}$, (ii) $\{h^k\}P_i^c\{h^\ell\} \iff k < \ell$, and (iii) $\{h^k\}P_i^c\emptyset$ for all $k \leq q_h \iff \{h\}P_i\emptyset$. Market (H^c, I, q^c, R^c) is a hospital-intern market including hospitals with capacity 1.

For $\mu \in \mathcal{M}(H, I, q, R)$, the *corresponding* matching $\mu^c \in \mathcal{M}(H^c, I, q^c, R^c)$ is defined as follows:

- (i) For any $i \in I$ and $h \in H$, $\mu_i = \{h\} \iff \mu_i^c = \{h^k\}$ for some $k \in \{1, 2, \dots, q_h\}$.
- (ii) For any $h \in H$ and $i \in I$, $i \in \mu_h \iff \mu_{h^k}^c = \{i\}$ for some $k \in \{1, 2, \dots, |\mu_h|\}$ and $\mu_{h^\ell}^c P_h \mu_{h^k}^c$ for all $\ell \in \{1, 2, \dots, k-1\}$.
- (iii) For any $i \in I$, $\mu_i = \emptyset \iff \mu_i^c = \emptyset$.
- (iv) For any $h \in H$ and for every $\dot{k} > |\mu_h|$, we have $\mu_{h^k}^c = \emptyset$.

Lemma 1 in Roth and Sotomayor (1989, p. 566) implies that a matching μ is in S(H, I, q, R) if and only if its corresponding matching μ^c is in $S(H^c, I, q^c, R^c)$. By Theorem 4 in Roth and Sotomayor (1989, p. 568), we know that for any $v, \eta \in S(H, I, q, R)$ if there is some $h \in H$ with $v_h P_h \eta_h$, then for all $i \in v_h$ and all $j \in \eta_h \setminus v_h$ we have $\{i\} P_h \{j\}$. Hence, each position of each hospital h gets its most preferred intern in matching $(\mu^H(H, I, q, R))^c$ among all stable matchings of (H^c, I, q^c, R^c) . This implies $(\mu^H(H, I, q, R))^c$ is the hospital-optimal stable matching of market (H^c, I, q^c, R^c) : $(\mu^H(H, I, q, R))^c = \mu^H(H^c, I, q^c, R^c)$. Similarly, each position h gets its least preferred intern in matching $(\mu^I(H, I, q, R))^c$ among all stable matchings of (H^c, I, q^c, R^c) . This implies that $(\mu^I(H, I, q, R))^c = \mu^I(H^c, I, q^c, R^c)$. By Proposition 2 of Gale and Sotomayor (1985b, p. 264), decreasing the number of positions in H^c does not harm any other position under the hospital position-optimal and intern-optimal stable rules in (H^c, I, q^c, R^c) . By responsiveness of preferences, decreasing the capacity of a hospital does not harm other hospitals under the hospital-optimal and intern-optimal stable rules in

(H, I, q, R). On the other hand, again by Proposition 2 of Gale and Sotomayor (1985b, p. 264), decreasing the number of positions in H^c does not make any intern better off under the hospital position-optimal and intern-optimal stable rules in (H^c, I, q^c, R^c) . This fact implies that decreasing the capacity of any hospital in H does not make any intern better off under the hospital-optimal and intern-optimal stable rules in (H, I, q, R).

Next, we examine the properties of pure strategy equilibria when they exist. First, consider the following example.

Example 1 Let (H, I, q, R) be a hospital-intern market such that $H = \{h_1, h_2\}$ and $I = \{i_1, i_2, i_3, i_4\}$. Let $q_1 = 4$ and $q_2 = 4$. Hospitals' preferences satisfy

$$\{i_1\}P_{h_1}\{i_3\}P_{h_1}\{i_2\}P_{h_1}\{i_4\}P_{h_1}\emptyset$$
 with $\{i_1\}P_{h_1}\{i_2,i_3,i_4\}$

$$\{i_3\}P_{h_2}\{i_2\}P_{h_2}\{i_4\}P_{h_2}\{i_1\}P_{h_2}\emptyset$$
 with $\{i_2\}P_{h_2}\{i_1, i_4\}$

and interns' preferences are stated as

$$\{h_2\}P_{i_1}\{h_1\}P_{i_1}\emptyset$$

 $\{h_1\}P_{i_2}\{h_2\}P_{i_2}\emptyset$
 $\{h_1\}P_{i_3}\{h_2\}P_{i_3}\emptyset$
 $\{h_1\}P_{i_4}\{h_2\}P_{i_4}\emptyset$

Below, we give the outcome matchings of the capacity-reporting games under φ^H and φ^I for each pure strategy $q' \in Q$, with equilibria of the two games highlighted in bold:

φ^{H}		1	$q_2' \\ 2$	3	4
	1	$\begin{pmatrix} h_1 & h_2 & \emptyset \\ \{i_1\} & \{i_3\} & i_2, i_4 \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 & \emptyset \\ \{i_1\} & \{i_2, i_3\} & i_4 \end{pmatrix}$	$\begin{pmatrix} \mathbf{h}_1 & \mathbf{h}_2 \\ \{\mathbf{i}_1\} & \{\mathbf{i}_2, \mathbf{i}_3, \mathbf{i}_4\} \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 \\ \{i_3\} & \{i_1, i_2, i_4\} \end{pmatrix}$
q_1'	2	$\begin{pmatrix} h_1 & h_2 & \emptyset \\ \{i_1, i_3\} & \{i_2\} & i_4 \end{pmatrix}$	$\begin{pmatrix} \mathbf{h}_1 & \mathbf{h}_2 \\ \{\mathbf{i}_1, \mathbf{i}_3\} & \{\mathbf{i}_2, \mathbf{i}_4\} \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 \\ \{i_2, i_3\} & \{i_1, i_4\} \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 \\ \{i_2, i_3\} & \{i_1, i_4\} \end{pmatrix}$
	3	$\begin{pmatrix} \mathbf{h}_1 & \mathbf{h}_2 \\ \{\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3\} & \{\mathbf{i}_4\} \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 \\ \{i_2, i_3, i_4\} & \{i_1\} \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 \\ \{i_2, i_3, i_4\} & \{i_1\} \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 \\ \{i_2, i_3, i_4\} & \{i_1\} \end{pmatrix}$
	4	$\begin{pmatrix} h_1 & h_2 \\ \{i_2, i_3, i_4\} & \{i_1\} \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 \\ \{i_2, i_3, i_4\} & \{i_1\} \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 \\ \{i_2, i_3, i_4\} & \{i_1\} \end{pmatrix}$	$\begin{pmatrix} \mathbf{h}_1 & \mathbf{h}_2 \\ \{\mathbf{i}_2, \mathbf{i}_3, \mathbf{i}_4\} & \{\mathbf{i}_1\} \end{pmatrix}$

In the capacity-reporting game under the hospital-optimal rule, every hospital weakly prefers each of the equilibrium matchings $\varphi^H(1,3)$, $\varphi^H(2,2)$, and $\varphi^H(3,1)$ to equilibrium matching $\varphi^H(4,4)$. Moreover, every intern weakly prefers $\varphi^H(4,4)$ to each of $\varphi^H(1,3)$, $\varphi^H(2,2)$, and $\varphi^H(3,1)$.

Similarly, in the game under the intern-optimal rule, every hospital weakly prefers each of the equilibrium matchings $\varphi^I(1,2)$ and $\varphi^I(2,1)$ to equilibrium matching $\varphi^I(4,4)$. Moreover, every intern weakly prefers $\varphi^I(4,4)$ to each of $\varphi^I(1,2)$ and $\varphi^I(2,1)$. The remaining three pure strategy equilibria, (3,4), (4,3), and (3,3), yield the same equilibrium outcomes as (4,4) does.

We observe a few interesting properties in Example 1. First, these two games have very different sets of equilibria. Except truthful capacity revelation, there is no other common equilibrium profile. This observation is not very surprising.

Second, when a pure strategy equilibrium exists, truthful capacity revelation need not be an equilibrium. Consider the capacity-reporting game under φ^H . If we restricted the true capacities of hospitals to $q_1=3$ and $q_2=3$, (3,3) – truthful capacity revelation – would not be an equilibrium, but (1,3), (2,2), and (3,1) would be. Next consider the capacity-reporting game under φ^I . If we restricted the true capacities as $q_1=2$ and $q_2=2$, (2,2) – truthful capacity revelation – would not be an equilibrium, but (1,2) and (2,1) would be.

Third, and more important, in both games, the truthful capacity revelation equilibrium generates a matching that is no better than the outcome of the other equilibria for hospitals and at least as good as the outcome of the other equilibria for the interns. This property is not an exception and does not depend on the choice of φ^H or φ^I . We can state a more general result about this observation as a theorem. The proof of the theorem is composed of simple applications of the Capacity Lemma.

Theorem 3 Let $V \in \{H, I\}$. If $q' \in Q$ is an equilibrium in the capacity reporting game under φ^V and $q'' \in Q$ such that $q' \leq q''$, then (i) every hospital weakly prefers $\varphi^V(q')$ to $\varphi^V(q'')$, and (ii) every intern weakly prefers $\varphi^V(q'')$ to $\varphi^V(q')$. Proof Let (H, I, q, R) be a hospital-intern market. First, we consider the capacity-reporting game under φ^H . Suppose that $q' \in Q$ is an equilibrium of this game-form and $q'' \in Q$ such that $q' \leq q''$. Let h be an arbitrary hospital. Since q' is an equilibrium, $\varphi_h^H(q')R_h\varphi_h^H(q''_n, q'_{-h})$ follows. By the Capacity Lemma, $\varphi_h^H(q''_n, q'_{-h})R_h\varphi_h^H(q'')$. Thus, $\varphi_h^H(q')R_h\varphi_h^H(q'')$. On the other hand, a direct application of the Capacity Lemma implies $\varphi_i^H(q'')R_i\varphi_i^H(q')$. A similar proof can be given for the capacity-reporting game under φ^I .

