

Beyond the Spanish MIR with Consent: (Hidden) Cooperation and Coordination in Matching*

José Alcalde[†]

*I thank Jordi Massó, Josep E. Peris, Marek Pycia, Antonio Romero-Medina and the anonymous referees for their useful comments that have helped to improve the paper. This work is partially supported by the Spanish *Ministerio de Economía y Competitividad*, under projects ECO2013-43119-P and ECO2016-77200-P

[†]IUDES[†] University of Alicante. jose.alcalde@ua.es

Abstract

Sequential mechanisms to solve matching problems are useful to promote (hidden) cooperation between agents. Taking as a starting point the MIRC mechanism, employed in Spain to match medical students and residency programs in privately owned hospitals, we find that:

- (1) In the current system, where the number of students that each program might enroll is limited, the unique equilibrium allocation can be unstable.
- (2) When the above limit is not (formally) imposed, instability is not expected to occur. Nevertheless, the multiplicity of equilibria shows that coordination failure might emerge, generating a social welfare loss.
- (3) When the role of students and hospitals is reversed in the MIRC mechanism, (hidden) cooperation is guaranteed. Moreover, coordination failure disappears.

Keywords: MIR with Consent, Hidden Cooperation, Coordination, Matching Mechanisms.

Journal of Economic Literature Classification Numbers: **C78, D61, D71, D78.**

1 Introduction

Some popular real-life matching mechanisms are described as hierarchical selection procedures. The aim of this paper is to shed some light on the open debates on how some allocation mechanisms could be reformulated. This is the case of the Spanish MIR, the internship program for graduating medical students, which annually matches over 6000 graduates with residency programs.

This allocation procedure can be described as follows. Each hospital announces how many students can be admitted in each of its programs. Then students take a highly competitive, comprehensive exam. Their score in such an exam has two consequences. First, the public hospitals rank the student accordingly to their scores,¹ and second, the students sequentially select the program they want to join from those that still have vacancies, where the order in which students are called to select their residency program coincides with that obtained from their score in the exam. There are two additional elements that influence the ability of each student to select a specific hospital, which are based on the distinction of two classes of hospitals: *public hospitals*, handled by regional governments, and *private hospitals*. The first restriction is determined by a rule that states that a student cannot select a medical specialty offered by some private hospital until all the positions offered by public hospitals, in that specialty, have been filled.² The second restriction comes from the different interpretations of the two categories of hospitals of what is the residency aim, as we describe below.

From the public hospitals' point of view, residency is an important part of the educational system for graduating medical students. This implies that, in practice, each public hospital must admit any student, unless all its positions have been filled; and all public hospitals must rank all students according to a common scale, derived from their score in the comprehensive exam. Therefore, given that the order in which students made choices is the same as the common scale, the allocation of internship places (in public hospitals) is done through a serial dictatorship procedure.³

The private hospitals perceive residencies as a training process for potential medical staff, so their participation is not guided by mere educational reasons. In practice, this implies that for a student to be enrolled in some of the private hospital residency programs,

¹ As in the U.S., the score of this exam is a relevant proxy variable for the hospitals to rank different students. In this matter, and related to the American system, Jayakumar (2016) states that "In the NRMP's 2014 report, 94% of residency programs in all specialties reported that USMLE Step 1 scores were a factor in selecting applicants to interview." Our aim in the present paper is not related to any potential modification in the USMLE motivated by the use that hospitals make from its score. Related to this recent debate, the interested readers are directed to Prober et al. (2016), Jayakumar (2016) and Katsufurakis et al. (2016).

² The 2016 call can be consulted at the Spanish Official Bulletin, September 13, 2016. The information about the available places in private hospitals is gathered at p. 65761.

³ Machado et al. (2012) analyze the MIR for public hospitals. They exploit the physicians' hospital choices to infer quality differentials among hospitals.

it is necessary that the hospital explicitly consents to accept the student. Since, in addition, some of the private hospitals are highly specialized in a few medical specialties,⁴ it is hardly assumable a common ranking hypothesis for these hospitals. This is why in this paper we assume that preferences exhibited by private hospital are not necessarily correlated.

In this paper, we concentrate on the MIR mechanism, applied to the private hospitals in Spain, and a variant of this mechanism. This is what it is commonly known as “MIR with consent”, and hereafter referred to as the MIRC mechanism.

A way to understand the relevance of the MIRC, related to the whole MIR system, comes from the comparative between offered positions and the students demanding them. In 2014, the private institutions that participate in the MIRC offered 44 positions associated to different medical and/or surgical specialties. This proposal caught the attention of a relevant group of students. As a result of such an interest 1000 candidates were interviewed to decide whether they qualify or not to obtain the hospital’s consent to be a candidate.⁵ Such a high demand is partially explained by the long-standing tradition and prestige of the hospital participating in the MIRC.

The mechanisms for (two-sided) matching problems studied in the literature associate polar roles to agents belonging to the two groups of involved agents, namely institutions (symbolized in this paper as hospitals) and individuals (exemplified here by students). The agents on one side of the market play an active role because they propose matching agreements to their potential mates, while the agents in the opposite side play a passive role because they just evaluate the offers they have received, and then either accept or reject each of these proposals. Therefore, associated to each matching mechanism, we can describe its dual by exchanging the roles of the agents on the two sides. According to this idea of duality in matching mechanisms, we not only study the consequences of maintaining the MIRC mechanism, but we also analyze the agents’ behavior when faced with the dual of the MIRC mechanism, which we call the DMIRC mechanism.

Since the mechanisms that we analyze in this paper are sequential, to study the agents’ behavior when faced to these mechanisms we concentrate on their Subgame Perfect Nash Equilibria. This simplifies our analysis because the agents in one side of the market (either students when faced to the MIRC mechanism, or hospitals if faced to the DMIRC mechanism) have a dominant strategy. Therefore each of these games can be reduced to a game where the agents not having a dominant strategy play simultaneously (Baron and Kalai, 1993). We then concentrate in the Nash equilibria of these games that survive to an iterative discarding of dominated strategies.

Our main conclusion is that the DMIRC mechanism implements the Student-Optimal Stable allocation. Beyond theoretical research, this result has some policy implications.

⁴For instance, the Barraquer Ophthalmology Center only offers a residency program in ophthalmology, whereas the CETIR Medical Center only offers a program in nuclear medicine.

⁵This information is extracted from the Redacción Médica’s web page, accessed on 04/09/2017.

The MIRC mechanism must be given up in favor of the DMIRC mechanism. This assertion is sustained because by keeping the MIRC mechanism there is no guarantee that, in equilibrium, the outcome is stable, whereas by adopting the DMIRC, there is only one equilibrium allocation: the stable matching which is optimal from the students' point of view. Our results also contribute to an open debate on which algorithm should be adopted by the NRMP (see Williams, 1995), which led to the adoption of the Student-Optimal Stable Matching mechanism in 1998 (see Section 6.1).

1.1 Related Literature

Since Gale and Shapley (1962), stability of the outcome has been a central requirement for many matching mechanisms to be adopted.⁶ Stability implicitly entails the existence of cooperation among agents. Nevertheless, as pointed out by Roth (1982), when agents are faced with stable matching mechanisms, there are always some agents that can benefit from misrepresenting their true characteristics. The impossibility result by Roth (1982) has motivated the growth of a literature exploring matching mechanisms and the outcomes that are attained in equilibrium.

When agents' decisions are taken simultaneously, Ma (1995) illustrates how the agents' collusion, when selecting their strategies, induces the stability of the equilibrium outcomes; Alcalde (1996) shows how the lack of agents' coordination might induce unstable allocations; and Pycia (2012) points out that any stable allocation can be decentralized through a collusive equilibrium. Summarizing, the literature points out that a collusive selection of the actions taken by the agents seems to be necessary and sufficient to attain stable outcomes. In other words, strategic cooperation should be explicit, rather than hidden, to attain stable outcomes.

Nevertheless, the literature also provides two (non-exclusive) escape routes to obtain stable outcomes without requiring the strategic collusion of agents when faced with the mechanism. In such a case we say that the agents' **cooperation is hidden**.

The first approach to reach this hidden cooperation comes from considering that either agents exhibit some rationalizable strategic behavior (Bernheim, 1984; Pearce, 1984), or they iteratively discard dominated strategies reaching a sophisticated equilibrium (Farquharson, 1969; Moulin, 1979). Alcalde (1996) shows that the Deferred Acceptance mechanism is dominance solvable. This result was extended by Ma (2010) for the Student Optimal Stable Matching mechanism, when students play truncated strategies. The results by Alcalde (1996) and Ma (2010) add a further insight to the description of the agents' behavior. Since there is only one equilibrium allocation, which can be obtained by easy-to-identify strategies, we can also suggest that agents exhibit **coordination** when facing the mechanism, which helps to avoid the utility losses associated with a possible mismatch.

⁶See, e.g., Abdulkadiroğlu and Sönmez (2003), Roth (2003, 2008) or Abdulkadiroğlu et al. (2009).

The second approach comes from the design of sequential mechanisms, and the prediction of agents' decisions via backward induction. In this context, Alcalde and Romero-Medina (2000) show that stable equilibrium allocations can be obtained without agents' strategic collusion. In a close framework, Alcalde et al. (1998) illustrate how the sequentiality in which agents decide also helps to attain stable configurations in job matching markets (Kelso and Crawford, 1982), with variable wages. Triossi (2009) introduces sunk costs in sequential mechanisms. He shows that stable equilibrium allocations are obtained. Haeringer and Wooders (2011) also explore the relationship between sequentiality in the agents' decision and the stability of the equilibrium outcomes.

Sequentiality has also been employed by some other authors as a tool to reach the agents' cooperation. For instance, Sotomayor (2003) explores the design of (non-revelation) sequential mechanisms for the marriage problem implementing stable allocations. The idea she underlines has been recently extended by Romero-Medina and Triossi (2014) to the more general case of many-to-one matching framework. The mechanisms explored by Romero-Medina and Triossi (2014) exhibit some similarities with the ones we introduce in the present paper. Nevertheless, our analysis of the agents' behavior differs from that developed by Romero-Medina and Triossi (2014).

In the mechanisms proposed in the present paper, the agents exhibiting the "passive role" have a dominant strategy, namely to accept the best proposal they receive. This reduces the mechanisms to one-stage games where the players are agents on the "active side of the market." In this reduced game the outcome function associates each profile of strategies with the best-response selection that agents in the "passive side of the market" ought to do. Therefore, in our analysis, we focus on Subgame Perfect Nash Equilibrium, applying a backward induction argument. To refine the equilibria, we apply a notion of stage-undominated strategies in the first stage. Similar arguments are employed in Baron and Kalai (1993), Austen-Smith and Banks (2005, Section 4.1), or Alcalde and Dahm (2016) among others.

The rest of the paper is organized as follows. Section 2 describes the framework and presents the main standard concepts. The MIRC mechanism is introduced in Section 3. In particular, Subsection 3.1 points out the main problems exhibited by the current Spanish system, which allocates medical students to residency programs. It is shown that the MIRC mechanism might not have any stable equilibrium allocation. Therefore, we concentrate on the analysis of its dual, the DMIRC mechanism and this is the aim of Section 4.1. We demonstrate that the DMIRC mechanism helps to conciliate (hidden) cooperation as well as coordination between agents. For completeness, Section 5 studies a slight modification of the MIRC mechanism, where students are not constrained when a hospital's quota is exceeded. It is shown (Section 5.1) that each stable allocation can be supported through an equilibrium. Finally, Section 6 states our main conclusions, as well as suggestions for future research. For presentation convenience, all the proofs, as well as some auxiliary results, are relegated to the Appendix.

2 The Model and Main Definitions

There are two finite, disjoint sets of agents, namely $S = \{s_1, \dots, s_i, \dots, s_n\}$ denoting the set of students; and the set of hospitals, $H = \{h_1, \dots, h_j, \dots, h_m\}$. Associated to hospital h_j there is an integer $q_j \geq 1$ indicating the maximal number of students it can enroll, and it is named its quota.

