Online Appendix

Near-Feasible Stable Matching with Couples

Thành Nguyen and Rakesh Vohra

Preferences and Stability

A1. Preferences

Doctor's preferences over hospitals are based on a model fitted from Israeli hospital preference data (see Kelner (2015)). We first generate a number p_h representing the 'popularity' of a hospital h, where $p_h = 0.99K(0.8)^{X_h} + 0.01 \times 18$, X_h is an integer chosen i.i.d. uniformly from 1 to 18, and K is the total number of slots, which is equal to the number of doctors.

Preferences of each single doctor are generated by selecting hospitals iteratively at random without replacement. At each iteration, the probability of selecting hospital h from among those that remain is proportional to p_h.

To generate the preferences of the couples, we assume that couples would like to be allocated to hospitals in the same region rather than different regions. So, we choose lotteries over ordered pairs of hospitals with the property that pairs in the same region are favored over pairs in different regions.

Choose $\lambda \in (0, 1)$ and set

$$\nu_{h,h'} = \begin{cases} \lambda p_h p_{h'} & \text{if hospital } h, h' \text{ are in the same region,} \\ (1-\lambda) p_h p_{h'} & \text{otherwise.} \end{cases} \tag{A1}$$

Preferences of each couple are generated by selecting ordered pairs of hospitals iteratively at random without replacement. At each iteration, the probability of selecting the ordered pair (h, h') from among the ordered pairs that remain is proportional to $\nu_{h,h'}$.

When λ is close to 1, hospitals in the same region are more likely to be at the top of a couple's preference ordering. If $\lambda < 0.5$, then couples prefer not to be in the same region. For the results reported we set $\lambda = 0.7$.

To generate the preferences of a couple over all hospital pairs including the outside option, we order all pairs of form $(h_1, \emptyset)$ or $(\emptyset, h_2)$ uniformly at random. Finally we construct the full preference ordering so that it is 'unemployment-averse', i.e.

$$\begin{aligned} &(h_1, h_2) \succ (h_3, \emptyset) \succ (\emptyset, \emptyset) \\ &(h_1, h_2) \succ (\emptyset, h_3) \succ (\emptyset, \emptyset) \end{aligned} \tag{A2}$$

for any $h_1, h_2, h_3 \in H$.

A2. Stability

Let H be the set of hospitals, D^1 the set of single doctors, and D^2 the set of couples. Each couple $c \in D^2$ is denoted $c = (f, m)$ where f_c and m_c are the first and second member of c, respectively. The set of all doctors, D, is given by $D^1 \cup \{m_c | c \in D^2\} \cup \{f_c | c \in D^2\}$.

Each single doctor $d \in D^1$ has a strict preference ordering $\succ_d$ over $H \cup \{\emptyset\}$ where $\emptyset$ denotes the outside option for each doctor. If $h \succ_d \emptyset$, we say that hospital h is acceptable for d. Each couple $c \in D^2$ has a strict preference ordering $\succ_c$ over $H \cup \{\emptyset\} \times H \cup \{\emptyset\}$–i.e., over pairs of hospitals, including the outside option.

Each hospital $h \in H$ has a fixed capacity $k_h > 0$. The preference of a hospital h over subsets of D is summarized by h's choice function $ch_h(.) : 2^D \to 2^D$. While a choice function can be associated with every strict preference ordering over subsets of D, the converse is not true. The information contained in a choice function is only sufficient to recover a partial order over the subsets of D. Therefore, it isn't always possible to say whether a hospital prefers a couple over some pair of single doctors.

We assume, as is standard in the literature, that $ch_h(.)$ is responsive. This means that h has a strict priority ordering $\succ_h$ over elements of $D \cup \{\emptyset\}$. If $\emptyset \succ_h d$, we say d is not acceptable to h. For any set $D^* \subset D$, hospital h's choice from that subset, $ch_h(D^*)$, consists of the (up to) k_h highest priority doctors among the acceptable doctors in D^*. Formally, $d \in ch_h(D^*)$ if and only if $d \in D^*; d \succ_h \emptyset$ and there exists no set $D' \subset D^* \setminus \{d\}$, such that $|D'| = k_h$ and $d' \succ_h d$ for all $d' \in D'$.

