The “Rural Hospital Theorem” and Market Design under Distributional Constraints

Fuhito Kojima\textsuperscript{1}

April 1, 2015

\textsuperscript{1}Stanford University.
Geographical distribution of medical doctors is a contentious (and often politicized) issue in health care.

“Hospitals in rural areas do not attract enough medical residents to meet their demands: 35 million Americans living in underserved areas and need 16,000 doctors.” (Washington Post)
Rural Doctor Shortages in Other Countries

Similar concerns about doctor shortages in rural areas appear around the globe. An Indian example:

“[E]xperts say at least 600,000 more doctors and 1 million nurses are needed to achieve the 1:1000 doctor-patient ratio recommended by the World Health Organization. It is a tall order for a country that graduates 30,000 new doctors and 45,000 nurses annually. It all but makes equitable access to healthcare impossible.” (Canadian Medical Association Journal)

Similar reports are easy to find (just google “doctor shortage” and “rural area”)

Fuhito Kojima

The “Rural Hospital Theorem” and Market Design under Distributional Constraints
A similar problem appears in terms of medical specialties: “[M]any medical students spoke of something called the R.O.A.D. to Happiness. “R.O.A.D.” stood for residencies in radiology, ophthalmology, anesthesiology, and dermatology. These specialties usually paid better than a career in family or internal medicine. ... A R.O.A.D. specialty also meant a better ability to have what doctors called a controllable lifestyle.” (“Match Day” by Brian Eule),

Other mathematically similar problems abound...

- Chinese Graduate School Admission (academic and professional master’s programs)
- College Admission in Ukraine (state-financed and privately-financed seats)
If the doctor imbalance is a problem (of course, we should also ask if it is really a problem), is there any policy to solve it?

Many different solutions have been proposed and tried.

1. Subsidies for doctors/hospitals in rural areas.
2. Fellowship for medical students if they promise to work in rural areas upon graduation.
3. “regional caps” (discussed below).
Matching Mechanism for Interns

The idea was discussed in earlier days of the American medical intern market (Sundarshan and Zisook, 1981):

“The United States suffers a terrible problem of maldistribution of physicians, with urban areas being relatively over-served and inner-city and rural areas being relatively under-served. At present, although approximately 100 hospitals fill every residency position, there are over 100 hospitals that do not receive a single application. This maldistribution would only be worsened by the only rational alternative to the present matching program, the mirror-image program, which favors students.”

The above quote says that the doctor-proposing DA would worsen the situation than the hospital-proposing DA (remember, before late 1990s, hospital-proposing DA was used).

More generally, any role of a mechanism for influencing geographical distributions of doctors?
The “Rural Hospital Theorem”

Rural Hospital Theorem (RS Theorem 2.22)

The set of doctors and hospitals that are unmatched is the same for all stable matchings.

- So the theorem says that it is impossible to increase doctors in rural areas as long as a stable matching mechanism is implemented!
- Another remark: If some students are matched in some stable matching and not in others, the latter may be unfair to him/her. The theorem says that there is no need to worry (this was a policy issue in 1990s; see Roth and Peranson 1999).
Proof of Rural Hospital Theorem

We use a well-known property of stable matchings:

**Result (Gale and Shapley 1962; RS Theorem 2.12)**

1. There exists (which is found by doctor-proposing DA) a **doctor-optimal stable matching**, that is, a stable matching that every doctor weakly prefers to any stable matching.
2. The doctor-optimal stable matching is hospital-pessimal, that is, every hospital weakly disprefers it to any stable matching.
3. A symmetric result holds for college-proposing DA.

Let $\mu^D$ be the doctor-optimal stable matching and $\mu$ be an arbitrary stable matching.

- Since $\mu^D$ is doctor-optimal, all the students that are matched in $\mu$ are matched in $\mu^D$.
- Since $\mu^D$ is hospital-pessimal, all the hospitals that are matched in $\mu^D$ are matched in $\mu$.
- But the number of matched doctors and hospitals are the same in any matching. So the same set of doctors and hospitals are matched in $\mu^D$ and $\mu$. 
As a matter of fact, an even stronger conclusion holds. Assume that hospitals have more than one position (“many-to-one” matching).