Each student s_i has a strict, complete, transitive, and asymmetric preference relation \succ_i over $H \cup \{s_i\}$.⁷ Given a set of options for student s_i , say $A \subseteq H \cup \{s_i\}$, $\mathcal{B}(A; \succ_i)$ denotes her maximal on A under preferences \succ_i . In the case of hospitals, each h_j has a strict, complete, transitive, and asymmetric preference relation \succ_j over the possible sets of students 2^S . Given a set of students $S' \subseteq S$, and hospital h_j , $Ch(S', \succ_j)$ denotes h_j choice on S' according to \succ_j , i.e. $S'' = Ch(S', \succ_j)$ whenever (i) $S'' \subseteq S'$ and, (ii) for each $\widehat{S} \subseteq S'$, $\widehat{S} \neq S''$, $S'' \succ_j \widehat{S}$. For agent $x \in S \cup H$, \succeq_x denotes its/her weak preferences; i.e. $y \succeq_x z$ means that either $y \succ_x z$ or $y = z$.

Throughout the paper, we assume that agents, as well as the quota of each hospital, are given. Therefore, a specific problem can be described through the agents' preferences, $P = (\succeq_x)_{x \in H \cup S}$. A solution for P , also called a **matching**, is a correspondence $\mu : S \cup H \rightarrow S \cup H$ such that

- (a) for each student s_i , $\mu(s_i) \in H \cup \{s_i\}$;
- (b) for each hospital h_j , $\mu(h_j) \subseteq S$; and
- (c) for each student-hospital pair (s_i, h_j) , $\mu(s_i) = h_j$ if and only if $s_i \in \mu(h_j)$.

Given a problem P , we say that matching μ is

(1) **Individually Rational** if

- (i) no student prefers her being unmatched option rather than the hospital she is assigned to under μ , if any; i.e., there is no student s_i such that $s_i \succ_i \mu(s_i)$; and
- (ii) no hospital prefers not to enroll some of the students it has been assigned; i.e. for each hospital h_j , $\mu(h_j) = Ch(\mu(h_j), \succ_j)$.

(2) **Stable** if it is individually rational and there is no hospital h_j and (non-empty) set of students S' such that

- (i) for each student $s_i \in S'$, $h_j \succ_i \mu(s_i)$; and
- (ii) $S' \subseteq Ch(\mu(h_j) \cup S', \succ_j)$.

⁷ When student s_i prefers not to be enrolled to any residency program rather than being attached to hospital h_j , it is denoted by $s_i \succ_i h_j$. This is why s_i is also referred as the "being unmatched option" for student s_i .

A pair (h_j, S') which satisfies conditions (i) and (ii) above is said to **block** matching μ .

An individually rational matching which is not blocked by any student-hospital pair is called **Pairwise Stable**.

For a given problem P , $\mathcal{IR}(P)$ denotes the set of its individually rational matchings; $\mathcal{PS}(P)$ is the set of its pairwise stable matchings; and $\mathcal{C}(P)$ denotes the set of its stable matchings.

It is well-known that for any given problem P , $\mathcal{C}(P) \subseteq \mathcal{PS}(P)$, whereas the opposite is not true. Moreover, there are some instances with no pairwise stable allocation, and thus no stable matching.

Roth (1984, 1985) explored the possibility of finding environments where all the problems have stable allocations. Roth (1984) shows that whenever the hospitals preferences satisfy *substitutability*, a stable matching always exists. Roth (1985) studies the case where the hospitals preferences satisfy *responsiveness*, a property guaranteeing the existence of stable matchings, which connects each hospital opinion on isolated students with its opinion over different sets of students.

Definition 1 [Substitutability]

Preferences for hospital h_j , \succ_j , satisfy substitutability when for each set of students $S' \subseteq S$, if $s_i \in Ch(S', \succ_j)$, then $s_i \in Ch(S' \setminus \{s_\ell\}, \succ_j)$ for each $s_\ell \neq s_i$.

A problem P is said to be substitutable if the preferences of all the hospitals satisfy substitutability.

Definition 2 [Responsiveness]

Let h_j be a hospital with quota q_j . We say that its preferences \succ_j satisfy responsiveness if the following holds.

- (i) For each $S' \subseteq S$, $|Ch(S', \succ_j)| \leq q_j$;⁸
- (ii) for each $S' \subset S$, with $|S'| < q_j$ and any two students s_i and s_ℓ not in S' , $S' \cup \{s_i\} \succ_j S' \cup \{s_\ell\}$ if, and only if $\{s_i\} \succ_j \{s_\ell\}$; and
- (iii) for each $S' \subseteq S$, with $|S'| \leq q_j$ and any student $s_i \in S'$, $s_i \in Ch(S', \succ_j)$ if, and only if, $\{s_i\} \succ_j \emptyset$.

A problem P is said to be responsive if the preferences of all the hospitals satisfy responsiveness. Under responsiveness, the following results hold.

Lemma 1 Let P be a responsive problem, then $\mathcal{PS}(P) = \mathcal{C}(P) \neq \emptyset$.

⁸ Given a set T , $|T|$ denotes its cardinality.

Lemma 2 Let P be a responsive problem. Then, there is a matching $\mu^{SO} \in \mathcal{C}(P)$ such that for each $s_i \in S$ and $\mu \in \mathcal{C}(P)$, $\mu^{SO}(s_i) \succeq_i \mu(s_i)$. Such an allocation is known as the *Student-Optimal Stable matching*.

Lemma 3 Let P be a responsive problem. Then, there is a matching $\mu^{HO} \in \mathcal{C}(P)$ such that for each $h_j \in H$ and $\mu \in \mathcal{C}(P)$, $\mu^{HO}(h_j) \succeq_j \mu(h_j)$. Such an allocation is known as the *Hospital-Optimal Stable matching*.

Lemmata 1 and 2 were proven in Roth (1984) for the case where hospitals' preferences satisfy substitutability. Since responsiveness implies substitutability, the results are still valid within this framework. Lemma 3 is borrowed from Roth (1985).

It is usual to assume that hospitals' preferences are responsive; nevertheless, some of our results are still valid without such an assumption.

3 The MIRC Mechanism

In this section we describe how the MIRC mechanism operates. We further develop a formal analysis of which outcomes are likely to occur when agents behave strategically.

This is an $(n + 1)$ -stage game. At the initial stage, each hospital announces a (possibly empty) set of students, which is interpreted as the students that eventually might be enrolled by this hospital. These announcements are made simultaneously. Then, each student selects, according to a given hierarchic ordering, the hospital whose internship program she wants to attend, provided that she meets the eligibility restrictions already determined by hospitals. For simplicity of presentation, we assume that students' decisions are made according their label.⁹ Therefore, student s_1 chooses a program among the hospitals declaring her as 'enrollable', if any; then, student s_2 makes her choice from the hospitals declaring their ability to enroll her, and having a vacancy once s_1 request has been handled; and so on. We formally describe the MIRC mechanism.

Definition 3 [The MIRC Mechanism]

The MIRC mechanism is the following sequential game, with $n + 1$ stages. The set of players is $S \cup H$. Actions are selected according to these (sequential) rules:

- (0) Each hospital h_j selects an action $\sigma_j \subseteq S$. This selection is simultaneously made by all the hospitals. This information becomes public before proceeding to the next stage.
- (1) Student s_1 , taking into account the actions already taken by all the hospitals, chooses her action $a_1 \in \{h_j \in H : s_1 \in \sigma_j\} \cup \{s_1\}$. Then, s_1 is matched to a_1 ; i.e. $\mu^{MIRC}(s_1) = a_1$. This information becomes publicly known by the remaining students.

⁹ Otherwise, relabel the students, and the description will be similar once this relabeling is done.

...

- (ℓ) Student s_ℓ , taking into account the actions taken by all the hospitals and the students preceding her, chooses her action $a_\ell \in H' \cup \{s_\ell\}$, where hospital $h_j \in H'$ if, and only if, $s_\ell \in \sigma_j$, and $|\{s_i \in \{s_1, \dots, s_{\ell-1}\} : a_i = h_j\}| < q_j$. Then, s_ℓ is matched to a_ℓ ; i.e., $\mu^{MIRC}(s_\ell) = a_\ell$. The remaining students are publicly informed about this.

Once all the agents have selected their actions, matching μ^{MIRC} is obtained, as described through stages (1) to (n) above.

3.1 Strategic Behavior and the MIRC Mechanism

We now concentrate on exploring how agents will act when faced with the MIRC mechanism. Note that, given the actions selected by the hospitals in stage (0), it is easy to determine which action each student ought to select.

Let $\mathcal{A}(s_1 | \sigma)$ be set of achievable options for student s_1 , when the actions chosen by the hospitals are $\sigma = (\sigma_1, \dots, \sigma_j, \dots, \sigma_m)$; that is,

$$\mathcal{A}(s_1 | \sigma) = \{h_j \in H : s_1 \in \sigma_j\} \cup \{s_1\}.$$

Therefore, the best-response of s_1 to the hospitals' actions is to choose $a_1^*(\sigma) = \mathcal{B}(\mathcal{A}(s_1 | \sigma); \succ_1)$. Note that the selection of $a_1^*(\sigma)$ as her best available option does not depend on what the other students could eventually select. Moreover, for each given combination of hospitals' actions, say σ , when s_1 selects her best-response, $\mu^{MIRC}(s_1) = a_1^*(\sigma)$.

An iterative argument helps to determine the best-response of s_i to her predecessors' actions, provided that all the students preceding her have already selected their best-response. In such a case a hospital, say h_j , is available for s_i whenever this hospital declares that s_i is an admissible student, and its positions are not yet exhausted; i.e.,

- (a) $s_i \in \sigma_j$, and
- (b) $|\{s_\ell \in \{s_1, \dots, s_{i-1}\} : a_\ell^*(\sigma) = h_j\}| < q_j$.

Let $\mathcal{A}(s_i | \sigma)$ be the set containing all the hospitals available for s_i , as already described, and her being unmatched option, s_i . Note that the arguments above, related to s_1 , can be replicated to show that s_i ' best-response to her predecessors' actions is to select $a_i^*(\sigma) = \mathcal{B}(\mathcal{A}(s_i | \sigma); \succ_i)$, inducing that $\mu^{MIRC}(s_i) = a_i^*(\sigma)$.

Following Baron and Kalai (1993), a reasonable way to predict how the agents will react to the MIRC mechanism is by assuming that hospitals exhibit perfect foresight on what students will select, and thus concentrating on the Subgame Perfect Nash Equilibria with stage-undominated strategies. Our framework, and taking into account the reasoning above about students' best-responses, allows us to reduce the sequential game previously

described to a simpler one-stage game. In this game the players are the hospitals; a strategy for each hospital is a (possibly empty) set of students, and thus the strategy space for each hospital coincides with its set of actions in the MIRC mechanism. The outcome associated to each profile of strategies is the outcome of the MIRC mechanism when each hospital action coincides with the strategy in this game and each student's action is her best-response to her predecessors' actions, as previously described. Let Γ^{HR} denote this hospital-reduced game.

In what follows, and abusing notation, $\mu^{HR}(\sigma)$ denotes the matching obtained when the hospitals, faced with Γ^{HR} , play the vector of (pure) strategies σ . For agent $x \in S \cup H$, $\mu^{HR}(x; \sigma)$ is what the above matching assigns to this agent.

Given the vector of strategies σ and hospital h_j , (σ_{-j}, σ'_j) denotes the vector of strategies obtained from σ by replacing this hospital strategy by σ'_j instead of σ_j .

A **strategy** σ_j is said to be **dominated** for hospital h_j at game Γ^{HR} if there is another strategy for it, say σ'_j , such that (a) for each vector of strategies for the remaining hospitals, say σ_{-j} , $\mu^{HR}(h_j; (\sigma_{-j}, \sigma'_j)) \succeq_j \mu^{HR}(h_j; (\sigma_{-j}, \sigma_j))$; and (b) there is a vector of strategies for the remaining hospitals, say σ'_{-j} such that $\mu^{HR}(h_j; (\sigma'_{-j}, \sigma'_j)) \succ_j \mu^{HR}(h_j; (\sigma'_{-j}, \sigma_j))$. Strategy σ_j is an **undominated strategy** for hospital h_j at game Γ^{HR} whenever no strategy dominates it.