A matching μ is an assignment of each single doctor to a hospital or his/her outside option, an assignment of couples to at most two positions (in the same or different hospitals) or their outside option, such that the total number of doctors assigned to any hospital h does not exceed its capacity k_h. Given matching μ, let μ_h denote the subset of doctors matched to h; μ_d and μ_{f_c}, μ_{m_c} denote the position(s) that the single doctor d, and the female and male members of the couple c obtain in the matching, respectively.

We say μ is *individually rational* if $ch_h(\mu_h) = \mu_h$ for any hospital h; $\mu_d \succeq_d \emptyset$ for any single doctor d and $(\mu_{f_c}, \mu_{m_c}) \succeq_c (\emptyset, \mu_{m_c})$; $(\mu_{f_c}, \mu_{m_c}) \succeq_c (\mu_{f_c}, \emptyset)$; $(\mu_{f_c}, \mu_{m_c}) \succeq_c (\emptyset, \emptyset)$ for any couple c.

Roth and Sotomayor (1992), we list the ways in which different small coalitions can block a matching μ.

DEFINITION 2: *The following are called blocking coalitions for a matching μ.*

1) *A pair $d \in D^1$ and $h \in H$ can block μ if $h \succ_d \mu(d)$ and $d \in ch_h(\mu(h) \cup d)$.*

2) *A triple $(c, h, h') \in D^2 \times (H \cup \{\emptyset\}) \times (H \cup \{\emptyset\})$ with $h \neq h'$ can block μ if $(h, h') \succ_c \mu(c)$, $f_c \in ch_h(\mu(h) \cup f_c)$ when $h \neq \emptyset$ and $m_c \in ch_{h'}(\mu(h') \cup m_c)$ when $h' \neq \emptyset$.*

3) A pair $(c, h) \in D^2 \times H$ can block μ if $(h, h) \succ_c \mu(c)$ and $(f_c, m_c) \subseteq ch_h(\mu(h) \cup c)$.

A3. Construction of $\succ_h^*$

DEFINITION 3: *Hospital h's priority ordering over the individual doctors, $\succ_h$, and the preferences of the couples $\{\succ_c : c \in D^2\}$ is used to construct a strict ordering, $\succ_h^*$, over the the coalitions representing the assignment of a doctor or a couple to at least one position at h–namely, coalitions of the form $(d, h), (c, hh')$, $(c, h'h)$, and (c, hh).*

Denote a generic instance of one of these coalitions by $(\cdot, h)$. For each coalition $(\cdot, h)$, let $d(\cdot, h)$ be the doctor assigned to h. If $(\cdot, h)$ represents the assignment of both members of a couple to h, let $d(\cdot, h)$ denote the least preferred member of the couple according to $\succ_h$. Then, $\succ_h^*$ is defined as follows. For two different coalitions $(a, h) \neq (b, h)$, if $d(a, h) \succ_h d(b, h)$, then $(a, h) \succ_h^* (b, h)$. If $d(a, h) = d(b, h)$, then (a, h) and (b, h) represent two different assignments of a couple c, in which case, $(a, h) \succ_h^* (b, h)$ if and only if $(a, h) \succ_c (b, h)$.

A4. Discussion of Stability

Under responsive choice functions, Definition 2 can have an undesirable implication. The following example suggested by a referee illustrates this.

Suppose two single doctors d, d', a couple $c = (f, m)$ and a hospital h with capacity 2. Recall, that for the couple c and hospital h to block a matching we require $\{f, m\} \subset ch_h(\mu(h) \cup \{f, m\})$, thus it is a stable matching for h to hire (d, d'), who are in 2^{nd} and 4^{th} positions, while the hospital may actually prefer the couple, whose members are ranked 1^{st} and 3^{rd}.

Because $\succ^*$ is defined based on the least preferred member of a couple, the stable matching we construct actually satisfies a stronger notion of stability. In particular, replace item 3 in Definition 2, with the following:

$$\begin{aligned} 3'.\ & \text{A pair } (c, h) \in D^2 \times H \text{ can block } \mu \text{ if } (h, h) \succ_c \mu(c) \\ & \text{and both } f_c \subseteq ch_h(\mu(h) \cup f_c) \text{ and } m_c \subseteq ch_h(\mu(h) \cup m_c). \end{aligned}$$

Under this definition, the matching in which $\mu(h) = \{d, d'\}$ is not stable because it is blocked by (c, h).