**Rural Hospital Theorem (RS Theorem 5.13)**

Any hospital that does not fill all of its positions at some stable matching is assigned precisely the same set of doctors at every stable matching.

- So, not only the number of students but also the set of doctors is unchanged at hospitals that do not fill all of their positions.
- The conclusions of these results seems to be pretty robust: More general results have been obtained by Martinez et. al (2000 JET), Hatfield and Milgrom (2005 AER), Hatfield and Kojima (2009 JET), among others.
Geographical Distribution of Medical Residents in Japan

So, what could be done?


- Prior to the reform, clinical departments in university hospitals allocated doctors.

Critics say that many rural hospitals fill fewer positions in the new matching mechanism.

Japanese government introduced a “regional cap” as a constraint, and modified the DA (JRMP mechanism).
Model of Regions

A hospital may have more than one position: let $q_h$ be hospital $h$'s capacity.

Each hospital belongs to exactly one region $r \in R$.

For each region $r$, there is a regional cap $q_r$ (a positive integer).

A matching is feasible if the number of doctors assigned in each region $r$ is at most $q_r$.

More generally, the model is “matching with distributional constraints”, with concrete applications like:

- Chinese Graduate School Admission (academic and professional master’s programs)
- College Admission in Ukraine (state-financed and privately-financed seats)
- Medical Matching in the U.K. (regional cap)
- Teacher Matching in Scotland (regional cap)
Example: The JRMP Mechanism

In Japan, government exogenously imposes a target capacity $\bar{q}_h \leq q_h$ for each hospital $h$ such that $\sum_{h \in H_r} \bar{q}_h = q_r$ for each region $r \in R$.

The JRMP mechanism (Japan Residency Matching Program mechanism): the deferred acceptance mechanism, except that it uses the target capacity.

Idea: In order to satisfy regional caps, simply force hospitals to be matched to a smaller number of doctors than their real capacities, but otherwise use the standard deferred acceptance algorithm.

But does the JRMP mechanism inherit good properties of DA?
There are two hospitals $h_1, h_2$ in one region with regional cap 10.

Each hospital has a capacity of 10 and a target capacity of 5.

There are 10 doctors, $d_1, \ldots, d_{10}$ such that

$$d_1 \succ_h d_2, \succ_h \ldots \succ_h d_{10} \succ_h \emptyset,$$

for both hospitals,

$d_1, d_2, d_3$ find only $h_1$ acceptable,

$d_4, \ldots, d_{10}$ find only $h_2$ acceptable.

The JRMP mechanism produces

$$\mu_{h_1} = \{d_1, d_2, d_3\}$$

$$\mu_{h_2} = \{d_4, d_5, d_6, d_7, d_8\}.$$

This matching is inefficient (even in the constrained sense).

Remark: Assumption that some doctors find some hospitals unacceptable is for simplicity. There exists an example where all doctors find all hospitals acceptable.
Outcome of JRMP May Be Unstable

The same example as before: There are two hospitals $h_1, h_2$ in one region with regional cap 10.

Each hospital has a capacity of 10 and a target capacity of 5.

There are 10 doctors, $d_1, \ldots, d_{10}$ such that

$$d_1 \succ_h d_2, \succ_h \ldots \succ_h d_{10} \succ_h \emptyset,$$

for both hospitals,

$d_1, d_2, d_3$ find only $h_1$ acceptable,

$d_4, \ldots, d_{10}$ find only $h_2$ acceptable.

The JRMP mechanism produces

$$\mu_{h_1} = \{d_1, d_2, d_3\}$$

$$\mu_{h_2} = \{d_4, d_5, d_6, d_7, d_8\}.$$

This matching is not “stable.”

Remark: Similar inefficiency and instability occur under other existing mechanisms.
Clearly, the JRMP matching can be unstable in the standard sense. We introduce a new stability concept.

→ Idea: The only blocking pair is caused because the regional cap is binding.