An **equilibrium** for Γ^{HR} is a vector strategies $\sigma^* = (\sigma_1^*, \dots, \sigma_j^*, \dots, \sigma_m^*)$ such that for each hospital h_j and strategy for it, σ_j

$$\mu^{HR}(h_j; \sigma^*) \succeq_j \mu^{HR}(h_j; (\sigma_{-j}^*, \sigma_j)).$$

When faced with Γ^{HR} , some of the available strategies for each hospital might be dominated. Assuming that rational hospitals are not expected to play dominated strategies, Γ^{HR} should be redefined throughout the elimination, for each hospital, of its dominated strategies. Since at the new, restricted game any hospital might have dominated strategies (that were undominated in the original game), an iterative process of eliminating dominated strategies yields a reduced set of strategies for each player, say Σ_j^F , in which no strategy is dominated. An **equilibrium** σ^* is said to be **sophisticated** whenever the strategy selected by each hospital belongs to its (reduced) set of strategies obtained by iterative elimination of dominated strategies, Σ_j^F .

We now deal with the strategic analysis of Γ^{HR} . We first see that any stable matching can be decentralized through an equilibrium when agents are faced with the MIRC mechanism. We want to stress that Proposition 1 below does not require that the hospitals' preferences are responsive. The only necessary condition is the existence of a stable allocation. Moreover, this result specifies the actions that should be taken by the hospitals, revealing that a high level of coordination is needed to reach the equilibrium. Our second result, Proposition 2, establishes that under responsiveness only individual rational matchings can be supported as equilibrium outcomes. But nothing can be guaranteed about the stability of

the equilibrium outcomes. This result is complemented with Example 1 which definitively suggests that the MIRC mechanism must be given up.

Proposition 1 Let P be a problem, and μ be a stable matching for P . Then, the profile of strategies σ^* such that, for each hospital h_j , $\sigma_j^* = \mu(h_j)$ constitutes an equilibrium for Γ^{HR} . Moreover, $\mu^{HR}(\sigma^*) = \mu$.

Proposition 2 Let P be a responsive problem, and Γ^{HR} the hospital-reduced game associated to this problem. If σ^* is an equilibrium for Γ^{HR} , then $\mu^{HR}(\sigma^*)$ is an individually rational matching for P .

Unfortunately, as the next example points out, there might be some equilibrium supporting unstable allocations. Such instability is the cost of the hospitals' admissibility restrictions induced by their quota.¹⁰

Example 1 Let us consider the following instance involving three students, $S = \{a, b, c\}$, and two hospitals $H = \{1, 2\}$, with $q_1 = 2$ and $q_2 = 1$. Hospitals preferences are responsive, and the description of the agents' preferences is¹¹

$$\begin{aligned} \gamma_a &:= 1, 2 & \gamma_1 &:= b, a, c \\ \gamma_b &:= 2, 1 & \gamma_2 &:= a, b, c \\ \gamma_c &:= 1, 2 \end{aligned}$$

It can be seen that this problem has a unique stable matching, $\hat{\mu}$, which is described by

$$\hat{\mu} := \begin{pmatrix} a & b & c \\ 1 & 2 & 1 \end{pmatrix}.$$

According to Proposition 1 above, this matching can be decentralized through an equilibrium for Γ^{HR} , namely $\hat{\sigma}$, with $\hat{\sigma}_1 = \{a, c\}$ and $\hat{\sigma}_2 = \{b\}$. Nevertheless, it is difficult to justify that hospital 2 might select strategy $\hat{\sigma}_2$. This is because such a strategy is a dominated strategy for this hospital. To be precise, let us consider strategies $\sigma_2^* = \{a, b, c\}$, $\sigma_2^{**} = \{a, b\}$, and any $\sigma_2' \notin \{\sigma_2^*, \sigma_2^{**}\}$. It can be seen that, for any strategy σ_1 played by hospital 1,

$$\begin{aligned} \mu^{HR}(2; (\sigma_1, \sigma_2^*)) &= \mu^{HR}(2; (\sigma_1, \sigma_2^{**})) = \{b\} \succ_2 \mu^{HR}(2; (\sigma_1, \sigma_2')) & \text{if } a \in \sigma_1, \text{ whereas} \\ \mu^{HR}(2; (\sigma_1, \sigma_2^*)) &= \mu^{HR}(2; (\sigma_1, \sigma_2^{**})) = \{a\} \succ_2 \mu^{HR}(2; (\sigma_1, \sigma_2')) & \text{if } a \notin \sigma_1. \end{aligned}$$

¹⁰ See Theorem 2 in Section 5.1.

¹¹ Since hospitals' preferences are responsive, given their quatae, the only information that we need related to the hospitals' preferences is how each hospital compares any pair of students, as well as which are admissible. This also applies to Example 2.

Then, provided that hospital 2 ought to select one strategy in $\{\sigma_2^*, \sigma_2^{**}\}$, the set of students $\{a, b\}$ is no longer available for hospital 1. But hospital 1 can obtain $\{b, c\}$, which turns out to be its best available set of students provided that it cannot enroll $\{a, b\}$. This objective is only reached by selecting $\sigma_1^* = \{b, c\}$ because for any $\sigma_2 \in \{\sigma_2^*, \sigma_2^{**}\}$, and each $\sigma_1 \neq \sigma_1^*$,

$$\mu^{HR}(1; (\sigma_1^*, \sigma_2)) = \{b, c\} \succ_1 \mu^{HR}(1; (\sigma_1, \sigma_2)).$$

Note that, using the terminology by Moulin (1979), Γ^{HR} is *dominance solvable* for problem P . Thus, the only equilibria supported by the agents' rationalizable strategic behavior are $\sigma^* = (\sigma_1^*, \sigma_2^*)$ and $\sigma^{**} = (\sigma_1^*, \sigma_2^{**})$, yielding matching

$$\mu^{HR}(\sigma^*) = \mu^{HR}(\sigma^{**}) = \begin{pmatrix} a & b & c \\ 2 & 1 & 1 \end{pmatrix},$$

which fails to be stable for problem P since the pair $(1, a)$ blocks $\mu^{HR}(\sigma^*)$. ■

4 The Dual MIRC Mechanism

In this section we explore the consequences of a slight reformulation of the MIRC mechanism. Our proposal here can be seen as a reform similar to the one which took place in the USA in 1998, when the NMRP moved from a Hospital-Optimal Stable matching mechanism to a Resident-Optimal Stable matching mechanism. This reform involved switching the roles of hospitals and students in the description of the mechanism: Hospitals that, previously were the agents making proposals, become the proposal-receivers; whereas the students, that were the agents receiving proposals, become the proposer-agents.

In this section we develop a game-theoretical analysis for the mechanism being dual to the Spanish MIR with Consent, where such a duality should be understood as the switch in the agents' roles.¹² Summarizing the results of the previous section and those of the present section, a reform in the allocation system employed in Spain, by adopting the Dual of the MIRC mechanism, would be undertaken.

We now proceed to give a formal definition for the **Dual MIRC mechanism**, DMIRC mechanism in short. This is a $(m + 1)$ -stage game involving students -who simultaneously select an action at the initial stage- and hospitals -whose action is decided sequentially according to an exogenously given hierarchical order-. For simplicity of presentation we assume that hospitals choose the students they enroll in their programs according to their labels.

¹²Strictly speaking, a dual formulation of the MIRC mechanism requires that, in the first step, each student selects only one hospital. Here we assume that each student can select as many hospitals as she wants. We are grateful to a referee for pointing out this fact.

Definition 4 [The DMIRC Mechanism]

The DMIRC mechanism is the following sequential game, with $m + 1$ stages. The set of players is $S \cup H$. Actions are selected according to the following (sequential) rules:

- (0) Each student s_i selects an action $\delta_i \subseteq H \cup \{s_i\}$. This selection is simultaneously made by all the students. This information becomes public before proceeding to the next stage.
- (1) Hospital h_1 , taking into account the actions already taken by all the students, decides its own action $a_1^D \subseteq \{s_i \in S : h_1 \in \delta_i\}$. Then, h_1 is assigned all the students in a_1^D ; i.e., $\mu^{DMIRC}(h_1) = a_1^D$. This information becomes publicly known by the remaining hospitals.
- ...
- (ℓ) Hospital h_ℓ , taking into account the actions taken by all the students and the hospitals preceding it, chooses its action $a_\ell^D \subseteq \{s_i : h_\ell \in \delta_i\} \setminus \cup_{j < \ell} \{a_j^D\}$. Then, $\mu^{DMIRC}(h_\ell) = a_\ell^D$. This information becomes publicly known by the remaining hospitals.

For the sake of completeness, any student s_i who is not assigned a hospital after stage m is unmatched (i.e., $\mu^{DMIRC}(s_i) = s_i$ for each such student).

The intuitive description of how the DMIRC mechanism operates is very simple. At the initial stage each student declares which hospitals (if any) are acceptable from her point of view. Then, hospital h_1 selects a set of students among the ones declaring that h_1 is acceptable. Students selected by h_1 are enrolled to h_1 , and thus they are no longer available for the remaining hospitals. For the other hospitals the argument is similar, just taking into account that for a student to be achievable it is not only necessary that she declares that the hospital is acceptable for her, but also that the student has not already been assigned to another hospital.

4.1 Strategic Behavior and the DMIRC Mechanism

A similar reasoning to that developed in Section 3.1 allows us to predict which action ought to be selected by each hospital when faced with the DMIRC mechanism. Given $\delta = (\delta_1, \dots, \delta_i, \dots, \delta_n)$, the vector of actions taken by all the students, the set of achievable students for hospital h_1 is

$$\mathcal{A}^D(h_1 | \delta) = \{s_i \in S : h_1 \in \delta_i\}.$$

The best-response by h_1 to the actions selected by the students is to choose $a_1^{D*}(\delta) = Ch(\mathcal{A}^D(h_1 | \delta); \succ_1)$. Note that when h_1 selects its best-response action it enrolls the students in $a_1^{D*}(\delta) = \mu^{DMIRC}(h_1)$. This decision only depends on the actions taken by the

students and thus it is independent of the actions that the remaining hospitals ought to eventually select. This action is then observed by the remaining hospitals.

An iterative argument on how the hospitals act allows us to anticipate that when h_ℓ has to select its action, given δ and assuming that the hospitals preceding it have selected their best-response to their preceding agents, then the set of available students for h_ℓ is

$$\mathcal{A}^D(h_\ell | \delta) = \{s_i \in S : h_\ell \in \delta_i\} \setminus \bigcup_{j=1}^{\ell-1} a_j^{D*}(\delta).$$

It is reasonable to expect that rational students anticipate that each hospital will select its best-response action to the actions chosen by its predecessors. This fact allows identification of the outcomes that result from agents' interaction when faced with the DMIRC mechanism with those of the student-reduced game Γ^{SR} where players are the students, a strategy for each player is a set of hospitals, and the outcome associated to profile of strategies δ , to be denoted as $\mu^{SR}(\delta)$, is the outcome of the DMIRC mechanism where the actions by the students are the strategies selected in Γ^{SR} , whereas each hospital action is its best response to its predecessors' actions as already described.

We now study which outcomes are likely when agents are faced with the DMIRC mechanism. To reach our objective, we concentrate on the equilibria for game Γ^{SR} .¹³

Our first concern relies on the more general case where no structure is assumed for the hospital preferences. Therefore, Proposition 3 below extends the analysis performed in Romero-Medina and Triossi (2014) related to their *Students Apply College Sequentially Choose mechanism*. This result establishes that each stable allocation can be decentralized through some equilibrium. Thus, the similarities between Propositions 1 and 3 agree with the duality of the DMIRC and the MIRC mechanisms.

Proposition 3 Let P be a problem, and μ be a pairwise-stable matching for P . Then, profile of strategies δ^* such that, for each student s_i , $\delta_i^* = \mu(s_i)$ constitutes an equilibrium for Γ^{SR} .

Corollary 1 Let P be a problem, and μ be a stable matching for P . Then, there is an equilibrium δ^* for Γ^{SR} such that $\mu^{SR}(\delta^*) = \mu$.

As argued in Example 1, it is difficult to assume that students select dominated strategies when playing game Γ^{SR} . Nevertheless, the strategies proposed in Proposition 3, might be dominated for some of the students. It can be shown that, despite what Proposition 3 reports, some stable outcomes might not be supported by any equilibrium when students exhibit a sophisticated behavior (Farquharson, 1969).

¹³ The notions of equilibrium and sophisticated equilibrium for Γ^{SR} are similar to the ones introduced in Section 3.1 for game Γ^{HR} , just taking into account that players in the former game are students, whereas hospitals are the players in Γ^{HR} .