The ordering $\succ^*$ in Definition 3 is *not* a primitive of the model but a technical device introduced to invoke Scarf's lemma. We prove domination with respect to $\succ^*$ and show in Lemma 3.2 that this corresponds to stability with respect to Definition 2.

The same referee points out that domination with respect to $\succ^*$ can be restrictive. Specifically, change the hospital's priority ordering in the previous example to $f \succ_h d \succ_h d' \succ_h m$. The hospital's modified ranking is $d \succ_h^* d' \succ_h^* c$. The only dominating extreme point will assign $\{d, d'\}$ to h. This might be considered restrictive because it is possible that the hospital will prefer the couple c to the pair (d, d'). However, to evaluate such choices, one needs to extend the standard model because hospitals are not endowed with orderings over pairs of doctors. This is beyond the scope of this paper.

A5. Proof of Lemma 2

The proof is by contradiction. Let $\bar{x}$ be an integral dominating solution of (1-2-3), and assume that the corresponding assignment μ in the residency matching with couples is not stable. This means that at least one of the three items below is true.

1) A pair $d \in D^1$ and $h \in H$ blocks μ because $h \succ_d \mu(s)$ and $d \in ch_h(\mu(h) \cup d)$.
2) A triple $(c, h, h') \in D^2 \times H \times H$ with $h \neq h'$ blocks μ because $(h, h') \succ_c \mu(c)$, $f_c \in ch_h(\mu(h) \cup f_c)$ and $m_c \in ch_{h'}(\mu(h') \cup m_c)$.
3) A pair $(c, h) \in D^2 \times H$ blocks μ because $(h, h) \succ_c \mu(c)$ and $(f_c, m_c) \subseteq ch_h(\mu(h) \cup \{f_c, m_c\})$.

The first type of blocking coalition corresponds to the column associated with variable (d, h). Now, because $ch_h(.)$ is a responsive choice function over *individual* doctors, $d \in ch_h(\mu(h) \cup d)$ implies that d is among the best k_h candidates among $\mu(h) \cup d$. Therefore, $\bar{x}$ does not dominate column (d, h): this is a contradiction because $\bar{x}$ is a dominating solution.

The second type of blocking coalition corresponds to column (c, h, h'). Following the same argument, the blocking coalition implies that f_c is among the best k_h candidates among $\mu(h) \cup f_c$ (similar for m_c and h'.) Together with the tie-breaking rule of $\succ_h^*$, this implies that $\bar{x}$ does not dominate the column (c, h, h').

In the third type of blocking coalition, the pair (f_c, m_c) and a hospital h correspond to a column (c, h, h). Because $(f_c, m_c) \subseteq ch_h(\mu(h) \cup c)$, *both* f_c and m_c are among the k_h best candidates, even when we consider the order $\succ^*$ for the columns, because both members are still ranked highly among $\mu_h \cup \{f_c, m_c\}$. In the matching μ, the couple c is not assigned to h, thus, either h's capacity is not fully allocated, or a doctor worse than both f_c and m_c is assigned to h. Both cases imply that $\bar{x}$ does not dominate column (c, h, h). ∎

Maintaining Stability in Rounding

B1. Proof of Lemma 3

First of all, x^* is a feasible matching with respect to capacities k^*. Because $\bar{x}$ only contains assignments of mutually acceptable hospital-doctors, so does x^*. Thus, x^* is individually rational. Given that $\bar{x}$ dominates all columns of $\mathcal{Q}$, and x^*

is obtained from $\bar{x}$, we show that under the new capacity vector k^*, x^* dominates all columns of $\mathcal{Q}$.

Consider the column associated with the assignment of couple c_0 to hospital h_1 and h_2, (c_0, h_1, h_2). (A similar argument will apply to the other columns). $\bar{x}$ dominates (c_0, h_1, h_2) either at the constraint corresponding to c_0 or at $h_1 \in H$ or at $h_2 \in H$.