We can state an immediate corollary to this theorem. If there is an equilibrium different from truthful capacity revelation, every hospital weakly prefers this equilibrium outcome to the outcome of truthful capacity revelation. On the other hand, every intern weakly prefers the outcome of truthful capacity revelation to this equilibrium outcome.

5 Positive results under restricted domains of preferences

In this section, we consider two preference restrictions each of which guarantees the existence of a pure strategy equilibrium. These restrictions are motivated by the observations that followed Theorem 1. The first restriction requires that hospitals prefer a larger number of acceptable interns irrespective of the preferability of each intern. Under this restriction, named "strong monotonicity in population," we will not observe best response cycles discussed in the third remark in the subsequent discussion for Theorem 1. The second restriction is to require common preferences for one group: we consider the cases where either (i) all hospitals have common preferences over interns and their preferences over groups of interns are responsive but otherwise arbitrary, or (ii) all interns have common preferences over hospitals. In either case, the sets of acceptable interns (and hospitals) are equal to the set of interns (and hospitals). Such a restriction, named "common preference rankings for one group," removes conflicts of interests between interns and hospitals. Thus, welfare decreasing swappings for hospitals due to truthful capacity revelation will not occur.

5.1 Strong monotonicity in population of acceptable interns

We start with the case where hospitals prefer larger number of acceptable interns irrespective of the preferability of each intern. Let $A_h = \{i \in I : \{i\}P_h\emptyset\}$ be the set of acceptable interns for hospital h. We say that hospitals' preferences satisfy *strong monotonicity in population*, if and only if for any $h \in H$ and for any $L, L' \subseteq A_h$ we have $|L| > |L'| \Rightarrow LP_hL'$. Let $\mathcal{R}^{\text{SM}} \subset \mathcal{R}$ be the domain of such profiles of responsive preference relations.

First, we consider the capacity-reporting game under φ^H . The following result establishes that if every hospital h reports $\left|\varphi_h^H(q)\right|$ as its capacity, this strategy will be an equilibrium of the game.

Theorem 4 Let (H, I, q, R) be a hospital-intern market such that $R \in \mathcal{R}^{SM}$. If $q_h^* = |\varphi_h^H(q)|$ for every hospital h, then q^* will be an equilibrium of the capacity-reporting game under φ^H .

Proof Let (H, I, q, R) be a hospital-intern market such that $R \in \mathcal{R}^{SM}$. Let $q_h^* = |\varphi_h^H(q)|$ for each $h \in H$. We will prove that q^* is an equilibrium using the following claim.

Claim If $\widehat{q} \in Q$ with $q^* \leq \widehat{q}$, then for each hospital h we have $\varphi_h^H(q^*) R_h \varphi_h^H(\widehat{q})$.

Proof of Claim Let $\widehat{q} \in \mathcal{Q}$ such that $q^* \leq \widehat{q}$. Matching $\varphi^H(q)$ is feasible in the market (H, I, \widehat{q}, R) . Suppose that $\varphi^H(q) \notin \mathcal{S}(H, I, \widehat{q}, R)$. We have either (i) there exists a pair (h, i) blocking $\varphi^H(q)$ in market (H, I, \widehat{q}, R) , or (ii) there exists some agent v and some agent $y \in \varphi^H_v(q)$ such that $\emptyset P_v y$. If case (i) holds, (h, i) will block $\varphi^H(q)$ also in market (H, I, q, R), since $q_h \geq \widehat{q}_h$. If case (ii) holds, agent v will block $\varphi^H(q)$ also in market (H, I, q, R). In both cases, the result will be a contradiction to $\varphi^H(q) \in \mathcal{S}(H, I, q, R)$. We showed that $\varphi^H(q) \in \mathcal{S}(H, I, \widehat{q}, R)$. Similarly, we can show that $\varphi^H(q) \in \mathcal{S}(H, I, q^*, R)$. By Theorem 1 in Roth and Sotomayor (1989, p. 566), for any $h \in H$, for any stable matching $\eta \in \mathcal{S}(H, I, q^*, R)$ we have $|\eta_h| = |\varphi_h^H(q)| = q_h^*$ and for any stable matching $v \in \mathcal{S}(H, I, \widehat{q}, R)$ we have $|v_h| = |\varphi_h^H(q)| = q_h^*$. So $\varphi^H(\widehat{q})$ is feasible in (H, I, q^*, R) . Any pair or any agent that blocks it in market (H, I, q^*, R) can block it also in market (H, I, \widehat{q}, R) . Therefore, $\varphi^H(\widehat{q}) \in \mathcal{S}(H, I, q^*, R)$. Since $\varphi^H(q^*)$ is hospital-optimal in (H, I, q^*, R) , $\varphi_h^H(q^*)R_h\varphi_h^H(\widehat{q})$ for each hospital h.