Our first concern about the students' behavior is related to how strategies can be disregarded based on dominance arguments; in particular when, for each student, there might be several strategies being equivalent.¹⁴

For game Γ^{SR} there are some undominated strategies that are rarely selected by the students. The reason is that they are *protectively dominated* strategies (see Definition 5 below).

The next example helps to illustrate the idea captured by protective domination.

Example 2 Consider the following 6-student-3-hospital responsive problem, with $S = \{a, b, c, d, e, f\}$, $H = \{1, 2, 3\}$, and $q_j = 2$ for each $h_j \in H$. Agents' preferences are

$$\begin{array}{lll} \succ_a := 3, 2, 1 & \succ_d := 2, 3, 1 & \succ_1 =: a, b, c, d, e, f \\ \succ_b := 2, 3, 1 & \succ_e := 3, 2, 1 & \succ_2 =: c, e, a, b, d, f \\ \succ_c := 3, 2, 1 & \succ_f := 2, 3, 1 & \succ_3 =: d, f, b, a, e, c \end{array}$$

Assume that, when students play game Γ^{SR} , students c and e play strategy $\delta_c^* = \delta_e^* = \{2\}$ and students d and f play $\delta_d^* = \delta_f^* = \{3\}$. Then, for any strategies played by a and b , say δ_a^* and δ_b^* respectively, $\mu^{SR}(a; \delta^*) \in \{1, a\}$. To be precise, $a = \mu^{SR}(a; \delta^*)$ when $1 \notin \delta_a^*$, whereas $1 = \mu^{SR}(a; \delta^*)$ whenever $1 \in \delta_a^*$. Since $1 \succ_a a$, this implies that strategy $\delta_a^* = \{1\}$ is not dominated by any strategy not containing hospital 1.

Taking into account that hospital 1 is the first hospital to select which students it enrolls, and that a is its preferred student, $\mu^{SR}(a; \delta) = 1$ for each profile of strategies δ such that $\{1\} \subseteq \delta_a$. That is, for student a , all the strategies including hospital 1 are equivalent when playing Γ^{SR} . Therefore, $\delta_a = \{1\}$ is an undominated strategy. ■

Taking into account the instance above, proposed in Example 2, is it reasonable to predict that some of the students declare to be admissible only their less preferred hospital, not including any additional admissible option? Note that the argument in this example can also be replicated for any problem including as many hospitals as wanted! Since no student incurs any cost by listing more hospitals, it does not seem to be a common behavior. How could we explain that no student is expected to list only her worst acceptable hospital?

A justification comes from the analysis of how, in real-life situations, sequential procedures are implemented. Assume that, at the very beginning, the order in which the hospitals select the students they enroll is established. In practice, there is a deadline at which each hospital has to settle its decision. If a hospital delays, it is not excluded, but it is relegated to be the last hospital to decide. No hospital has any interest in being delayed, but there is a small (positive) probability that some hospital incurs a delay situation, so that the order

¹⁴ Two strategies are said to be equivalent for a given student whenever her outcome is the same, no matter which of the two strategies is selected, for any profile of strategies selected by her rivals. For the case of Γ^{SR} , δ_i and δ'_i are equivalent for student s_i whenever for any δ_{-i} played by the other students, $\mu^{SR}(s_i; (\delta_{-i}, \delta_i)) = \mu^{SR}(s_i; (\delta_{-i}, \delta'_i))$.

in which the hospitals made their selections might be altered. Taking into account that such an alteration of the initial order might occur, even with a small probability, it is reasonable that students try to protect against the effects of such a possibility. This protection might be reached by including additional ‘admissible’ hospitals.

Before formally defining protective domination, we need some additional notation. Let $\pi : \{1, \dots, j, \dots, m\} \rightarrow \{1, \dots, j, \dots, m\}$ a permutation of the hospitals. For a given permutation π , Γ_π^{SR} is the student-reduced game associated to the DMIRC mechanism where the hospitals select according to π .¹⁵ $\mu_\pi^{SR}(\delta)$ is the matching that Γ_π^{SR} determines when students play the profile of strategies δ .

Definition 5 Let P be a given problem and Γ^{SR} its associated student-reduced game. We say that δ_i **protectively dominates** δ'_i for student s_i if for any strategies for the other students δ_{-i} , and each permutation of the hospitals π

$$\mu_\pi^{SR}(s_i; (\delta_{-i}, \delta_i)) \succeq_i \mu_\pi^{SR}(s_i; (\delta_{-i}, \delta'_i)),$$

and there is a profile of strategies for the other students δ_{-i}^* and a permutation π^* such that

$$\mu_{\pi^*}^{SR}(s_i; (\delta_{-i}^*, \delta_i)) \succ_i \mu_{\pi^*}^{SR}(s_i; (\delta_{-i}^*, \delta'_i)).$$

Remark 1 Note that the domination of strategy δ_i by δ'_i , for student s_i , is protective when there is a source of stochastic domination when comparing the hospitals that s_i can achieve for any given strategy by the other students, provided that the order in which the hospitals choose their students is selected at random, once the students have decided their strategies.

A similar reasoning to that employed to an iterative deletion of dominated strategies, and the definition of sophisticated equilibrium in Section 3.1, allows us to introduce the notion of **protectively sophisticated equilibrium** below.

Provided that no student selects protectively dominated strategies, we can reduce the set of strategies for each student to the ones not being protectively dominated. An iterative deletion of protectively dominated strategies allows us to determine, for each student, a set of strategies in which no remaining strategy is protectively dominated, provided that no student plays protectively dominated strategies. Let Σ^{PF} the set of profiles of strategies obtained by iterative deletion of protective dominated strategies. We say that $\delta^* \in \Sigma^{PF}$ is a protectively sophisticated equilibrium, PSE in short, whenever for each student s_i and any strategy for her $\delta_i \in \Sigma_i^{PF}$,

$$\mu^{SR}(s_i; \delta^*) \succeq_i \mu^{SR}(s_i; (\delta_{-i}^*, \delta_i)).$$

¹⁵ $\pi(j)$ denotes the turn of h_j ; i.e. if, for instance $\pi(j) = 3$, then h_j is the third hospital that chooses the students it enrolls. As assumed through this paper, Γ^{SR} stands for the game Γ_π^{SR} in which π is the identity permutation, $\pi(j) = j$ for each j .

We now introduce the main result in this section. It establishes that, under responsiveness, any equilibrium remaining after an iterative deletion of protectively dominated strategies decentralizes the Student-Optimal Stable matching. Therefore, this result points out how the DMIRC mechanism helps to achieve (hidden) cooperation as well as coordination between agents.

Theorem 1 Let P be a responsive problem, and Γ^{SR} be the student-reduced game associated to P . Then, Γ^{SR} has a protectively sophisticated equilibrium. Moreover, for each PSE for Γ^{SR} , say δ^* , $\mu^{SR}(\delta^*)$ is the Student-Optimal Stable matching for problem P , μ^{SO} .

Although the Appendix proposes a formal proof of Theorem 1, we illustrate below why this result holds as the consequence of two facts, whose precise statement is relegated to the Appendix. These facts are:

- (a) Lemma 4: If δ^* is an equilibrium for Γ^{SR} , then $\mu^{SR}(\delta^*)$ is a stable matching for the original responsive problem P ;
- (b) Lemma 7: Given the responsive problem P , let μ^{SO} be its Student-Optimal Stable matching. Then, when protectively dominated strategies are iteratively deleted, the set of remaining strategies for each student, say Σ_i^{PF} , contains all the hospitals that are at least as preferred as her assigned hospital at μ^{SO} ; i.e. if $\delta_i \in \Sigma_i^{PF}$ then¹⁶

$$\{h_j \in H : h_j \succsim_i \mu^{SO}(s_i)\} \cap \{h_j \in H : s_i \succ_j \emptyset\} \subseteq \delta_i; \quad (1)$$

The analysis of how students faced with Γ^{SR} leads us to identify a focal point equilibrium (Schelling, 1980) where each student selects all the hospitals that are at least as preferred as that assigned to her by the Student-Optimal Stable matching; i.e. δ^* is such that for each s_i , $\delta_i^* = \{h_j \in H : h_j \succsim_i \mu^{SO}(s_i)\}$.

Note that, since the notion of protective dominance involves all the possible orderings in which the hospitals choose which students to enroll, as a consequence of Theorem 1 above, we can establish that the profile of strategies δ^* (i.e., $\delta_i^* = \{h_j \in H : h_j \succsim_i \mu^{SO}(s_i)\}$ for each student s_i) is also a focal point equilibrium for a slight perturbation of Γ^{SR} . Here, the students select the hospitals they would like to be enrolled in, having no information on the ordering in which the hospitals will choose the students they enroll. When no student plays strategies being protectively dominated (at some level), playing δ_i^* is an optimal strategy for each student, independent of the specific ordering the hospitals use to choose the students they enroll in their programs.

¹⁶ To be precise, any student is free to include in her strategy any hospital which considers that this student is inadmissible. Note that, since the hospital will never select this student, the final outcome of Γ^{SR} is not affected by the inclusion or not of the hospital in the student's strategy.

5 The Unbounded MIRC Mechanism

As Example 1 shows, the MIRC mechanism can result in unstable allocations when hospitals act strategically. This example also points out that the limit of hospitals' ability to enroll students (no hospital can enroll more students than its quota determines) can be the origin of such instability.

In this section we explore the agents' behavior when faced with the MIRC mechanism, but the ability of each student to select a hospital is not limited by its quota restriction. We call this mechanism the **Unbounded MIRC mechanism**, UMIRC mechanism henceforth. Relying on the description of the MIRC mechanism in Definition 3, the UMIRC mechanism can be introduced as a modification of the former in which it is assumed that the quota restrictions are not binding; i.e., $q_j \geq n$ for each h_j .

Definition 6 [The Unbounded MIRC Mechanism]

The UMIRC mechanism is the following sequential game, with $n + 1$ stages. The set of players is $S \cup H$. Actions are selected according to these (sequential) rules:

- (0) Each hospital h_j selects an action $\sigma_j \subseteq S$. This selection is simultaneously made by all the hospitals. This information becomes public before proceeding to the next stage.
- (1) Student s_1 , taking into account the actions already taken by all the hospitals, chooses her action $a_1 \in \{h_j \in H : s_1 \in \sigma_j\} \cup \{s_1\}$. Then, s_1 is matched to a_1 ; i.e. $\mu^{UMIRC}(s_1) = a_1$. This information becomes publicly known by the remaining students.
- ...
- (ℓ) Student s_ℓ , taking into account the actions taken by all the hospitals and the students preceding her, chooses her action $a_\ell \in \{h_j \in H : s_\ell \in \sigma_j\} \cup \{s_\ell\}$. Then, s_ℓ is matched to a_ℓ ; i.e., $\mu^{UMIRC}(s_\ell) = a_\ell$. The remaining students are publicly informed about this.

Once all the agents have selected their actions, matching μ^{UMIRC} is obtained, as described through stages (1) to (n) above.

5.1 Strategic Behavior and the UMIRC Mechanism

This section parallels the analysis developed in Section 3.1 for the MIRC mechanism. Therefore, the arguments associated to the agents' rationality when faced with the MIRC mechanism are still valid in this section that concentrates in the UMIRC mechanism.

Note that, for a given vector of actions σ selected by the hospitals, the set of available hospitals for student s_i at stage i , $i = 1, \dots, n$, is

$$\mathcal{A}^U(s_i | \sigma) = \{h_j \in H : s_i \in \sigma_j\} \cup \{s_i\},$$

which is independent of the other students' selected actions.

Any rational student selects her preferred hospital (if any) on $\mathcal{A}^U(s_i | \sigma)$; i.e., each student s_i seeking to maximize her utility, when faced with the UMIRC mechanism, chooses action $a_i^*(\sigma) = \mathcal{B}(\mathcal{A}^U(s_i | \sigma); \succ_i)$. Therefore, for any such student, $\mu^{UMIRC}(s_i) = a_i^*(\sigma)$.