Suppose first $\bar{x}$ dominates (c_0, h_1, h_2) at c_0. Then $\sum_{h,h'} \bar{x}_{(c_0,h,h')} = 1$, and couple c_0 does not like the allocation h_1, h_2 strictly more than any of the assignments that they obtained under $\bar{x}$. Now because x^* is a $0-1$ vector rounded from $\bar{x}$ that satisfies Lemma 3:

(i.) ${x^*}_{(c_0,h,h')} > 0 \;\Rightarrow \bar{x}_{(c_0,h,h')} > 0$

(ii.) $\sum_{h,h'} \bar{x}_{(c_0,h,h')} = 1 \;\Rightarrow\; \sum_{h,h'} x^*_{(c_0,h,h')} = 1.$

These imply that c_0 (weakly) prefers the assignments that they gets in x^* more than (h_1, h_2) (we use 'weakly prefers' because it is possible that $x^*_{(c_0,h_1,h_2)} = 1$).

Next, suppose $\bar{x}$ dominates (c_0, h_1, h_2) at h_1 (a similar argument will apply to h_2). This implies that the capacity of hospital h_1 binds: $\mathcal{H}_{h_1}\bar{x} = k_{h_1}$. Furthermore, h_1 weakly prefers all columns in which the corresponding component of $\bar{x}$ is positive to (c_0, h_1, h_2). Now because of property (i) in Lemma 3, a component of x^* can be positive only when the corresponding component of $\bar{x}$ is positive. Thus, $\bar{x}$ dominates (c_0, h_1, h_2) when we change the capacity at h_1 to be $k^*_{h_1} := \mathcal{H}_{h_1}x^*$.

B2. When a Hospital's Capacity Constraint Does not Bind

Given a fractional dominating solution $\bar{x}$, let H^0 be the set of hospitals for which (1) does not bind. Denote the total slack in these non-binding constraints by K (not necessarily integral).

Introduce $\lceil K \rceil$ dummy single doctors $d_1, \ldots, d_{\lceil K \rceil}$. Choose a strict ordering over the hospitals in H^0, and assign it to each of the dummy doctors. The remaining hospitals will be ranked below $\emptyset$ by all the dummy doctors. Augment the priority ordering of hospitals in H^0 by appending $d_1 \succ \ldots \succ d_{\lceil K \rceil}$ to the bottom of these hospitals' orderings but above $\emptyset$. The priority ordering of hospitals not in H^0 is augmented by appending $d_1 \succ \ldots \succ d_{\lceil K \rceil}$ to the bottom of these hospitals' preference above $\emptyset$.

Extend $\bar{x}$ to include the dummy doctors so that all slots in H^0 are filled. We can do this by going through the list of dummy doctors from d_1 to $d_{\lceil K \rceil}$ and assigning each doctor to the best position available. Because we are working with a fractional assignment, a doctor can be split between different positions. Let $\bar{\bar{x}}$ be the resulting assignment. It is straightforward to see that $\bar{\bar{x}}$ is a dominating solution of the instance with dummy doctors, and this solution fully allocates all positions. Let x^{**} be an integral solution obtained by rounding $\bar{\bar{x}}$ according to the IR algorithm. Let k^{**} be the new capacity of the hospitals–that is, $k^{**} := \mathcal{H} \cdot x^{**}$.

According to Lemma 3, x^{**} is a stable solution with respect to k^{**}, and our algorithm bounds the difference between k^{**} and k.

We show that after eliminating the variables corresponding to dummy doctors from x^{**}, the resulting assignment, x^*, is stable with respect to k^{**}. This is true because under $\bar{x}$, the constraints (1) corresponding to hospitals in H^0 do not bind. Hence, $\bar{x}$ dominates all columns of the constraint matrix $\mathcal{Q}$ either at a couple/doctor constraint or at a hospital h constraint where $h \notin H^0$. As dummy doctors are never assigned to hospitals outside of H^0, it follows that for all $h \notin H^0$, $\mathcal{H}_h \cdot x^{**} = \mathcal{H}_h \cdot x^*$. Hence,

$$k_h^{**} = \mathcal{H}_h \cdot x^{**} = \mathcal{H}_h \cdot x^* = k^* \text{ for } h \notin H^0.$$

With these observations, and following the same argument as in Section B.B1, we obtain that x^* is stable with respect to k^{**}.

B3. Termination of the IR algorithm

To show that the IR algorithm terminates with an integral solution, we prove that if it has not yet terminated, we can always eliminate a constraint. It relies on the following lemma (Lemma 2.1.4, page 14, Lau, Ravi and Singh (2011)).