We will consider “strong stability” and “weak stability”.

**Definition**

A matching $\mu$ is **strongly stable** if it is feasible, individually rational and, if $(d, h)$ is a blocking pair, then

1. number of doctors matched in $h$’s region = its regional cap,
2. $d' \succ_h d$ for all $d' \in \mu_h$, and
3. $d$ is not matched in $h$’s region.

In other words, satisfying the blocking pair will lead to a violation of a regional cap.
There is one region with regional cap 1, with two hospitals $h_1$ and $h_2$ with capacity 1 each and two doctors, $d_1$ and $d_2$, with preferences

$$
\succ h_1 : d_1, d_2, \quad \succ h_2 : d_2, d_1, \\
\succ d_1 : h_2, h_1, \quad \succ d_2 : h_1, h_2.
$$

1. No matching in which two doctors are matched is feasible because it violates the regional cap.
2. If no doctor is matched, then there is a blocking pair ($d_1$ and $h_1$ for example).
3. $\mu_{h_1} = \{d_2\} \implies (d_1, h_1)$ is a blocking pair ($h_1$ can reject $d_2$ to be paired with $d_1$).
4. $\mu_{h_1} = \{d_1\} \implies (d_1, h_2)$ is a blocking pair.
5. $\mu_{h_2} = \{d_2\}$ and $\mu_{h_2} = \{d_1\}$ is not strongly stable (symmetric argument).
Weak Stability

Definition

A matching $\mu$ is **weakly stable** if it is feasible, individually rational and, if $(d, h)$ is a blocking pair, then

1. number of doctors matched in $h$'s region = its regional cap, and
2. $d' \succ_h d$ for all $d' \in \mu_h$

Idea: We tolerate some more blocking pair, while still requiring that the only blocks involve vacant positions and regional cap is already full.

Remark: The paper considers **stability**, which is stronger than weak stability but still guarantees existence.
We define the **flexible deferred acceptance mechanism** below, given target capacity profile \((\bar{q}_h)_{h \in H}\). Start with a matching in which no one is matched.

**Application Step:**
Choose a currently unmatched doctor, and let her apply to her most preferred hospital that has not rejected her so far (if any).

**Acceptance/Rejection Step:**
Consider the region of the hospital receiving the new application. Each hospital in the region chooses from the tentatively matched doctors and the new applicant (if any):

1. First, each hospital chooses its most preferred acceptable applicants up to its target capacity.
2. Then, one by one, each hospital in the region takes turns (following a fixed order) to choose the most preferred remaining applicant who are applying to it until (i) the regional quota is filled or (ii) the capacity of the hospital is filled or (iii) no doctor remains to be matched.
Example of flexible DA

The same example as before: There are two hospitals $h_1, h_2$ in one region with regional cap 10.

Each hospital has a capacity of 10 and a target capacity of 5.

There are 10 doctors, $d_1, \ldots, d_{10}$ such that

$$d_1 \succ_h d_2, \succ_h \ldots \succ_h d_{10} \succ_h \emptyset,$$

for both hospitals,

$d_1, d_2, d_3$ find only $h_1$ acceptable,

$d_4, \ldots, d_{10}$ find only $h_2$ acceptable.

The flexible DA mechanism produces

$$\mu_{h_1} = \{d_1, d_2, d_3\}$$

$$\mu_{h_2} = \{d_4, d_5, d_6, d_7, d_8, d_9, d_{10}\},$$

which is (weakly) stable and efficient.
The flexible deferred acceptance mechanism produces a weakly stable matching for any input.

Intuition:

Unlike the JRMP mechanism, the target capacity of each hospital is not rigid.

As long as the regional cap is not violated, hospitals can tentatively accept doctors beyond the target capacities.

Like DA, an acceptable doctor rejected from a more preferred hospital was rejected either because there are enough better doctors in that hospital, or regional quota was filled by other doctors.

→ The doctor cannot form a legitimate blocking pair!
Theorem

Any weakly stable matching is efficient.