Since the hospitals can anticipate each student's best-response to their actions, we can identify the analysis of the equilibrium outcomes for the UMIRC mechanism with that for the reduced game involving only the hospitals, which anticipate that students behave rationally. Therefore, as argued in Section 3.1, we only need to restrict our analysis to the hospitals' behavior when faced with the **Unbounded Hospital-Reduced game**, to be denoted as Γ^{UHR} , where hospitals are the players; the strategies for each hospital are the sets of students; and associated to each vector of strategies, $\sigma = (\sigma_1, \dots, \sigma_j, \dots, \sigma_m)$ there is a matching $\mu^{UHR}(\sigma)$ assigning student s_i to her best available option,

$$\mu^{UHR}(s_i; \sigma) = \mathcal{B}(\mathcal{A}^U(s_i | \sigma); \succ_i).$$

Theorem 2 Let P be a problem. Then, game Γ^{UHR} has an equilibrium if, and only if, $\mathcal{C}(P) \neq \emptyset$. Moreover, for each equilibrium for Γ^{UHR} , say σ^* , $\mu^{UHR}(\sigma^*) \in \mathcal{C}(P)$. Similarly, associated to each $\mu \in \mathcal{C}(P)$ there is an equilibrium for Γ^{UHR} , say σ' , such that $\mu = \mu^{UHR}(\sigma')$.

Theorem 2 above points out that stable allocations are reached without imposing any cooperation between the agents. This is, in part, derived from the sequentiality in which agents take their decisions. Therefore, we can understand the equilibria of Γ^{UHR} as the result of a non-explicit (or hidden) cooperation between agents. Nevertheless, since the equilibrium is not unique, there is the risk of utility losses induced by a (possible) coordination failure by hospitals.

6 Conclusions and Open Questions

This paper explores, from a game-theoretical point of view, the mechanism employed in Spain to match medical graduates to residency programs in private hospitals. We observe that this mechanism might select, as equilibrium outcomes, unstable matchings.

The lack of (guaranteed) stability motivates us to explore how agents behave when faced with its dual mechanism, obtained from the MIRC mechanism by swapping the roles of hospitals and students. It is remarkable that no hospital is (out of equilibrium) over-occupied unless it exhibits an irrational behavior. This is because each hospital selects the students it wants to enroll in its program, from the ones available. A detailed analysis of the agents' behavior, when faced with the DMIRC mechanism, guarantees that only stable allocations can be expected when this mechanism is adopted. Furthermore, we detect a focal point equilibrium yielding the Student-Optimal Stable matching.

The focality of this equilibrium is maintained for a wide diversity of mechanisms sharing a common structure: Students simultaneously select the hospitals whose residency programs they feel are admissible and then, hospitals sequentially choose the students they enroll from those still available. The difference between these mechanisms is determined by the order in which the hospitals make their choice, and the moment at which this order is publicized, in particular for our purpose, the specific order, and whether the students know it before selecting their strategy or it is decided once the students' strategies have been settled are irrelevant. The reason is that, when the students choose their (protectively undominated) strategies, they are protecting against any eventually "unfavorable" ordering in which the hospitals make their decisions.

For the sake of completeness, we also explore a variant of the MIRC mechanism in which residency programs over-occupancy is allowed. Its adoption guarantees that, under responsiveness, only stable allocations can be reached. Nevertheless, since the set of stable outcomes of a given problem can be large, the lack of agents' coordination when faced with the UMIRC mechanism might induce welfare losses.

A further aspect that could be relevant to clarify relies with a robustness analysis of the mechanisms. The common structure shared by the MIRC, the DMIRC and the UMIRC is that agents in a side of the market (the proposers) simultaneously select the agents in the opposite side with whom they would like to sign a contract. Then, the agents in the opposite side (the receivers) decide which contracts (if any) to sign. This sequential structure allows to guarantee that each receiver has a dominant strategy. And the essence of this strategy does not depend on the order in which the receivers evaluate the proposals: they just select their preferred proposal among the ones that are still available. Nevertheless, the order in which the receivers take their decisions might be relevant for the proposers to select which strategy to play. Therefore, the question that we deal with is the following. Imagine that the proposers do not know the exact order in which the receivers will act. Then, should the proposers' strategies depend on their uncertainty about the order in which the receivers act? Note that, when faced to the DMIRC mechanism, we argued that students' strategies should satisfy condition (1), which is independent from the order in which the hospitals select their students. Therefore, we can say that the DMIRC mechanism is robust with respect to the order in which the hospitals play, and also with respect to any misinformation by the students about this order. A similar assertion can be made for the UMIRC because for any two orders describing how the students play, say π and π' , and each strategy played by the hospitals, say σ , $\mu_{\pi}^{UHR}(\sigma) = \mu_{\pi'}^{UHR}(\sigma)$. This holds because the number of students that can be assigned to each hospital is unbounded. Nevertheless, for the case of the MIRC mechanism, since hospitals cannot admit a number of students exceeding their quota, we cannot guarantee that the hospitals strategies were independent from the order in which students act. This lack of robustness might have negative implications in terms of welfare when the hospitals exhibit some uncertainty about the exact order in which the students

will made their selection.

To conclude, we want to propose some open questions that might deserve future research.

6.1 DMIRC vs. Student-Optimal Stable Mechanism

The NRMP adopted in 1998 the Student-Optimal Stable mechanism. As Ma (2010, Theorem 6) points out, it is expected that the outcome is either unstable or it is the Hospital-Optimal Stable matching. So the hospitals' strategic behavior means that the main objective of motivating the reform undertaken in 1998 -to reach a Student-Optimal Stable matching- is barely reached. The adoption of a stable matching mechanism within this framework has a potential additional problem, which has not been considered in this paper and, as far as we know, has not been addressed yet by any author.

Sönmez (1997) pointed out that, when faced with a stable mechanism, the hospitals have two sources of strategic behavior: their preferences, and their quotas. In Ma (2010), as well as in the present paper, the quotas by the hospitals are considered as given. It could be that the results by Ma (2010) fail to be valid when the hospitals also exhibit a strategic behavior when declaring their quotas. Nevertheless, our Theorem 1 is immune to such a behavior. The reason is that in our description of the DMIRC mechanism no hospital is required to announce its quota. Note that, since the DMIRC mechanism is not a stable mechanism, Theorem 1 in Sönmez (1997) does not apply.

Therefore, the adoption of the DMIRC mechanism exhibits some advantages related to the application of the of the Student-Optimal Stable mechanism. Since there is a general difficulty of designing non-manipulable stable mechanism (Alcalde and Barberà, 1994), we focus our comparison in terms of the simplicity of the mechanisms, as well as the outcomes reached at equilibrium. Related to the simplicity of the two mechanisms, from the agents' point of view, we think that it is barely arguable that the Student-Optimal Stable mechanism is simpler than the DMIRC. Related to the properties exhibited by the equilibria outcomes of the two mechanisms, the arguments above clarify the benefits of using the DMIRC mechanism instead of the Student-Optimal Stable mechanism.

6.2 Extensions

In the case of job matching markets (Kelso and Crawford, 1982), where salaries are a relevant variable, the literature provides some results which are, in essence, similar to those obtained for matching markets. Alcalde and Revilla (1999) illustrate that the lack of sequentiality in how the agents decide might induce instability. This instability can be avoided either by assuming collusive behavior by the agents, when their actions are simultaneously selected (Alcalde and Revilla, 1999, Theorem 4.2); or by resorting to sequential decisions (Alcalde et al., 1998) and thus concentrating on Subgame Perfect Nash Equilibria. These

results are obtained under the assumption that firms' preferences satisfy gross substitutability.¹⁷ Therefore, a natural extension of our results comes from the analysis of sequential mechanisms when institutions and individuals are not only worried about their matches but also about the monetary transfers associated to their connection. In such a case, it might be relevant to study agents' behavior when gross substitutability is not fulfilled.

Hatfield and Milgrom (2005) propose a general framework where firms and workers negotiate not only about salaries but also overall contractual terms. These authors propose a notion of substitutability to guarantee the existence of stable allocations. Recently, Echenique (2012) found a close relationship between Hatfield and Milgrom's substitutability and Kelso and Crawford's gross substitutability. Therefore, this connection suggests that the results by Alcalde et al. (1998) can be adapted to the framework of matching with contracts. What remains an interesting open question is whether, in a general setting of matching with contracts, sequentiality helps to (stealthily) promote the agents' coordination when selecting their strategies.

References

- Abdulkadiroğlu, A., Pathak, P.A., Roth, A.E., 2009. Strategy-Proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match. *American Economic Review* 99, 1954 – 1978.
- Abdulkadiroğlu, A., Sönmez, T., 2003. School Choice: A Mechanism Design Approach. *American Economic Review* 93, 729 – 747.
- Alcalde, J., 1996. Implementation of Stable Solutions to Marriage Problems. *Journal of Economic Theory* 69, 240 – 254.
- Alcalde, J., Barberà, S., 1994. Top Dominance and the Possibility of Strategy-Proof Stable Solutions to Matching Problems. *Economic Theory* 4, 417 – 437.
- Alcalde, J., Dahm, M., 2016. Dual Sourcing with Price Discovery. CeDEx DP n. 2016-03, The University of Nottingham.
- Alcalde, J., Pérez-Castrillo, D., Romero-Medina, A., 1998. Hiring Procedures to Implement Stable Allocations. *Journal of Economic Theory* 84, 469 – 480.
- Alcalde, J., Revilla, P., 1999. The Role of Unions in Hiring Procedures for Job Markets. *Economics Letters* 62, 153 – 158.

¹⁷ This assumption is closely related to the substitutability condition defined in this paper (see, e.g., Roth, 1984, footnote 7).

- Alcalde, J., Romero-Medina, A., 2000. Simple Mechanisms to Implement the Core of College Admissions Problems. *Games and Economic Behavior* 31, 294 – 302.
- Austen-Smith, D., Banks, J.S., 2005. *Positive Political Theory II: Strategy and Structure*. The University of Michigan Press.
- Baron, D., Kalai, E., 1993. The Simplest Equilibrium of a Majority-Rule Division Game. *Journal of Economic Theory* 61, 290–301.
- Bernheim, B.D., 1984. Rationalizable Strategic Behavior. *Econometrica* 52, 1007 – 1028.
- Echenique, F., 2012. Contracts versus Salaries in Matching. *American Economic Review* 102, 594–601.
- Farquharson, R., 1969. *Theory of Voting*. Yale University Press, New Haven, Connecticut.
- Gale, D., Shapley, L.S., 1962. College Admissions and the Stability of Marriage. *American Mathematical Monthly* 69, 9 – 15.
- Haeringer, G., Wooders, M., 2011. Decentralized Job Matching. *International Journal of Game Theory* 40, 1 – 28.
- Hatfield, J.W., Milgrom, P.R., 2005. Matching with Contracts. *American Economic Review* 95, 913–935.
- Jayakumar, K.L., 2016. Numerical USMLE Step 1 Scores Are Still Important in Selection of Residency Applicants. *Academic Medicine* 91, 1470–1471.
- Katsurfrakis, P.J., Uhler, T.A., Jones, L.D., 2016. The Residency Application Process: Pursuing Improved Outcomes Through Better Understanding of the Issues. *Academic Medicine* 91, 1483–1487.
- Kelso, A., Crawford, V.P., 1982. Job Matching, Coalition Formation, and Gross Substitutes. *Econometrica* 50, 1483 – 1504.
- Ma, J., 1995. Stable Matchings and Rematching-Proof Equilibria in a Two-sided Matching Markets. *Journal of Economic Theory* 66, 352 – 369.
- Ma, J., 2010. The Singleton Core in the College Admissions Problem and its Application to the National Resident Matching Program (NRMP). *Games and Economic Behavior* 69, 150–164.
- Machado, M.P., Mora, R., Romero-Medina, A., 2012. Can We Infer Hospital Quality from Medical Graduates' Residency Choices? *Journal of the European Economic Association* 10, 1400–1424.