Consider capacity q^* . Let $h \in H$. By the Claim,

$$\varphi_h^{\mathrm{H}}(q^*)R_h\varphi_h^{\mathrm{H}}(\widehat{q}_h,q_{-h}^*)$$
 for any $\widehat{q}_h > q_h^*$.

By strong monotonicity in population,

$$\varphi_h^{\mathrm{H}}(q^*)P_h\varphi_h^{\mathrm{H}}(\widehat{q}_h,q_{-h}^*)$$
 for any $\widehat{q}_h < q_h^*$.

Hence, q^* is an equilibrium of the capacity-reporting game under φ^H .

Note that this theorem says *neither* (i) truthful capacity revelation strategy profile q is an equilibrium, nor (ii) $\varphi^H(q^*) = \varphi^H(q)$. These points can be seen in the following example, which is a modified version of Example 1.

Example 2 Let (H, I, q, R) be a hospital-intern market such that $H = \{h_1, h_2\}$ and $I = \{i_1, i_2, i_3, i_4\}$. Let $q_1 = 3$ and $q_2 = 3$. The profile of preference relations, R, which satisfies strong monotonicity in population $(R \in \mathbb{R}^{SM})$, is stated as

$$\begin{split} \{i_1\}P_{h_1}\{i_3\}P_{h_1}\{i_2\}P_{h_1}\{i_4\}P_{h_1}\varnothing\\ \{i_3\}P_{h_2}\{i_2\}P_{h_2}\{i_4\}P_{h_2}\{i_1\}P_{h_2}\varnothing\\ \{h_2\}P_{i_1}\{h_1\}P_{i_1}\varnothing\\ \{h_1\}P_{i_2}\{h_2\}P_{i_2}\varnothing\\ \{h_1\}P_{i_3}\{h_2\}P_{i_3}\varnothing\\ \{h_1\}P_{i_4}\{h_2\}P_{i_4}\varnothing \end{split}$$

Below, we give the outcome matchings of the capacity-reporting game under φ^H for each pure strategy $q' \in Q$. Equilibria of this game are highlighted in bold:

$arphi^{ m H}$		1	$q_2' \ 2$	3
	1	$\begin{pmatrix} h_1 & h_2 & \emptyset \\ \{i_1\} & \{i_3\} & i_2, i_4 \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 & \emptyset \\ \{i_1\} & \{i_2, i_3\} & i_4 \end{pmatrix}$	$\begin{pmatrix} \mathbf{h}_1 & \mathbf{h}_2 \\ \{\mathbf{i}_1\} & \{\mathbf{i}_2, \mathbf{i}_3, \mathbf{i}_4\} \end{pmatrix}$
q_1'	2	$\begin{pmatrix} h_1 & h_2 & \emptyset \\ \{i_1, i_3\} & \{i_2\} & i_4 \end{pmatrix}$	$\begin{pmatrix} \mathbf{h}_1 & \mathbf{h}_2 \\ \{\mathbf{i}_1, \mathbf{i}_3\} & \{\mathbf{i}_2, \mathbf{i}_4\} \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 \\ \{i_2, i_3\} & \{i_1, i_4\} \end{pmatrix}$
	3	$\begin{pmatrix} \mathbf{h}_1 & \mathbf{h}_2 \\ \{\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3\} & \{\mathbf{i}_4\} \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 \\ \{i_2, i_3, i_4\} \ \{i_1\} \end{pmatrix}$	$\begin{pmatrix} h_1 & h_2 \\ \{i_2, i_3, i_4\} & \{i_1\} \end{pmatrix}$

Note that $q^*=(3,1)$, and $q'=q^*$ is an equilibrium of the game (as Theorem 4 says). However, truthful capacity revelation is not an equilibrium, and $\varphi^H(q)$ does not coincide with $\varphi^H(q^*)$.

We see a very different result in the capacity-reporting game under φ^{I} : truthful capacity revelation is a weakly dominant strategy when preferences satisfy strong monotonicity in population.

Theorem 5 Let (H, I, q, R) be a hospital-intern market such that $R \in \mathbb{R}^{SM}$. In the capacity-reporting game under φ^I , q_h is a weakly-dominant strategy for each $h \in H$.