LEMMA 4 (Rank Lemma): *Let $P = \{x : Ax \geq b, x \geq 0\}$ and let x be an extreme point of P such that $x_j > 0$ for every i. Then, the maximal number of linearly independent binding constraints of the form $A_i x = b_i$ for some row i of A equals the number of variables.*

We reformulate Lemma 4 below, to apply in our setting.

LEMMA 5: *Let x be an extreme point of $Q = \{x : \mathcal{Q}x = q, 0 \leq x \leq 1\}$. Let J be the index set of non-integral components of x. Let $\mathcal{Q}|_J$ be the submatrix of $\mathcal{Q}$ consisting of the columns indexed by J. Then, the number of non-integral components of x, $|J|$, is equal to the maximum number of linearly independent rows of $\mathcal{Q}|_J$.*

To prove Lemma 5, let I be the index set of integral components of x, that is x_j is either 0 or 1 for all $j \in I$. We can rewrite $\mathcal{Q}x = \mathcal{Q}|_J \cdot x|_J + \mathcal{Q}|_I \cdot x|_I = q$. Let $q' := \mathcal{Q}|_J \cdot x|_J = q - \mathcal{Q}|_I \cdot x|_I$, and consider $Q|_J = \{y \in \mathbb{R}^{|J|} : \mathcal{Q}|_J \cdot y = q', y \geq 0\}$. The solution $x|_J$ is an extreme point of $Q|_J$ and all of its components are strictly positive. Applying Lemma 4 to $Q|_J$ and $x|_J$ we obtain Lemma 5.

To see how to use this lemma in our proof, let $\mathcal{D}^*, \mathcal{A}^*$ be the submatrices of $\mathcal{D}$ and $\mathcal{A}$, respectively, corresponding to the binding constraints of the linear program in Step 1. Thus, x is an extreme solution of $\left\{ \begin{bmatrix} \mathcal{D}^* \\ \mathcal{A}^* \end{bmatrix} x = \begin{bmatrix} 1 \\ b^* \end{bmatrix}; 0 \leq x \leq 1 \right\}$. Let J be the index of a non-integral component of x. Assume, for a contradiction, that we cannot eliminate any binding constraints. Credit every component of $x|_J$ with one token. Subsequently, we redistribute these tokens to the constraints (rows) of $\begin{bmatrix} \mathcal{D}^*|_J \\ \mathcal{A}^*|_J \end{bmatrix}$ in such a way that each constraint will get at least 1 token. We show this to be possible because each column of the matrix has a relatively small number of non-zero entries. This redistribution shows that the number of binding constraints is at most the number of non-integral components. Furthermore, we show that equality arises only when the binding constraints are linearly dependent. This implies that the maximum number of linearly independent constraints is less than the number of non-integral components, which contradicts Lemma 5.

Token Distribution

To complete the proof we show that if the algorithm has not yet terminated, we can always find a constraint to eliminate. Suppose, for a contradiction, we are at an iteration where no constraint can be eliminated and each component of $x|_J$ is fractional. Endow each fractional component of $x|_J$ with 1 token and redistribute that token among the constraints in (4) and (5) as follows:

- The 1 token associated with the variable $x_{(c,h,h')}$ is apportioned as follows: a $\frac{1}{4}$ tokens to each of the constraints $\mathcal{H}_h \cdot x = k_h$ and $\mathcal{H}_{h'} \cdot x = k_{h'}$ (if $h = h'$, then $\mathcal{H}_h \cdot x = k_h$ gets $\frac{1}{2}$ tokens) and the remaining $\frac{1}{2}$ token assigned to the couple c constraint–that is, $\sum_{h,h'} x_{(c,h,h')} \leq 1$.
- The one token associated with the variable $x_{(d,h)}$ is apportioned as follows: a $\frac{1}{4}$ tokens to the constraints $\mathcal{H}_h \cdot x = k_h$; the remaining $\frac{3}{4}$ tokens are allotted to the doctor d constraint–that is, $\sum_h x_{(d,h)} \leq 1$.