- Moulin, H., 1979. Dominance-Solvable Voting Schemes. *Econometrica* 47, 1337 – 1351.
- Pearce, D.G., 1984. Rationalizable Strategic Behavior and the Problem of Perfection. *Econometrica* 52, 1029 – 1050.
- Prober, C.G., Kolars, J.C., First, L.R., Melnick, D.E., 2016. A Plea to Reassess the Role of United States Medical Licensing Examination Step 1 Scores in Residency Selection. *Academic Medicine* 91, 12–15.
- Pycia, M., 2012. Stability and Preference Alignment in Matching and Coalition Formation. *Econometrica* 80, 323 – 362.
- Romero-Medina, A., Triossi, M., 2014. Non-Revelation Mechanisms in Many-to-One Markets. *Games and Economic Behavior* 87, 624 – 630.
- Roth, A.E., 1982. The Economics of Matching: Stability and Incentives. *Mathematics of Operations Research* 7, 617 – 628.
- Roth, A.E., 1984. Stability and Polarization of Interests in Job Matching. *Econometrica* 52, 47 – 58.
- Roth, A.E., 1985. The College Admissions Problem is not Equivalent to the Marriage Problem. *Journal of Economic Theory* 36, 277 – 288.
- Roth, A.E., 2003. The Origins, History, and Design of Resident Match. *Journal of the American Medical Association* 289, 909 – 911.
- Roth, A.E., 2008. What Have We Learned from Market Design? *The Economic Journal* 118, 285 – 310.
- Roth, A.E., Sotomayor, M.O., 1990. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press, New York.
- Schelling, T., 1980. *The Strategy of Conflict*. Harvard University Press, Cambridge, Massachusetts. Second edition.
- Sönmez, T., 1997. Manipulation via Capacities in Two-Sided Matching Markets. *Journal of Economic Theory* 77, 197–204.
- Sotomayor, M.O., 2003. Reaching the Core of the Marriage Market through a Non-Revelation Matching Mechanism. *International Journal of Game Theory* 32, 241 – 251.
- Triossi, M., 2009. Hiring Mechanisms, Application Costs and Stability. *Games and Economic Behavior* 66, 566 – 575.
- Williams, K.J., 1995. A Reexamination of the NRMP Matching Algorithm. *Academic Medicine* 70, 470–476.

APPENDIX

Proposition 1 establishes a simple way to identify some equilibria for Γ^{HR} . As a consequence of this result, each stable matching can be decentralized through some equilibrium.

Proof of Proposition 1

Let P be a problem, and μ be a stable matching for P . Now, as stated in Proposition 1, assume that each hospital h_j , when faced with Γ^{HR} , selects strategy $\sigma_j^* = \mu(h_j)$. Since μ is stable, and thus individually rational, for each student s_i , $\mu(s_i) \succsim_i s_i$. Therefore, by construction, when hospitals select the strategies above, the set of achievable options for student s_i is $\mathcal{A}(s_i | \sigma^*) = \{\mu(s_i), s_i\}$, and thus $\mu = \mu^{HR}(\sigma^*)$.

To demonstrate Proposition 1 we proceed by contradiction. Let us assume that σ^* is not an equilibrium for Γ^{HR} . Then, there should be a hospital, say h_j , and strategy for it, σ'_j such that

$$\mu^{HR}(h_j; (\sigma_{-j}^*, \sigma'_j)) \succ_j \mu^{HR}(h_j; \sigma^*) = \mu(h_j). \quad (2)$$

Since μ is individually rational, for each h_j , $Ch(\mu(h_j); \succ_j) = \mu(h_j)$. Therefore, condition (2) above implies that

$$\mu^{HR}(h_j; (\sigma_{-j}^*, \sigma'_j)) \setminus \mu^{HR}(h_j; \sigma^*) \neq \emptyset.$$

Note that each student in $\mu^{HR}(h_j; (\sigma_{-j}^*, \sigma'_j)) \setminus \mu^{HR}(h_j; \sigma^*)$ prefers her mate under $\mu^{HR}(\sigma_{-j}^*, \sigma'_j)$ rather than her mate under $\mu^{HR}(\sigma^*)$, which coincides with her mate under μ . This implies that h_j and the (non-empty) set of students $\mu^{HR}(h_j; (\sigma_{-j}^*, \sigma'_j)) \setminus \mu^{HR}(h_j; \sigma^*)$ block matching μ , contradicting that it was stable. ■

Proposition 2 establishes that, under responsiveness, each equilibrium for Γ^{HR} yields an individually rational matching with respect to the primitives of problem P .

Proof of Proposition 2

Recall that, given how the outcome of Γ^{HR} is described, no student is assigned a hospital being worse -according her preferences- than her being unmatched option. Therefore, if $\mu^{HR}(\sigma^*)$ fails to be individually rational, it must be a hospital h_j such that $Ch(\mu^{HR}(h_j, \sigma^*); \succ_j) \subsetneq \mu^{HR}(h_j, \sigma^*)$.

For h_j given, consider the alternative strategy $\sigma'_j = \{s_i \in \mu^{HR}(h_j; \sigma^*) : \{s_i\} \succ_j \emptyset\}$, i.e. σ'_j is obtained by removing from $\mu^{HR}(h_j; \sigma^*)$, the students being inadmissible for this hospital. Therefore, $\mu^{HR}(h_j, (\sigma_{-j}^*, \sigma'_j)) \subseteq \mu^{HR}(h_j, \sigma^*) \cap \{s_i \in S : \{s_i\} \succ_j \emptyset\}$.

Note that, for each student $s_i \in \sigma'_j$, $h_j \in \mathcal{A}(s_i | (\sigma_{-j}^*, \sigma'_j))$. Moreover, for each student, her set of achievable options reduces when hospitals actions are $(\sigma_{-j}^*, \sigma'_j)$ instead of σ^* ,

and $h_j \in \mathcal{A}(s_i | (\sigma_{-j}^*, \sigma_j'))$ only for students in σ_j' . This implies that $\mu^{HR}(h_j, (\sigma_{-j}^*, \sigma_j')) = \mu^{HR}(h_j, \sigma^*) \cap \{s_i \in S : \{s_i\} \succ_j \emptyset\} \succ_j \mu^{HR}(h_j, \sigma^*)$, which contradicts that σ^* is an equilibrium for Γ^{HR} . ■

Proof of Proposition 3

Let P be a problem, and μ be a stable matching for P . Let us assume that, when faced with Γ^{SR} , each student s_i selects strategy $\delta_i^* = \mu(s_i)$. Since μ is pairwise stable, and thus individually rational, for each hospital h_j , $\mu(h_j) = Ch(\mu(h_j); \succ_j)$. By construction, when students select the strategies above, the set of achievable options for hospital h_j is $\mathcal{A}^D(h_j | \delta^*) = \mu(h_j)$, and thus $\mu = \mu^{SR}(\delta^*)$.

To prove Proposition 3 we proceed by contradiction. Let us assume that δ^* above is not an equilibrium for Γ^{SR} . Then, there should be a student s_i , and strategy for her, δ_i' such that

$$\mu^{SR}(s_i; (\delta_{-i}^*, \delta_i')) \succ_i \mu^{SR}(s_i; \delta^*) = \mu(s_i). \quad (3)$$

Since μ is individually rational, and thus $\mu(s_i) \succ_i s_i$, by transitivity $\mu^{SR}(s_i; (\delta_{-i}^*, \delta_i')) \in H$. Let $h_i = \mu^{SR}(s_i; (\delta_{-i}^*, \delta_i'))$. Since $s_i \in \mu^{SR}(h_j; (\delta_{-i}^*, \delta_i'))$, $s_i \in Ch(\mathcal{A}^D(h_j | \delta^*) \cup \{s_i\}; \succ_j)$. Therefore, the pair (s_i, h_j) blocks μ , contradicting the pairwise stability of μ . ■

Note that Corollary 1 follows because any stable matching is also pairwise stable.

Proving Theorem 1

We now deal with proving Theorem 1. As previously anticipated in Section 4.1, to prove this theorem we need some auxiliary, intermediate results. In what follows, we consider a given problem P satisfying responsiveness.

Prior to introducing the auxiliary results, we introduce some useful notation. Let us assume that each student s_i has to select a strategy in a restricted set $\Sigma_i' \subseteq \Sigma_i$, so that each profile of strategies is in the Cartesian product $\Sigma' = \prod_{i=1}^n \Sigma_i'$. Then, hospital h_j is available for s_i if there is a profile of strategies $\delta \in \Sigma'$ such that $\mu^{SR}(s_i; \delta) = h_j$. $\mathcal{O}(s_i; \Sigma')$ denotes the set of available hospitals for student s_i , if any, when playing Γ^{SR} under the restriction that the profiles of strategy belong to Σ' .

Our first result establishes that, under responsiveness, any equilibrium for Γ^{SR} yields a stable matching.

Lemma 4 Let P be a responsive problem, and δ^* be an equilibrium for Γ^{SR} . Then $\mu^{SR}(\delta^*) \in \mathcal{C}(P)$.

Proof

Assume that δ^* is an equilibrium for Γ^{SR} and, contrary to our statement, $\mu^{SR}(\delta^*)$ is unstable for P .

First, note that $\mu^{SR}(\delta^*)$ is individually rational. Since, by construction, each hospital selects its best set of available students, then any lack of individual rationality must come from some student. Assume that there is a student s_i such that $s_i \succ_i \mu^{SR}(s_i; \delta^*)$. Then, by selecting $\delta'_i = \emptyset$ she guarantees that $s_i = \mu^{SR}(s_i; (\delta^*_{-i}, \delta'_i)) \succ_i \mu^{SR}(s_i; \delta^*)$, which contradicts that δ^* is an equilibrium.

Since P is responsive, by Lemma 1, $\mu^{SR}(\delta^*)$ must be pairwise unstable. Therefore, since $\mu^{SR}(\delta^*)$ is individually rational, there is a student-hospital pair (s_i, h_j) blocking this matching; i.e. $h_j \succ_i \mu^{SR}(s_i; \delta^*)$, and $s_i \in Ch(\mu^{SR}(h_j; \delta^*) \cup \{s_i\}; \succ_j)$. Assume that, given the strategies played by the remaining students, s_i plays $\delta'_i = \{h_\ell \in H : h_\ell \succ_i h_j\}$, and consider the following cases, that exhaust all the possibilities:

(i) $\mu^{SR}(s_i; \delta^*) = s_i$. This implies that, when applying the DMIRC mechanism, and the actions chosen by the students are δ^* , no hospital has selected s_i . When s_i selects δ'_i instead of δ^*_i , there are two options.

- (1) If s_i is an achievable student for h_j then no hospital preceding h_j has selected s_i . This implies that for any $k < j$, $\mu^{SR}(h_k; \delta^*) = \mu^{SR}(h_k; (\delta^*_{-i}, \delta'_i))$. When h_j has to choose its action, given that $h_j \in \delta'_i$ and $s_i \in Ch(\mu^{SR}(h_j; \delta^*) \cup \{s_i\}; \succ_j)$, h_j includes s_i among the students it enrolls, that is $s_i \in a_j^{D^*}(\delta^*_{-i}, \delta'_i)$, and thus $h_j = \mu^{SR}(s_i; (\delta^*_{-i}, \delta'_i))$. This means that δ^* is not an equilibrium.
- (2) If s_i is not an achievable student for h_j then, when applying the DMIRC mechanism and s_i chooses action δ'_i instead of δ^*_i , there is some hospital h_k , preceding h_j , such that $s_i \in a_k^{D^*}(\delta^*_{-i}, \delta'_i)$. Therefore, $h_k = \mu^{SR}(s_i; (\delta^*_{-i}, \delta'_i))$. Since $h_k \in \delta'_i \setminus \{h_j\}$, then $h_k \succ_i h_j$. Thus, by transitivity, $\mu^{SR}(s_i; (\delta^*_{-i}, \delta'_i)) = h_k \succ_i s_i = \mu^{SR}(s_i; \delta^*)$, which means that δ^* is not an equilibrium.

(ii) There is a successor for h_j in the DMIRC mechanism, say h_ℓ with $\ell > j$, such that $\mu^{SR}(s_i; \delta^*) = h_\ell$. Note that, since $s_i \in Ch(\mu^{SR}(h_j; \delta^*) \cup \{s_i\}; \succ_j)$, it must be the case that $h_j \notin \delta^*_i$. Therefore, the arguments provided in case (i) above can be replicated to prove that δ^* cannot be an equilibrium for Γ^{SR} .

(iii) There is a hospital h_ℓ , preceding h_j in the DMIRC mechanism, such that $\mu^{SR}(s_i; \delta^*) = h_\ell$. Since $h_j \succ_i h_\ell$, when s_i is faced with the DMIRC and selects δ'_i instead of δ^*_i , $s_i \notin \mathcal{A}^D(h_\ell | (\delta^*_{-i}, \delta'_i))$, and thus $h_\ell \neq \mu^{SR}(s_i; (\delta^*_{-i}, \delta'_i))$.