Proof Let (H, I, q, R) be a hospital-intern market such that $R \in \mathcal{R}^{SM}$. Consider the capacity-reporting game under φ^I . Let h be an arbitrary hospital. Suppose that q'_{-h} is the reported capacities of all the remaining hospitals. Let $\left|\varphi_h^I(q_h, q'_{-h})\right| = q'_h$. Let $\widehat{q}_h \in Q_h$ be a capacity for hospital h such that $q'_h \leq \widehat{q}_h$. Then $\varphi^I(\widehat{q}_h, q'_{-h})$ is a feasible matching in the market $(H, I, (q_h, q'_{-h}), R)$. Furthermore $\varphi^I(q_h, q'_{-h}) \in \mathcal{S}(H, I, (\widehat{q}_h, q'_{-h}), R)$. Otherwise the blocking pair or the blocking agent in market $(H, I, (\widehat{q}_h, q'_{-h}), R)$ would also block it in market $(H, I, (q_h, q'_{-h}), R)$. Since by Corollary 5.30 in Roth and Sotomayor (1990, p. 163) every hospital weakly prefers any other stable matching to the intern-optimal stable matching in market $(H, I, (\widehat{q}_h, q'_{-h}), R)$, it follows that

$$\varphi_h^{\mathrm{I}}(q_h, q'_{-h}) R_h \varphi_h^{\mathrm{I}}(\widehat{q}_h, q'_{-h}).$$

By strong monotonicity in population,

$$\varphi_h^{\mathrm{I}}(q_h, q'_{-h}) P_h \varphi_h^{\mathrm{I}}(\widetilde{q}_h, q'_{-h})$$
 for all $\widetilde{q}_h < q'_h$.

Hence, q_h is a weakly-dominant strategy for hospital h.

The immediate implication of this result is that truthful capacity revelation is an equilibrium of the capacity-reporting game under φ^I . This result is in sharp contrast with the game under φ^H , although strong monotonicity in population guarantees existence of equilibrium in both games.

5.2 Common preferences for one group over agents

Finally, we consider the cases where either all hospitals or all interns have perfectly correlated preferences over the agents of the other group. Preference profile $R \in \mathcal{R}$ satisfies common preferences for hospitals over individual interns if and only if for any $h, h' \in H$ (i) for any $i, i' \in I$ we have $\{i\}P_h\{i'\} \iff \{i\}P_{h'}\{i'\}$ and (ii) for any $i \in I$, we have $\{i\}P_h\emptyset$ and $\{i\}P_{h'}\emptyset$. Preference profile $R \in \mathcal{R}$ satisfies common preferences for interns if and only if for any $i, i' \in I$, (i) for any $h, h' \in H$, we have $\{h\}P_i\{h'\} \iff \{h\}P_{i'}\{h'\}$, and (ii) for any $h \in H$, we have $\{h\}P_i\emptyset$ and $\{h\}P_{i'}\emptyset$. Let $\mathcal{R}^{CH} \subset \mathcal{R}$ and $\mathcal{R}^{CI} \subset \mathcal{R}$ be the domains of such profiles of preference relations, respectively.

²¹ Perfectly correlated preferences may be observed in real life. The same region hospitals and interns may be competing to be matched with each other, as it is in Britain. In this case, the central hospital is almost always preferred to smaller hospitals in the region (Roth 1991). Also, common preferences for schools are observed in student placement for some high schools done by a central authority in Turkey (Balinski and Sönmez 1999). Unlike college admissions, test score is unidimensional (only one skill category), so high schools' preferences are common: they all want the students who got high scores in the test.

²² Here, we assume $A_h = I$ for all $h \in H$ only for simplicity. We can obtain the same result (with some modification of the argument) even when hospitals have heterogeneous sets of acceptable interns: there exists a common ordering \succ over I such that (i) for any $h \in H$ and $i, i' \in A_h$, $\{i\}P_h\{i'\}\iff i \succ i'$, and (ii) for any $h \in H$, $i' \in A_h$, and $i \in I$, $i \succ i' \implies i \in A_h$. For common preference restriction for interns over hospitals, we can weaken the assumption by employing the counter-part assumption for interns' preferences.

First, we consider preference domain \mathcal{R}^{CH} . We can reorder acceptable interns by hospitals' common preference ordering as i_1, i_2, \ldots, i_m . Hospital h has the same preference ordering as this ordering:

$$\{i_1\}P_h\{i_2\}P_h\dots P_h\{i_n\}P_h\emptyset$$

Let $Ch_i(G)$ be the most preferable acceptable hospital in $G \subseteq H$ for intern i, i.e., $Ch_i(G) = \{h \in G : \{h\}R_i\{h'\} \text{ for any } h' \in G, \text{ and } \{h\}P_i\emptyset\}$. It is useful to have a matching μ^* generated by the following *serial-dictatorship*: ²³

Step 1 Let $G^1 = H$ and $q_h^1 = q_h$ for each $h \in H$. Set $\mu_{i_1}^*(H, I, q, R) = Ch_{i_1}(G^1)$.

: :

Step t For each $h \in H$ let $q_h^t = q_h^{t-1} - 1$ if $\mu_{i_{t-1}}^* = \{h\}$, and $q_h^t = q_h^{t-1}$ otherwise. Let $G^t = \{h \in H : q_h^t \neq 0\}$. Set $\mu_{i_t}^* = Ch_{i_t}(G^t)$.

The algorithm terminates after n steps, and μ^* becomes a matching of the hospital-intern market.