Note: This result is well-known when there is no regional cap, and is a straightforward implication of the fact that stability is equivalent to core.

But with regional caps, there is no obvious way to define the core. Fortunately the statement still goes through.

Corollary

The flexible deferred acceptance mechanism produces an efficient matching for any input.
Theorem

The flexible DA mechanism is (group) strategy-proof for doctors: Truth-telling is a dominant strategy for every doctor.

A (very rough) intuition: a doctor doesn’t need to give up trying for her first choice because, even if she is rejected, she will be able to apply to her second choice etc. The deferred acceptance guarantees that she will be treated equally if she applies to a position later than others.

Truth-telling is not necessarily a dominant strategy for hospitals (Roth 1982: There is no strategy-proof and stable mechanism.)
Other Mechanisms

Chinese Graduate School Admission (academic and professional master’s programs)

1. Set target capacity for each academic program.
2. Students apply to one program.

Medical Matching in the U.K., Teacher Matching in Scotland (two-round process)

1. First round: matching to a region.
2. Second round: matching to a program.
Additional Topics

1. Welfare Comparison (Theory and Simulation) → Detail.
2. A mechanism that does not work → Detail.
3. Stability → Detail.
4. Failure of the Rural Hospital Theorem → Detail.
Related Literature

- **Regional Maximum Quotas** Abraham et al. (2007), Biro et al. (2010), Budish et al. (2013), Kamada and Kojima (2013, 2015a,b,c), Goto et al. (2014a), Milgrom (2009)
- **Minimum Quotas** Fragiadakis et al. (2012), Goto et al. (2014), Fragiadakis and Troyan (2014)
Conclusion

In markets like the one for medical residents, some balances of the assignments (across regions, professional/academic masters, etc.) are often important.

The rural hospital theorem says that, as long as stable mechanisms are used, the market design cannot help.

Taking the Japanese residency match as a case study, we analyzed other possibilities.

- Conceptual issues of “right” definition of stability
- Mechanisms, like FDA, to get stability and strategy-proofness for one side.

More generally, design of matching mechanisms with distributional constraints are a largely untouched area, but potentially with lots of applications.
Additional Slides
We compare the (unconstrained) DA, FDA, and JRMP.

**Theorem**

For any preference profile,

1. Each doctor weakly prefers a matching produced by DA to FDA to JRMP.
2. If a doctor is unmatched in DA, she is unmatched in FDA. If a doctor is unmatched in FDA, she is unmatched in JRMP mechanism.
Simulation Based (Partly) on Japanese Data: Doctors

Figure: The numbers of matched doctors under different mechanisms
Figure: Cumulative number of doctors matched to their \( k \)-th or better choices.
Table: The number of hospitals that are matched to more doctors in DA, FDA, and JRMP.

<table>
<thead>
<tr>
<th>From\To</th>
<th>DA</th>
<th>FDA</th>
<th>JRMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>DA</td>
<td>0</td>
<td>138 (10.2%)</td>
<td>222 (16.4%)</td>
</tr>
<tr>
<td>FDA</td>
<td>104 (7.7%)</td>
<td>0</td>
<td>158 (11.6%)</td>
</tr>
<tr>
<td>JRMP</td>
<td>366 (27.0%)</td>
<td>376 (27.7%)</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure: Distributional consequence of the change from JRMP to FDA.
A Mechanism That Does Not Work

The **iterated deferred acceptance** mechanism:

1. Run the JRMP mechanism.
2. If the resulting matching fills all the target capacities, then terminate.
3. Otherwise, modify target capacities and re-run JRMP.
4. Repeat (following some termination rule).

A very simple idea. We often hear suggestions along this line.
It turns out that this mechanism is not strategy-proof for doctors.

Example

1. Doctors $d_1$ and $d_2$, and hospitals $h_1$ and $h_2$ in a single region with regional cap 2.
2. Each doctor prefers $h_1$ to $h_2$ to being unmatched.
3. Each hospital has capacity of 2 and a target capacity of 1, and prefers $d_1$ to $d_2$ to being unmatched.