It is straightforwardly verifiable that for each $k < j$

$$\mathcal{A}^D(h_k | (\delta^*_{-i}, \delta'_i)) \setminus \{s_i\} \subseteq \mathcal{A}^D(h_k | \delta^*) \setminus \{s_i\}.$$

Moreover, since (a) $h_j \in \delta'_i$, (b) $s_i \in Ch(\mu^{SR}(h_j; \delta^*) \cup \{s_i\}; \succ_j)$, and (c) preferences \succ_j are responsive, then either $s_i \in \mu^{SR}(h_j; (\delta^*_{-i}, \delta'_i))$, which contradicts that δ^* is an equilibrium; or $s_i \notin \mathcal{A}^D(h_j | (\delta^*_{-i}, \delta'_i))$. Note that, the latter case implies that there is

some hospital $h_r \in \delta'_i \setminus \{h_j\}$ such that $s_i \in a_r^{D^*}(\delta_{-i}^*, \delta'_i) = \mu^{SR}(h_r; (\delta_{-i}^*, \delta'_i))$. Since $h_r \succ_i h_j$, this means that δ^* is not an equilibrium. ■

We now deal with the analysis of which strategies are protectively dominated for each student. Lemma 5 is very intuitive. It establishes that no student should include in her strategy an inadmissible hospital. Then, Lemma 6 establishes that for each student, any strategy not including her best achievable hospital is protectively dominated. Finally, Lemma 7 extends the idea of Lemma 6 to any ulterior set of profiles of strategies that are not erased based on protective domination arguments.

From now on, for each student s_i , her set of strategies Σ_i is the set of all the possible subsets of hospitals. $\Sigma = \prod_{i=1}^n \Sigma_i$ is the set of profiles of strategies.

Lemma 5 Let P be a responsive problem, and $h_j \in \mathcal{O}(s_i; \Sigma)$ an achievable hospital for student s_i such that $s_i \succ_i h_j$, when playing Γ^{SR} . Then, any strategy $\delta_i \in \Sigma_i$ such that $h_j \in \delta_i$ is protectively dominated.

Proof

Given strategy δ_i containing some achievable hospital being inadmissible for student s_i , let δ'_i the strategy for s_i containing only the hospitals in δ_i being admissible for this student; i.e., $\delta'_i = \{h_j \in \delta_i : h_j \succ_i s_i\}$. Note that for any permutation of the hospitals π , and any strategy for the remaining students, say δ_{-i} ,

(i) if $\mu_\pi^{SR}(s_i, \delta) \succsim_i s_i$, then $\mu_\pi^{SR}(s_i, (\delta_{-i}, \delta'_i)) = \mu_\pi^{SR}(s_i, \delta)$; and

(ii) if $s_i \succ_i \mu_\pi^{SR}(s_i, \delta)$, then $\mu_\pi^{SR}(s_i, (\delta_{-i}, \delta'_i)) \succsim_i s_i \succ_i \mu_\pi^{SR}(s_i, \delta)$.

To conclude the proof, we just need to show that there is a profile of strategies, in which s_i plays δ_i , and permutation π , such that Γ_π^{SR} assigns to s_i an inadmissible hospital.

Let assume that h_r is an achievable hospital for student s_i such that $s_i \succ_i h_r$. Since $h_r \in \mathcal{O}(s_i; \Sigma)$, there is a profile of strategies $\delta \in \Sigma$ such that $h_r = \mu^{SR}(s_i; \delta)$, and thus $s_i \in a_r^{D^*}(\delta)$. Since h_r preferences are responsive, this implies that $s_i \succ_h \emptyset$. Now, assume that each student $s_\ell \neq s_i$ selects a strategy δ'_ℓ not including h_r . This implies that, for any permutation of the hospitals, say π , $\mu_\pi^{SR}(h_r; (\sigma'_{-i}, \sigma_i)) \subseteq \{s_i\}$. Moreover, for any permutation π' such that $\pi'(r) = 1$, i.e. h_r is the first hospital to choose in the DMIRC mechanism, $\mu_{\pi'}^{SR}(s_i; (\sigma'_{-i}, \sigma_i)) = h_r$. ■

Let Σ_i^0 denote the restricted set of strategies for student s_i not containing any achievable hospital that she considers inadmissible. The set of profiles of strategies where each student's strategy is in her restricted set of strategies Σ_i^0 is denoted by Σ^0 .

Lemma 6 Let P be a responsive problem. Assume that students play game Γ^{SR} , and each student s_i is restricted to select strategies in Σ_i^0 . Then any strategy δ_i such that $\mathcal{B}(\mathcal{O}(s_i; \Sigma^0); \succ_i) \notin \delta_i$ is protectively dominated for s_i .

Proof

Assume that student s_i selects strategy δ_i not including $\mathcal{B}(\mathcal{O}(s_i; \Sigma^0); \succ_i)$. Note that, in particular, this implies that there is at least one hospital that s_i considers admissible.¹⁸ Let us assume that $h_j = \mathcal{B}(\mathcal{O}(s_i; \Sigma^0); \succ_i)$, and let denote $\delta'_i = \delta_i \cup \{h_j\}$. Since $\delta_i \in \Sigma_i^0$, then $\delta'_i \in \Sigma_i^0$. We now demonstrate that δ'_i protectively dominates δ_i .

Note that the only difference between strategies δ_i and δ'_i is the inclusion or not of h_j as an admissible hospital for student s_i . Therefore, for any strategy played by the other students, say δ_{-i} , and each permutation of the hospitals, say π , either $\mu_\pi^{SR}(s_i; \delta) = \mu_\pi^{SR}(s_i; (\delta_{-i}, \delta'_i))$ or $\mu_\pi^{SR}(s_i; (\delta_{-i}, \delta'_i)) = h_j \succ_i \mu_\pi^{SR}(s_i; \delta)$. Thus, we just need to show that there is a permutation π and some profile of strategies for the students other than s_i , say δ'_{-i} , under which s_i finds profitable to include h_j as an admissible hospital.

Let δ'_{-i} be a profile of strategies for students other than s_i such that, for all $s_\ell \neq s_i$, $h_j \notin \sigma'_\ell \in \Sigma_\ell^0$. Let π a permutation of hospitals satisfying that $\pi(1) = j$. Since h_j is the first hospital to choose when the order is determined by π , $\mu_\pi^{SR}(h_j; (\delta'_{-i}, \delta_i)) = \emptyset$, whereas $\mu_\pi^{SR}(h_j; \delta') = s_i$. Therefore, since $h_j = \mathcal{B}(\mathcal{O}(s_i; \Sigma^0); \succ_i)$, $\mu_\pi^{SR}(s_i; \delta') = h_i \succ_i \mu_\pi^{SR}(s_i; (\delta'_{-i}, \delta_i))$. ■

Let $\Sigma_i^1 \subseteq \Sigma_i^0$ denote the restricted set of strategies for student s_i where each strategy includes as acceptable the best achievable hospital for this student. Σ^1 denotes the profiles of strategies δ such that for each student s_i , $\delta_i \in \Sigma_i^1$.

The idea underlying Lemma 6 above can be replicated in an iterative way; i.e., we can describe for student s , Σ_i^2 as her set of protectively undominated strategies, provided that each student s_ℓ is restricted to select strategies in Σ_ℓ^1 ; and so forth. In general, for any $t > 1$, Σ_i^t denotes the protectively undominated strategies for student s_i provided that the remaining students are selecting strategies in $\Sigma_{-i}^{t-1} = \prod_{\ell \neq i} \Sigma_\ell^{t-1}$.

The following result proposes a generalization for Lemma 6.

Lemma 7 Let P be a responsive problem. Assume that students play Γ^{SR} , and each student s_i is restricted to select strategies in Σ_i^t . Then any strategy δ_i such that $\mathcal{B}(\mathcal{O}(s_i; \Sigma^t); \succ_i) \notin \delta_i$ is protectively dominated for s_i .

Proof

Note that the arguments provided in the proof of Lemma 6 above can be adapted to prove this result.

Assume that $h_j = \mathcal{B}(\mathcal{O}(s_i; \Sigma^t); \succ_i) \notin \delta_i$, and define $\delta'_i = \delta_i \cup \{h_j\}$. As argued in the proof of Lemma 6, for each π and any $\delta_{-i} \in \Sigma_{-i}$, $\mu_\pi^{SR}(s_i; (\delta_{-i}, \delta'_i)) \in \{\mu_\pi^{SR}(s_i; \delta); h_j\}$.

Now, consider strategies for students other than s_i , say δ'_ℓ , such that for each $s_\ell \neq s_i$, $h_j \in \delta'_\ell$ if, and only if, $h_j \in \delta_\ell$ for all $\delta_\ell \in \Sigma_\ell^t$.¹⁹ Then, for any π such that $\pi(1) = h_j$, $\mu_\pi^{SR}(s_i; \delta') = h_j \succ_i \mu_\pi^{SR}(s_i; (\delta'_{-i}, \delta_i))$. ■

¹⁸ Otherwise, Σ_i^0 must only contain the strategy $\delta_i = \emptyset$, and thus $\mathcal{B}(\mathcal{O}(s_i; \Sigma^0); \succ_i) = \emptyset$.

¹⁹ That is to include h_j as an acceptable hospital is a necessary condition for δ'_ℓ to be in Σ_ℓ^t .

Recall that, under responsiveness, the Student-Optimal Stable matching is (weakly) efficient from the students' point of view (Roth, 1985, Theorem 2*). Therefore, the process of an iterative elimination of protectively dominated strategies by applying Lemma 7 above stops at the step T ,²⁰ where for each student s_i all the strategies in Σ_i^T must contain $\mu^{SO}(s_i)$, and eventually some hospitals being preferred to $\mu^{SO}(s_i)$ by s_i . In fact, we do not lose generality when assuming that the iterative application of Lemma 7 allows us to describe for each student s_i a set of strategies Σ_i^T at which no strategy can be protectively dominated, where for each $\sigma_i \in \Sigma_i^T$,

$$\{h_j \in H : h_j \succsim_i \mu^{SO}(s_i)\} \subseteq \delta_i \subseteq \{h_j \in H : h_j \succ_i s_i\}. \quad (4)$$

As the following example illustrates, condition (4) is not necessary to describe the 'limit' set Σ^T , because it might include, for some students, more hospitals than a necessary condition-derived from the iterative application of Lemma 7- imposes. Nevertheless, the unnecessary inclusion of some hospitals in a strategy to satisfy condition (4) above has no effect on the outcome of the game Γ^{SR} .

Example 3 Consider the following 6-student-3-hospital responsive problem, with $S = \{a, b, c, d, e, f\}$, $H = \{1, 2, 3\}$, with quotas $q_1 = 1$, and $q_2 = q_3 = 2$. Agents' preferences are

$$\begin{array}{lll} \succ_a := 1, 2, 3 & \succ_d := 1, 2, 3 & \succ_1 := d, f, b, e, c \\ \succ_b := 2, 1, 3 & \succ_e := 1, 2, 3 & \succ_2 := c, a, f, b, d \\ \succ_c := 1, 2, 3 & \succ_f := 2, 1, 3 & \succ_3 := a, b, d, e, f \end{array}$$

The Student-Optimal Stable matching for P is

$$\mu^{SO} := \begin{pmatrix} a & b & c & d & e & f \\ 2 & 3 & 2 & 1 & 3 & - \end{pmatrix}$$

Condition (4) above, applied to f establishes that this student must select strategy $\delta_f^* = \{1, 2, 3\}$. Nevertheless, the iterative deletion of protectively dominated strategies allows student f to choose either δ_f^* above or $\delta_f' = \{2, 3\}$. Note that, when describing Σ^1 , student d must include 1 among her acceptable hospitals; i.e., $\{1\} \subseteq \delta_d$ for each $\delta_d \in \Sigma_d^1$. Since d is the preferred student for hospital 1, and this hospital is the first to select its students, when the profiles of preferences have to be in Σ^1 , 1 becomes an unachievable hospital for all the students other than d . This implies that for any $t > 1$, and any student other than d , it is irrelevant that the student, when playing a strategy in Σ_i^t , includes 1 as an admissible hospital or not.