We have the following lemma.²⁴

Lemma 2 Let (H, I, q, R) be a hospital-intern market such that $R \in \mathbb{R}^{CH}$. Then μ^* is the unique stable matching.

Proof Let (H, I, q, R) be a hospital-intern market such that $R \in \mathcal{R}^{\text{CH}}$. We will prove that μ^* is the unique stable matching of this market. Before proving the statement, note $G^t \supseteq G^{t+1}$ applies for any $t \le n-1$. This can be seen from the fact that $G^t = \{h \in H : q_h^t \ne 0\}$ monotonically shrinks (weakly) by construction. First, $\mu_{i_1} = Ch_{i_1}(G^1)$ at any stable matching μ . Otherwise, pair $(i_1, Ch_{i_1}(G^1))$ will block μ if $Ch_{i_1}(G^1) \ne \emptyset$; and i_1 would be better of by staying unmatched if $Ch_{i_1}(G^1) = \emptyset$. Given this, $\mu_{i_2} = Ch_{i_2}(G^2)$ at any stable matching μ . Otherwise, $(i_2, Ch_{i_2}(G^2))$ will block μ if $Ch_{i_2}(G^2) \ne \emptyset$; and i_2 would be better off by staying unmatched if $Ch_{i_2}(G^2) = \emptyset$. Similarly, for any $t \le n$, $\mu_{i_t} = Ch_{i_t}(G^t)$ at any stable matching.

We can now prove that truthful capacity revelation is a weakly dominant strategy in the capacity-reporting game under the stable matching rule. ²⁵

Theorem 6 Let (H, I, q, R) be a hospital-intern market such that $R \in \mathcal{R}^{CH}$. In the capacity-reporting game under stable matching rule, q_h is a weakly dominant strategy for each $h \in H$.

²³ Loosely speaking, stability and common preference property induces a hierarchy very much like the hierarchy inherent in serial dictatorships in the context of one-sided matching markets. See Svensson (1994), Abdulkadiroğlu and Sönmez (1998), and Papai (2000).

 $^{^{24}}$ Alcalde (1995) and Banerjee et al. (2001) prove uniqueness of stable allocations for different kinds of markets by imposing preference restrictions called α -reducibility and top-coalition property, respectively. The proofs of the following two lemmas in the current paper are similar to theirs.

²⁵ Since stable matching is unique under \mathcal{R}^{CH} , we do not need to specify the stable matching rule to be φ^H or φ^I . The same comment applies to Theorem 7.

Proof Let (H, I, q, R) be a hospital-intern market such that $R \in \mathcal{R}^{\mathrm{CH}}$. Let φ^* be the unique stable matching rule in (H, I, q, R). Pick a hospital $h \in H$, and consider a possible capacity report $q'_h < q_h$. Let $\varphi^*_h(q_h, q'_{-h}) = \{i_k, i_l, \ldots, i_r\}$, where $k < l < \cdots < r \le n$. For any t < k, and any $q'_h < q_h, i_t \notin \varphi^*_h(q')$ follows by the construction of $\varphi^*(q')$. As long as $q'_h \ge 1$, $i_k \in \varphi^*_h(q')$ also follows. By the same argument, (i) for any t < l with $t \ne k$ and any $q'_h < q_h, i_t \notin \varphi^*_h(q')$, and (ii) as long as $q'_h \ge 2$, $i_l \in \varphi^*_h(q')$. Thus, by iteration of the same argument, $\varphi^*_h(q') \subseteq \varphi^*_h(q_h, q'_{-h})$ for any $q'_h < q_h$; moreover, $\varphi^*_h(q_h, q'_{-h}) \setminus \varphi^*_h(q')$ contains an acceptable intern if it is non-empty. By responsiveness of hospital preferences, hospital h is not better off by reporting q'_h . Hence, there is no incentive for a hospital to misreport its capacity.

Second, we consider preference domain \mathcal{R}^{CI} . We can reorder acceptable hospitals by interns' common preference ordering as h_1, h_2, \ldots, h_m . Intern i has the same preference ordering as this ordering:

$${h_1}P_i{h_2}P_i \dots P_i{h_m}P_i\emptyset$$

Recall that A_h denotes the set of acceptable interns for hospital h. Let $Ch_h(J,q_h)$ be the set of most preferable acceptable interns among $J \subseteq I$ for hospital h under capacity $q'_h \in Q_h$, i.e., $Ch_h(J,q'_h) = \{L \subseteq J \cap A_h \text{ with } |L| \le q'_h : LR_hL' \text{ for any } L' \subseteq J \text{ with } |L'| \le q'_h\}$. It is useful to have a matching μ^{**} generated by the following serial-dictatorship:

Step 1 Let
$$J^1 = I$$
. Set $\mu_{h_1}^{**} = Ch_{h_1}(J^1, q_{h_1})$.

Step t Let
$$J^t = J^{t-1} \setminus Ch_{h_{t-1}}(J^{t-1}, q_{h_{t-1}})$$
. Set $\mu_{h_t}^{**} = Ch_{h_t}(J^t, q_{h_t})$.