Under iterated DA,

$d_2$ tells the truth $\rightarrow \mu = \begin{pmatrix} h_1 & h_2 \\ d_1 & d_2 \end{pmatrix}$.

$d_2$ declares $h_2$ unacceptable $\rightarrow \mu' = \begin{pmatrix} h_1 & h_2 \\ d_1, d_2 & \emptyset \end{pmatrix}$.
A General Framework for Stability

Let **regional preferences** $\succeq_r$ be a weak ordering over nonnegative-valued integer vectors $W_r := \{ w = (w_h)_{h \in H_r} \mid w_h \in \mathbb{Z}_+ \}$.

**Definition**

A matching $\mu$ is **stable** if it is feasible, individually rational, and if $(d, h)$ is a blocking pair then (i) $|\mu_{r(h)}| = q_{r(h)}$, (ii) $d' \succ_h d$ for all doctors $d' \in \mu_h$, and (iii') either $\mu_d \notin H_{r(h)}$ or $w \succeq_{r(h)} w'$, where $w_h' = |\mu_{h'}|$ for all $h' \in H_{r(h)}$ and $w_h' = w_h + 1$, $w_{\mu_d} = w_{\mu_d} - 1$ and $w_{h'} = w_{h'}$ for all other $h' \in H_{r(h)}$.

When regional preferences satisfy (an appropriately defined) substitutability, we can generalize FDA that finds a stable matching and strategy-proof for doctors.
Further Generalization

Studying more general constraint structures may be interesting.
- Regional caps; one for each prefecture, one for each district in the prefecture.
- Government would like to impose caps based on prefectures and specialities.

When there is a hierarchy of regional caps, we show that a stable matching can be found by a generalization of our FDA.

By contrast, without a hierarchy, a stable matching does not necessarily exist.
Further Generalization

Studying more general constraint structures may be interesting.
- Regional caps; one for each prefecture, one for each district in the prefecture.
- Government would like to impose caps based on prefectures and specialities.

When there is a hierarchy of regional caps, we show that a stable matching can be found by a generalization of our FDA.

By contrast, without a hierarchy, a stable matching does not necessarily exist.
Let \textbf{regional preferences} \( \succeq_r \) be a weak ordering over nonnegative-valued integer vectors 
\( W_r := \{ w = (w_h)_{h \in H_r} | w_h \in \mathbb{Z}_+ \} \).

\begin{definition}
A matching \( \mu \) is \textbf{stable} if it is feasible, individually rational, and if \((d, h)\) is a blocking pair then (i) \( |\mu_r(h)| = q_{r(h)} \), (ii) \( d' \succ_h d \) for all doctors \( d' \in \mu_h \), and (iii') either \( \mu_d \notin H_{r(h)} \) or \( w \succeq_{r(h)} w' \), where \( w_{h'} = |\mu_{h'}| \) for all \( h' \in H_{r(h)} \) and \( w'_h = w_h + 1 \), \( w'_{\mu_d} = w_{\mu_d} - 1 \) and \( w'_{h'} = w_{h'} \) for all other \( h' \in H_{r(h)} \).
\end{definition}

When regional preferences satisfy (a n appropriately defined) substitutability, we can generalize FDA that finds a stable matching and strategy-proof for doctors.

→ Back.
The conclusion of the rural hospital theorem fails: The set of unmatched doctors and hospitals can differ across stable matchings.

There are two regions $r$ and $r'$ with regional cap of one each. Hospitals $h_1$ and $h_2$ are in $r$ and $h_3$ is in $r'$ with capacity one each.

Preferences are

$\succ_{h_1} : d_1, d_2, \quad \succ_{h_2} : d_2, d_1, \quad \succ_{h_3} : d_2,$

$\succ_{d_1} : h_1, h_2, \quad \succ_{d_2} : h_2, h_3.$

One stable matching matches $d_1$ to $h_1$ and $d_2$ to $h_3$. Another stable matching matches $d_2$ to $h_2$ only.

Given this, design of the mechanism may influence geographical distributions of doctors.