Similarly, when considering condition (4), student a plays either strategy $\delta_a = \{1, 2, 3\}$ or $\delta_a' = \{1, 2\}$. Nevertheless, when applying Lemma 6 and then iteratively Lemma 7 nothing

²⁰ I.e., by applying Lemma 7 we have that $\Sigma^{T+1} = \Sigma^T \subsetneq \Sigma^{T-1}$.

is required about including hospital 1 as an acceptable hospital for student a . Note that, since $\emptyset \succ_1 a$, hospital 1 is not available for student a , and thus it is irrelevant, for this student, to consider whether to include 1 as an admissible hospital or not. ■

Note that for each student s_i , and strategy for her $\delta_i \in \Sigma_i^T$, if δ_i does not satisfy condition (4) above is because the hospitals that are not listed in δ_i , but listed in any strategy fulfilling (4), do not affect the final outcome as it is the cases of hospital 1 for either student f or student a in the example above. This is why we say that no generality is lost when identifying for any student a strategy in Σ_i^T and the requirement that condition (4) is fulfilled.

We can now provide a formal proof of Theorem 1.

Proof of Theorem 1

Consider a responsive problem P , and let μ^{SO} be its Student-Optimal Stable matching. Let us assume that, when playing Γ^{SR} , each student selects strategy $\delta_i^* = \{h_j \in H : h_j \succsim_i \mu^{SO}(s_i)\}$. Since each student's strategy satisfies condition (4), none of them is employing a protectively dominated strategy. It can also be seen that δ^* constitutes an equilibrium for Γ^{SR} , with $\mu^{SR}(\delta^*) = \mu^{SO}$.

Now, let us consider a protectively sophisticated equilibrium, say δ' , with $\mu^{SR}(\delta') \neq \mu^{SR}(\delta^*) = \mu^{SO}$. Since σ' is a PSE, without loss of generality, we can assume that for each $s_i \in S$, $\delta_i^* \subseteq \delta'_i$. Moreover, by Lemma 4, $\mu^{SR}(\delta')$ is stable for P . Therefore, by Lemma 2, for each student s_i , either (a) $\mu^{SR}(s_i; \delta') = \mu^{SO}(s_i)$, or (b) $\mu^{SO}(s_i) \succ_i \mu^{SR}(s_i; \delta')$, where condition (b) is fulfilled by, at least, one student.

When comparing matchings μ^{SO} and $\mu^{SR}(\delta')$, since the two matchings are stable, it is well-known that for each $s_i \in S$, $\mu^{SO}(s_i) \in H$ if, and only if $\mu^{SR}(s_i; \delta') \in H$, and for each s_i fulfilling condition (b) above, $\mu^{SR}(h_j; \delta') \succ_j \mu^{SO}(h_j)$, for $h_j = \mu^{SO}(s_i)$.²¹

Let h_{j_1} be the hospital such that $\mu^{SO}(h_{j_1}) \neq \mu^{SR}(h_{j_1}; \delta')$, and for each $j < j_1$, $\mu^{SO}(h_j) = \mu^{SR}(h_j; \delta')$; and let $s_{i_1} \in \mu^{SR}(h_{j_1}; \delta') \setminus \mu^{SO}(h_{j_1})$. Since $\mu^{SO}(h_{i_1}) \succ_{i_1} \mu^{SR}(h_{i_1}; \delta')$, it must be the case that $\delta_{i_1}^* \subsetneq \delta'_{i_1}$.

Since δ' is an equilibrium,

$$\mu^{SR}(s_{i_1}; (\delta'_{-i_1}, \delta_{i_1}^*)) = s_{i_1}. \quad (5)$$

Note that, otherwise, $\mu^{SR}(s_{i_1}; (\delta'_{-i_1}, \delta_{i_1}^*)) \succ_{i_1} \mu^{SO}(s_{i_1}) \succ_{i_1} \mu^{SR}(s_{i_1}; \delta')$, showing that δ' is not a SPE.

Note that, for each $j' < j_1$, $\mu^{SR}(h_{j'}; (\delta'_{-i_1}, \delta_{i_1}^*)) = \mu^{SO}(h_{j'}) = \mu^{SR}(h_{j'}; \delta')$. Moreover, all the students in $\mu^{SO}(h_{j_1})$ are achievable by h_{j_1} when applying the DMIRC mechanism and the students' actions are $(\delta'_{-i_1}, \delta_{i_1}^*)$. Therefore, since \succ_{j_1} satisfy substitutability, for each $s_{i'} \in \mu^{SR}(h_{j_1}; (\delta'_{-i_1}, \delta_{i_1}^*)) \setminus \mu^{SO}(h_{j_1})$ and any $s_{i''} \in \mu^{SO}(h_{j_1}) \setminus \mu^{SR}(h_{j_1}; (\delta'_{-i_1}, \delta_{i_1}^*))$, $s_{i'} \succ_{j_1} s_{i''}$. Thus, for each $j'' > j_1$ the following holds.

²¹ See Theorems 2.22 and 5.12 in Roth and Sotomayor (1990).

- (a) If $\mu^{SR}(h_k; (\delta'_{-i_1}, \delta^*_{i_1})) \cap \mu^{SR}(h_{j''}; \delta') = \emptyset$ for each $k < j''$, then $\mu^{SR}(h_{j''}; (\delta'_{-i_1}, \delta^*_{i_1})) = \mu^{SR}(h_{j''}; \delta')$; and
- (b) there is at most one student $s_{i'} \in \mu^{SR}(h_{j''}; \delta') \setminus \mu^{SR}(h_{j''}; (\delta'_{-i_1}, \delta^*_{i_1}))$. Moreover, if such a student exists, then $\mu^{SR}(s_{i''}; (\delta'_{-i_1}, \delta^*_{i_1})) = h_k$ for some $k < j''$.

Therefore, by (5), there should be a hospital $h_{j'}$, with $j' > j_1$, such that either $|\mu^{SR}(h_{j'}; (\delta'_{-i_1}, \delta^*_{i_1}))| < |\mu^{SR}(h_{j'}; \delta')| \leq q_{j'}$ or there is a student $s_{i'} \in \mu^{SR}(h_{j'}; (\delta'_{-i_1}, \delta^*_{i_1}))$ such that $\mu^{SO}(s_{i'}) = s_{i'}$. Thus, there must be a student $s_{i_2} \in \mu^{SO}(h_{j'}) \setminus \mu^{SR}(h_{j'}; (\delta'_{-i_1}, \delta^*_{i_1}))$ such that $s_{i_2} \in Ch(\mu^{SR}(h_{j'}; (\delta'_{-i_1}, \delta^*_{i_1})) \cup \{w_{i_2}\}; \succ_{j'})$.

Hence, by (b) above, there must be $j_2 < j'$ such that $\mu^{SR}(s_{i_2}; \delta') = h_{j_2}$, and thus $h_{j'} \succ_{i_2} h_{j_2}$. Therefore, $\mu^{SR}(s_{i_2}; (\delta'_{-i_2}, \delta^*_{i_2})) \neq h_{j_2}$. Consider the following two cases, that exhaust all the possibilities,

- (i) $h_{j_2} = h_{j_1}$. Then $\mu^{SR}(h_{j_1}; (\delta'_{-i_2}, \delta^*_{i_2})) \setminus \mu^{SR}(h_{j_1}; \delta') = \mu^{SR}(h_{j_1}; (\delta'_{-i_1}, \delta^*_{i_1})) \setminus \mu^{SR}(h_{j_1}; \delta')$. Therefore, for each hospital h_j , $j_1 < j < j'$, $\mathcal{A}^D(h_j | (\delta'_{-i_1}, \delta^*_{i_1})) = \mathcal{A}^D(h_j | (\delta'_{-i_2}, \delta^*_{i_2}))$, and thus $h_{j'} = \mu^{SR}(s_{i_2}; (\delta'_{-i_2}, \delta^*_{i_2})) = \mu^{SO}(s_{i_2}) \succ_{i_2} \mu^{SR}(s_{i_2}; \delta')$. This shows that δ' is not a SPE.
- (ii) $j_1 < j_2 < j'$. Since δ' is a SPE, then $\mu^{SR}(s_{i_2}; (\delta'_{-i_2}, \delta^*_{i_2})) = s_{i_2}$. Now, the arguments related to h_{j_1} can be replicated for h_{j_2} .

Since the number of hospitals is finite, an iterative argument necessarily yields the existence of a student, s_{i_k} such that $\mu^{SO}(s_{i_k}) \succ_{i_k} \mu^{SR}(s_{i_k}; \delta')$, and $\mu^{SR}(s_{i_k}; (\delta'_{-i_k}, \delta^*_{i_k})) = \mu^{SO}(s_{i_k})$, showing that δ' is not a SPE. ■

Proof of Theorem 2

Let P be a problem, and σ^* an equilibrium for Γ^{UHR} . Let us assume that $\mu^{UHR}(\sigma^*)$ fails to be stable for P .

First, note that $\mu^{UHR}(\sigma^*) \in \mathcal{JR}(P)$. Since each student always selects her best achievable option, $\mu^{UHR}(s_i; \sigma^*) \succeq_i s_i$ for each student s_i . Assume that there is some hospital, say h_j , such that $Ch(\mu^{UHR}(h_j; \sigma^*); \succ_j) \neq \mu^{UHR}(h_j; \sigma^*)$. Then, when h_j plays strategy $\sigma'_j = Ch(\mu^{UHR}(h_j; \sigma^*); \succ_j)$ instead of σ^*_j , it gets

$$\mu^{UHR}(h_j; (\sigma^*_{-j}, \sigma'_j)) = Ch(\mu^{UHR}(h_j; \sigma^*); \succ_j) \succ_j \mu^{UHR}(h_j; \sigma^*),$$

showing that σ^* is not an equilibrium.

Since $\mu^{UHR}(\sigma^*)$ is an individually rational matching for P , if it fails to be stable, there should be a hospital, say h_j and a set of students, $S' \subseteq S$, blocking $\mu^{UHR}(\sigma^*)$. Let us define

$\sigma'_j = Ch(\mu^{UHR}(h_j; \sigma^*) \cup S'; h_j)$. Note that, since (S', h_j) blocks $\mu^{UHR}(\sigma^*)$, for any $s_i \in S'$, $h_j \succ_i \mu^{UHR}(s_i; \sigma^*)$. Therefore, when faced with the UMIRC mechanism, for each $s_i \in \sigma'_j$, $a_i^*(\sigma^*, \sigma'_j) = \mathcal{B}(\mathcal{A}^U(s_i | (\sigma^*, \sigma'_j)); \succ_i) = h_j$, and thus $\mu^{UHR}(h_j; (\sigma^*, \sigma'_j)) = \sigma'_j \succ_j \mu^{UHR}(h_j; \sigma^*)$, showing that σ^* is not an equilibrium. Note that, in particular, it is also proven that when $\mathcal{C}(P) = \emptyset$, no profile of strategies σ^* constitutes an equilibrium for Γ^{UHR} .

Now, let us consider a stable matching for P , say μ . Note that, when faced with the UMIRC mechanism, if each hospital h_j chooses action $\sigma'_j = \mu(h_j)$, then the set of available options for student s_i is $\mathcal{A}^U(s_i | \sigma') = \mu(s_i) \cup \{s_i\}$. Therefore, since μ is stable and thus individually rational, for each student s_i , $\mathcal{B}(\mathcal{A}^U(s_i | \sigma'); \succ_i) = \mu(s_i)$. Hence, for $\sigma' = (\sigma'_1, \dots, \sigma'_j, \dots, \sigma'_m)$, $\mu^{UHR}(\sigma') = \mu$.

Assume that σ' above is not an equilibrium for Γ^{UHR} . Then, there must be a hospital, say h_ℓ , and strategy for it σ''_ℓ such that $\mu^{UHR}(h_\ell; (\sigma'_\ell, \sigma''_\ell)) \succ_\ell \mu^{UHR}(h_\ell; \sigma') = \mu(h_\ell)$. Since μ is individually rational, this implies that there is a non-empty set of students $S' = \mu^{UHR}(h_\ell; (\sigma'_\ell, \sigma''_\ell)) \setminus \mu^{UHR}(h_\ell; \sigma')$ such that for each $s_i \in S'$, $h_\ell \succ_i \mu(s_i)$.²² This implies that (S', h_ℓ) blocks μ , contradicting its stability. ■

²²Note that h_ℓ was not achievable for students in S' when faced with the UMIRD mechanism and each hospital h_j action was σ'_j . Nevertheless, when the action by h_ℓ shifts to σ''_ℓ , h_ℓ becomes the best achievable option for students in S' .