The algorithm terminates after m steps, and μ^{**} is a matching of the hospital-intern market.

We have the following lemma.

Lemma 3 Let (H, I, q, R) be a hospital-intern market such that $R \in \mathcal{R}^{CI}$. Then μ^{**} is the unique stable matching.

Proof Let (H, I, q, R) be a hospital-intern market such that $R \in \mathbb{R}^{CI}$. We will prove that μ^{**} is the unique stable matching of this market. Before proving the statement, note $J^t \supseteq J^{t+1}$ applies for any $t \le m-1$, since the set $J^t = J^{t-1} \setminus Ch_{h_{t-1}}(J^{t-1}, q_{h_{t-1}})$ monotonically shrinks (weakly) as t increases. First, $\mu_{h_1} = Ch_{h_1}(I)$ at any stable matching μ . Otherwise, if $Ch_{h_1}(I, q_{h_1}) \ne \emptyset$, then for some $i \in Ch_{h_1}(I, q_{h_1})$, pair (i, h_1) will block μ ; and if $Ch_{h_1}(I, q_{h_1}) = \emptyset$, then h_1 will be better off by staying unmatched and deviating from μ . Given this, $\mu_{h_2} = Ch_{i_2}(J^2, q_{h_2})$ at any stable matching μ . Otherwise, if $Ch_{h_2}(J^2, q_{h_2}) \ne \emptyset$, then for some $i \in Ch_{h_2}(J^2, q_{h_2})$, pair (i, h_2) will block matching μ ; and if $Ch_{h_2}(J^2, q_{h_2}) = \emptyset$, then h_2 will be better off by staying unmatched. Similarly, for any $t \le m$, $\mu_{h_1} = Ch_{h_1}(J^t, q_{h_1})$ at any stable matching.

We can now prove that truthful capacity revelation is a weakly dominant strategy in the capacity-reporting game under the stable matching rule.

Theorem 7 Let (H, I, q, R) be a hospital-intern market such that $R \in \mathbb{R}^{CI}$. In the capacity-reporting game under the stable matching rule, q_h is a weakly dominant strategy for each $h \in H$.

Proof Let (H, I, q, R) be a hospital-intern market such that $R \in \mathcal{R}^{\text{CI}}$. Let φ^{**} be the unique stable matching rule in (H, I, q, R). Pick a hospital $h_t \in H$. Since J^t is solely determined by reports $\{q'_{h_1}, \ldots, q'_{h_{t-1}}\}$, hospital h_t cannot affect J^t . Pick $q' \in Q$. We have $\varphi^{**}_{h_t}(q') = Ch_{h_t}(J^t, q'_{h_t})$ and $\varphi^{**}_{h_t}(q_{h_t}, q'_{-h_t}) = Ch_{h_t}(J^t, q_{h_t})$. We have $\varphi^{**}_{h_t}(q_{h_t}, q'_{-h_t}) R_{h_t} \varphi^{**}_{h_t}(q')$ since (i) $\varphi^{**}_{h_t}(q') \subseteq \varphi^{**}_{h_t}(q_{h_t}, q'_{-h_t})$, and (ii) interns in $\varphi^{**}_{h_t}(q_{h_t}, q'_{-h_t})$ are acceptable for h_t . Thus, hospital h_t cannot do better by underreporting.

6 Conclusions

In this paper, we examined strategic interactions among hospitals in manipulating their capacities in hospital-intern matching markets. Hospitals report their capacities taking the matching rule and preferences of agents as given. We found that there may not be any pure strategy equilibrium in the standard preference domain (strict and responsive preferences). Even if equilibria exist, truthful capacity revelation may not be an equilibrium. When there is an equilibrium different from truthful capacity revelation, every hospital weakly prefers this equilibrium outcome to the outcome of any larger capacity profile. This makes capacity manipulation very likely to occur in real life. To make matters worse, there is an aftermarket for hospitals and interns in real life. If there are unmatched interns, hospitals can come back to hire these interns in the aftermarket. This strengthens hospitals' incentives to underreport. They can get better interns by underreporting their capacities in the regular market, and then fill the rest of their capacities with unmatched acceptable interns in the aftermarket. Our welfare result (the Capacity Lemma) says that such practices benefit hospitals but are detrimental to interns.

We also find preference restrictions that guarantee the existence of a pure strategy equilibrium. Restrictions on the domain of preferences can make truthful capacity revelation a weakly dominant strategy in some cases, depending on the choice of matching rule. Under strong monotonicity in population, truthful capacity revelation is a weakly dominant strategy under the intern-optimal stable rule, but not under the hospital-optimal stable rule. Interestingly, the National Resident Matching Program in the USA recently replaced the hospital-optimal stable rule with the intern-optimal stable rule based on recommendation by Roth and Peranson (1999).