We now argue that each binding constraint in (4) and (5) receives at least one token. Consider a binding constraint $\mathcal{H}_h \cdot x = k_h$ associated with hospital h. By the assumption that no constraint can be eliminated, we know that $\mathcal{H}_h \cdot (\lceil x \rceil - \lfloor x \rfloor) \geq 4$. Keep in mind that $\lceil x_i \rceil - \lfloor x_i \rfloor = 1$ if x_i is non-integral, and 0 otherwise. According to the token distribution scheme, a non-integral component of x gives the hospital h constraint $\frac{1}{4}$ or $\frac{1}{2}$ tokens if the corresponding assignment requires 1 or 2 slots from h, respectively. Thus, the number of tokens constraint $\mathcal{H}_h \cdot x = k_h$ gets is at least

$$\frac{1}{4}\mathcal{H}_h \cdot (\lceil x \rceil - \lfloor x \rfloor) \geq 1.$$

Next, consider a binding constraint corresponding to couple c. As this constraint binds–that is, $\sum_{h,h'} x_{(c,h,h')} = 1$–and it contains at least 1 non-integral variable, it must contain at least 2. Each of the fractional variables contributes $\frac{1}{2}$ a token, thus this constraint also obtains at least 1 token.

Similarly, for the constraint corresponding to a single doctor d. If this constraint binds and contains at least one non-integral variable, it must contains at least 2. Therefore, it also gets at least $2 \times \frac{3}{4} \geq 1$ token.

The total number of tokens distributed cannot exceed the number of fractional components of $x|_J$ which is $|J|$. By Lemma 5, total number of tokens received by binding constraints in (4) and (5) is at least the number of such binding constraints, $|J|-1$. This is because the aggregate capacity constraint may bind. We have two cases.

Case 1: The aggregate capacity constraint has not yet been eliminated.
We know that the total number of tokens allocated to binding constraints in (4) and (5) is at least $|J|-1$. Because the aggregate constraint has not yet been eliminated, there are at least three *non binding* doctor/ couple constraints that contain fractional variables. According to the token distribution scheme, we gave to these constraints at least $3 \times \frac{1}{2}$ tokens. Hence, the total number of tokens assigned to constraints in (4) and (5), binding or not, is at least $|J| + \frac{1}{2}$. This exceeds the the total number of tokens to be distributed, a contradiction.

Case 2: The aggregate constraint was eliminated at some earlier iteration.
By the extreme point property of $x|_J$, the $|J|$ binding constraints belong to (4) and (5). Each one of the binding constraint receives at least one token. Hence, none can receive strictly more than one token. This means no constraint in (2) can bind. Similarly, no non-binding constraint can receive any tokens. Hence, in $x|_J$, all variables associated with single doctors take the value zero. Furthermore, if $x(c,h,h') > 0$, the capacity constraints associated with h and h' must bind. If we apply these observations to the system (1, 2, 3), the relevant binding constraints have the form:

$$\text{(B1)} \qquad \sum_{c\in D^2}\sum_{h'\neq h} x_{(c,h,h')} + \sum_{c\in D^2}\sum_{h'\neq h} x_{(c,h',h)} + \sum_{c\in D^2\cup\{\emptyset\}} 2x_{(c,h,h)} = k_h$$

$$\text{(B2)} \qquad \sum_{h,h'\in H\cup\{\emptyset\}} x_{(c,h,h')} = 1$$

If we add up the binding constraints of the form (B2) we get the sum of the binding constraints of the form (B1). This violates the assumption that the binding constraints must be linearly independent. Hence, if we add up the binding constraints in (3) we get the sum of the binding constraints in (1). This violates the assumption of linear independence.

B4. Tightness

We outline why the token argument we used cannot be modified to give an improved bound. We will allow the quantity of tokens allocated to hospital h to depend on h.[14] For each hospital h let $r_h = \mathcal{H}_h \cdot (\lceil x \rceil - \lfloor x \rfloor)$. As before, suppose we are at an iteration where no constraint can be eliminated and each component of $x|_J$ is fractional. Endow each fractional component of $x|_J$ with 1 token and

[14] The same conclusion will be reached even if we allow the quantity of tokens to depend on both the hospital and the identity of the doctors.

redistribute the tokens among the constraints in (1-2-3) as follows:

- The 1 token associated with the variable $x_{(c,h,h')}$ is apportioned as follows: $\frac{1}{r_h}$ tokens to each of the constraints $\mathcal{H}_h \cdot x = k_h$ and $\mathcal{H}_{h'} \cdot x = k_{h'}$ (if $h = h'$, then $\mathcal{H}_h \cdot x = k_h$ gets $\frac{2}{r_h}$ tokens) and the remaining $1 - \frac{2}{r_h}$ token assigned to the couple c constraint–that is, $\sum_{h,h'} x_{(c,h,h')} \leq 1$.
- The 1 token associated with the variable $x_{(d,h)}$ is apportioned as follows: $\frac{1}{r_h}$ tokens to the constraints $\mathcal{H}_h \cdot x = k_h$; the remaining $1 - \frac{1}{r_h}$ tokens are allotted to the doctor d constraint–that is, $\sum_h x_{(d,h)} \leq 1$.

It is straightforward to see that the number of tokens allocated to each hospital h is at least

$$\frac{\mathcal{H}_h \cdot (\lceil x \rceil - \lfloor x \rfloor)}{r_h} = 1.$$

Now, consider the number of tokens allocated to a single doctor d constraint. There must be at least two hospitals h and h' such that $x(d,h), x(d,h') > 0$. Hence, the number of tokens allocated to this constraint is at least $1 - \frac{1}{r_h} + 1 - \frac{1}{r_{h'}}$. We need this sum to be at least 1. Hence, $r_h, r_{h'} \geq 2$. A similar argument for a couples, c, constraint requires that

$$1 - \frac{2}{r_h} + 1 - \frac{2}{r_{h'}} \geq 1 \ \Rightarrow r_h, r_{h'} \geq 4.$$

Hence, for our token argument to work we need $r_h \geq 4$ for all hospitals h which is precisely what we have assumed.

Additional Results

C1. *Proof of Theorem 2*

Let H^R be the set of rural hospitals, to which we assume no couple applies. Let H^U be the remaining (urban) hospitals. The main change in the IR algorithm is that we never drop any constraint corresponding to $h \in H^R$. Thus, at each iteration

$$\mathcal{H}_h x = k_h \text{ for all } h \in H^R.$$

The modified version of the IR algorithm, called IR1, is described in Figure C1.

To show that the IR1 algorithm returns a near-feasible stable matching that does not violate the capacity of $h \in H^R$, we follow the proof of Theorem 1. It is enough to show that if IR1 algorithm has not terminated, we can always find an active constraint to delete.

First, because the IR1 algorithm always maintains a solution satisfying the capacity constraints of rural hospitals, the aggregate constraint can be rewritten

Step 0. Start from $x := \bar{x}$ a dominating solution satisfying (4) and (5).
Initialize the active constraints to be the constraints:
$\mathcal{H}_h \cdot x = k_h$ for $h \in H^U$ and the aggregate constraint $a \cdot x \leq \sum k_h$

Step 1. If x is integral, stop; otherwise, among the active constraints that bind at the solution x, we eliminate one of them. The rule for selecting the constraint to eliminate is described:

- Choose any binding urban hospital constraint, $\mathcal{H}_h \cdot x = k_h$, such that $\mathcal{H}_h \cdot (\lceil x \rceil - \lfloor x \rfloor) \leq 3$ and eliminate it.
- If no urban hospital constraint can be eliminated, eliminate the aggregate capacity constraint.

If no constraint can be found to eliminate, stop, x must be integral. If a constraint is eliminated, denote by $\mathcal{A}x \leq b$ the system of remaining (active) constraints in (5).

Step 2. Find an extreme point z to maximize the number of jobs allocated:

$$
\max a \cdot z : \quad
\begin{aligned}
& z_i = x_i \text{ if } x_i \text{ is either 0 or 1} \textit{(fix the integral components)} \\
& \mathcal{D}_0 \cdot z = 1 \\
& \mathcal{D}_1 \cdot z \leq 1 \ \textit{(doctor/couple constraints as in } (4)\textit{)} \\
& z \geq 0 \\
& \mathcal{H}_h x = k_h \text{ for all } h \in H^R \textit{(rural hospital constraints)} \\
& \mathcal{A} z \leq b \ \ \textit{(active hospital constraints.)}
\end{aligned}
$$

Step 3. Update x to be the extreme point solution z^* found in Step 2.
Update $\mathcal{D}_0$ to include the new constraints from (4) that become binding at z^* from Step 2.
Update $\mathcal{D}_1$ to remove the new constraints from (4) that become binding at z^* from Step 2.
Return to Step 1.

Figure C1. IR1 algorithm

in terms of urban hospitals only. Namely,

$$\sum_{d,h:h\in H^U} x_{(d,h)} + \sum_{c,h,h':h,h'\in H^U} 2x_{(c,h,h')} \leq \sum_{h\in H^U} k_h.$$

Absent from this constraint is any variable $x_{(c,h,h')}$ where among the pair (h, h'), one is urban and the other is rural because of our assumption that only single doctors apply to rural hospitals.

Second, we modify the token distribution scheme by changing how the token associated with $x_{(d,h)}$ for $h \in H^R$ is allocated. Namely, assign $\frac{1}{2}$ a token to the constraint $\mathcal{H}_h \cdot x = k_h$; the remaining 1/2 token is given to the doctor d constraint–that is, $\sum_h x_{(d,h)} \leq 1$. For the other variables, the token distribution remains the same as in Section B.B3.

Each urban hospital constraint receives at least 1 token. To see why, observe that if a hospital constraint contains a non-integral variable, it must contain at least two of them. Each non-integral variable contributes 1/2 a token to the relevant constraint. Thus, the relevant constraint obtains at least 1 token.

Each couple constraint has at least two non-integral variables or none. When none, we can ignore this constraint because it does not affect any non-integral variables. As before, the number of tokens allocated to a couple constraint is at least 1.

Each fractional variable in in a single doctor constraint contributes either 1/2 or 3/4 of a token depending on whether the corresponding hospital is rural or urban. Thus, such a constraint also receives at least 1 token and *strictly* more than that if one of the variables is associated with an urban hospital.

Hence, as in case 1 in Section B.B3, we can always eliminate one active constraint if the IR1 algorithm has not terminated. When there are no active constraints left (as in case 2 of Section B.B3), the remaining constraints and variables are associated with the single doctors and rural hospitals only. This corresponds to the standard linear program of a many-to-one matching *without* couples. An extreme point of this linear program is integral.

C2. *Using Different Objective Functions to Prioritize Hospitals*

The IR algorithm described in Figure 1 uses an objective function, $a \cdot x$, to maximize the number of jobs allocated. Termination of the IR algorithm does not depend on this specific choice of objective function. The IR algorithm works for *any* linear objective function, $c \cdot x$. This can be used to reflect the fact that assigning extra slots to one hospital may be cheaper than allocating them to another.

In particular, replacing $\max a \cdot x$ with any linear objective function $c \cdot x$, the IR algorithm in Figure 1, starting from the fractional stable matching $\bar{x}$, will terminate in a 2-feasible stable matching in which the aggregate capacity does not increase by more than 4. Furthermore, $c \cdot x^* \geq c \cdot \bar{x}$.

Because the choice of the linear objective function, c is arbitrary, we can round $\bar{x}$ in any "direction". This implies the following result. (See Figure C2 for an illustration.)

CLAIM 1: *The fractional stable matching $\bar{x}$ can be expressed as a lottery over 2-feasible stable matchings that do not violate the aggregate constraint by more than* 4*.*

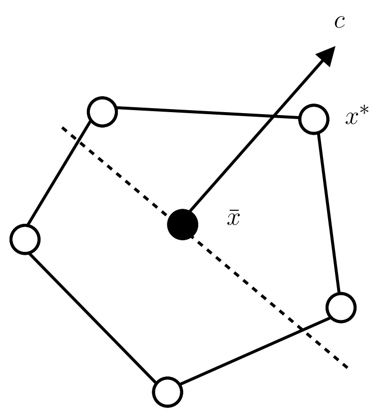

FIGURE C2. FRACTIONAL STABLE MATCHING CAN BE EXPRESSED AS A LOTTERY OVER NEAR-FEASIBLE STABLE MATCHINGS

Claim 1 is true because otherwise, $\bar{x}$ lies outside the convex hull of the near-feasible stable matchings, and therefore we can separate $\bar{x}$ from these near-feasible stable matchings with a linear function.

Claim 1 provides a randomized algorithm to round $\bar{x}$ so that it is ex-ante feasible (but ex-post is 2-feasible).