Acknowledgements We would like to thank Alvin Roth, Tayfun Sönmez, an associate editor and three anonymous referees of the journal for their helpful comments and suggestions which significantly improved the paper. This study was completed when M. Utku Ünver was a faculty member at Koç University, Istanbul, Turkey.

References

Abdulkadiroğlu A, Sönmez T (1998) Random serial dictatorship and the core from random endowments in house allocation problems. Econometrica 66:689–701

J Econ Theory 66:352–369

Abdulkadiroğlu A, Sönmez T (2003) School choice: a mechanism design approach. Am Econ Rev 93:729-747

Alcalde J (1995) Exchange-proofness or divorce-proofness? Stability in one-sided matching markets. Econ Des 1:275-287

Alcalde J (1996) Implementation of stable solutions to the marriage problem. J Econ Theory 69:240-254

Alcalde J, Barberà S (1994) Top dominance and the possibility of the strategy-proof stable solutions to matching problems. Econ Theory 4:417–435

Alcalde J, Romero-Medina A (2000) Simple mechanisms to implement the core of college admissions problem. Games Econ Behav 31:294–302

Balinski M, Sönmez T (1999) A tale of two mechanisms: student placement. J Econ Theory 84:73-94

Banerjee S, Konishi H, Sönmez T (2001) Core in a simple coalition formation game. Soc Choice Welf 18:135–153

Dubins LE, Freedman DA (1981) Machiavelli and the Gale-Shapley algorithm. Am Math Mon 88:485-494

Gale D, Shapley LS (1962) College admissions and the stability of marriage. Am Math Mon

Gale D, Sotomayor M (1985a) Some remarks on the stable marriage problem. Discrete Appl Math 11:223-232

Gale D, Sotomayor M (1985b) Ms. Machiavelli and the stable matching problem. Am Math Mon 92:261-268

Irving RW (1998) Matching medical students to pairs of hospitals: a new variation on a wellknown theme. In:Proceedings of ESA '98, 6th annual European symposium on algorithms, Springer, Berlin Heidelberg New York, pp. 381–392

Kara T, Sönmez T (1996) Nash implementation of matching-rules. J Econ Theory 68:425–439 Kara T, Sönmez T (1997) Implementation of college admission rules. Econ Theory 9:197–218 Ma J (1995) Stable matchings and rematching-proof equilibria in a two-sided matching market.

Ma J (1997) Manipulation and stability in a college admissions problem. Rutgers University working paper

Milgrom P, Shannon C (1994) Monotone comparative statics. Econometrica 62:157–180

Papai S (2000) Strategyproof assignment by hierarchical exchange. Econometrica 68:1403–1433 Postlewaite A (1979) Manipulation via endowments. Rev Econ Stud 46:255-262

Roth AE (1982) The economics of matching: stability and incentives. Math Oper Res 7:617-628 Roth AE (1984) The evolution of the labor market for medical interns and residents: a case study in game theory. J Polit Econ 92:991–1016

Roth AE (1985) The college admissions problem is not equivalent to the marriage problem. J Econ Theory 36:277-288

Roth AE (1991) A natural experiment in the organization of entry level labor markets: regional markets for new physicians and surgeons in the U.K. Am Econ Rev 81:415-440

Roth AE, Peranson E (1999) The redesign of the matching market for American physicians: some engineering aspects of economic design. Am Econ Rev 89:748–780

Roth AE, Rothblum U (1999) Truncation strategies in matching markets - in search of advice for participants. Econometrica 67:21–44

Roth AE, Sotomayor M (1989) The college admissions problem revisited. Econometrica 57:559–

Roth AE, Sotomayor M (1990) Two-sided matching: a study in game theoretic modeling and analysis. Cambridge University Press, Cambridge

Roth AE, Vande Vate JH (1990) Random paths to stability in two-sided matching. Econometrica 58:1475-1480

Roth AE, Xing X (1994) Jumping the gun: imperfections and institutions related to the timing of market transactions. Am Econ Rev 84:992-1044

Sertel MR (1994) Manipulating Lindahl equilibrium via endowments. Econ Lett 46:167-171

Shin S, Suh SC (1996) A mechanism implementing the stable rule in marriage problems. Econ Lett 51:185-189

Sönmez T (1996) Strategy-proofness in many-to-one matching problems. Econ Design 1:365–

Sönmez T (1997a) Games of manipulation in marriage problems. Games Econ Behav 20:169–176

Sönmez T (1997b) Manipulation via capacities in two-sided matching markets. J Econ Theory 77:197-204

Sönmez T (1999) Can pre-arranged matches be avoided in two-sided matching markets? J Econ Theory 86:148-156

Svensson LG (1994) Queue Allocation of Indivisible Goods. Soc Choice Welfare 11:323–330 Thomson W (1987a) Monotonic allocation mechanisms. University of Rochester working paper Thomson W (1987b) Monotonic allocation mechanisms in economies with public goods. University of Rochester working paper

Thomson \dot{W} (1995) Endowment monotonicity in economies with single-peaked preferences. University of Rochester working paper