Evaluating the Efficiency of Regulation in Matching Markets with Distributional Disparities

Kei Ikegami
(Yokohama City U)

Atsushi Iwasaki
(U Electro-Communications)

Akira Matsushita
(U Tokyo)

Kyohei Okumura
(UW Madison)

introduction

- outcomes in matching markets may differ from socially desirable ones
- policymakers often intervene in these markets
    - e.g., affirmative action in school choice, gender quotas in elections
- a common intervention is cap-based regulation
    - limits the number of matches in certain categories

Japan Residency Matching Program (JRMP)

Japan Residency Matching Program (JRMP)

- introduced in 2004
  - students and hospitals are matched for two-year residency programs

Japan Residency Matching Program (JRMP)

- introduced in 2004
  - students and hospitals are matched for two-year residency programs

- geographic disparities in the distrib. of residents have increased
  - **before**: majority of students trained at the **univ. hospital** affiliated with their schools
  - **after**: students can freely express their preferences, leading to a concentration of applicants in non-university and urban programs

Japan Residency Matching Program (JRMP)

- introduced in 2004
  - students and hospitals are matched for two-year residency programs

- geographic disparities in the distrib. of residents have increased
  - **before**: majority of students trained at the **univ. hospital** affiliated with their schools
  - **after**: students can freely express their preferences, leading to a concentration of applicants in non-university and urban programs
  - in 2009, **48.6**% of residents were matched with hospitals in **6** of Japan's **47** prefectures

Japan Residency Matching Program (JRMP)

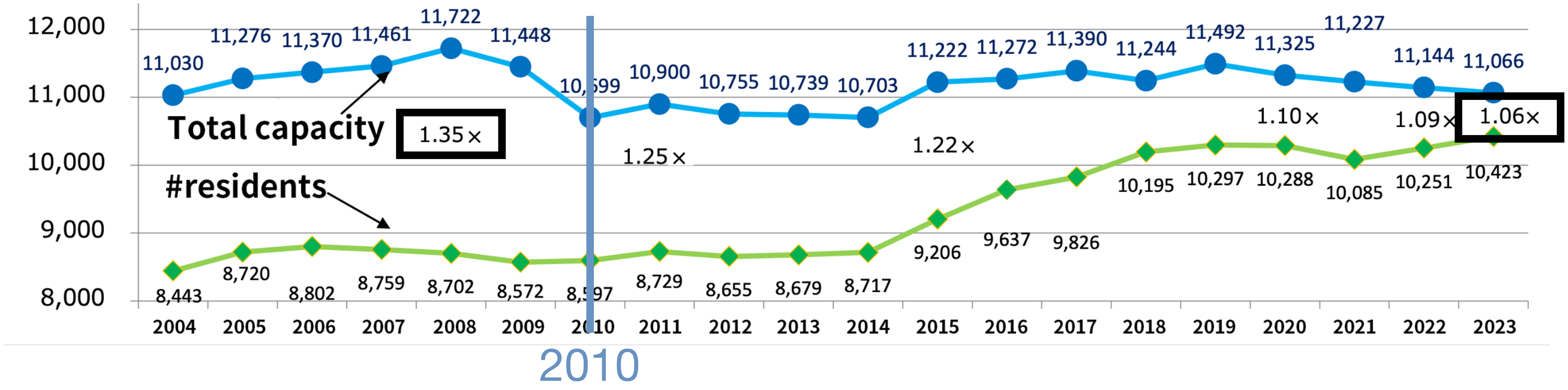
- introduced in 2004
  - students and hospitals are matched for two-year residency programs

- geographic disparities in the distrib. of residents have increased
  - **before**: majority of students trained at the **univ. hospital** affiliated with their schools
  - **after**: students can freely express their preferences, leading to a concentration of applicants in non-university and urban programs
  - in 2009, **48.6**% of residents were matched with hospitals in **6** of Japan's **47** prefectures

- policymakers sought to secure adequate resident coverage in all areas

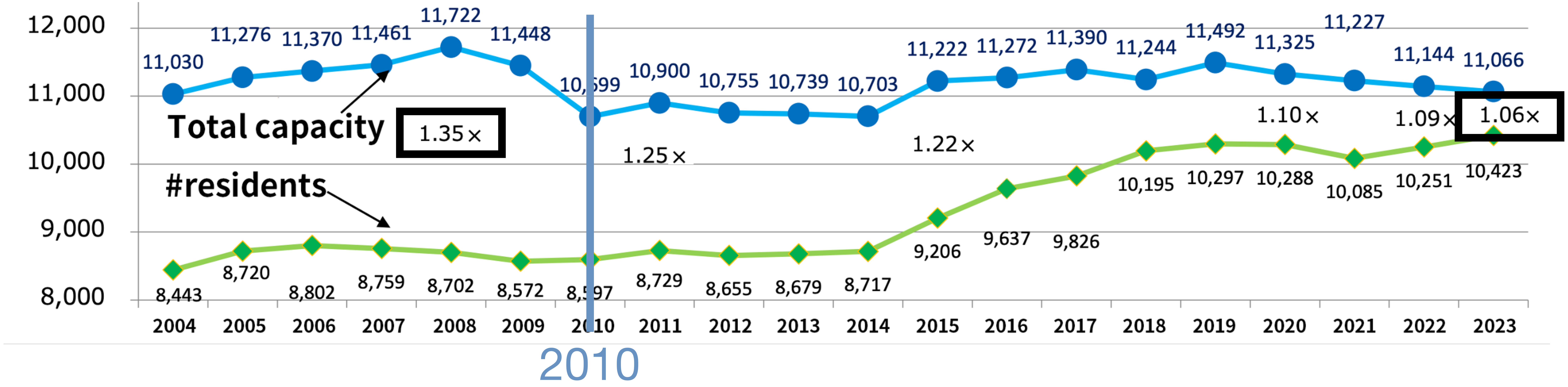
cap-based regulation in JRMP

- government has imposed cap-based regulation **since 2010.**



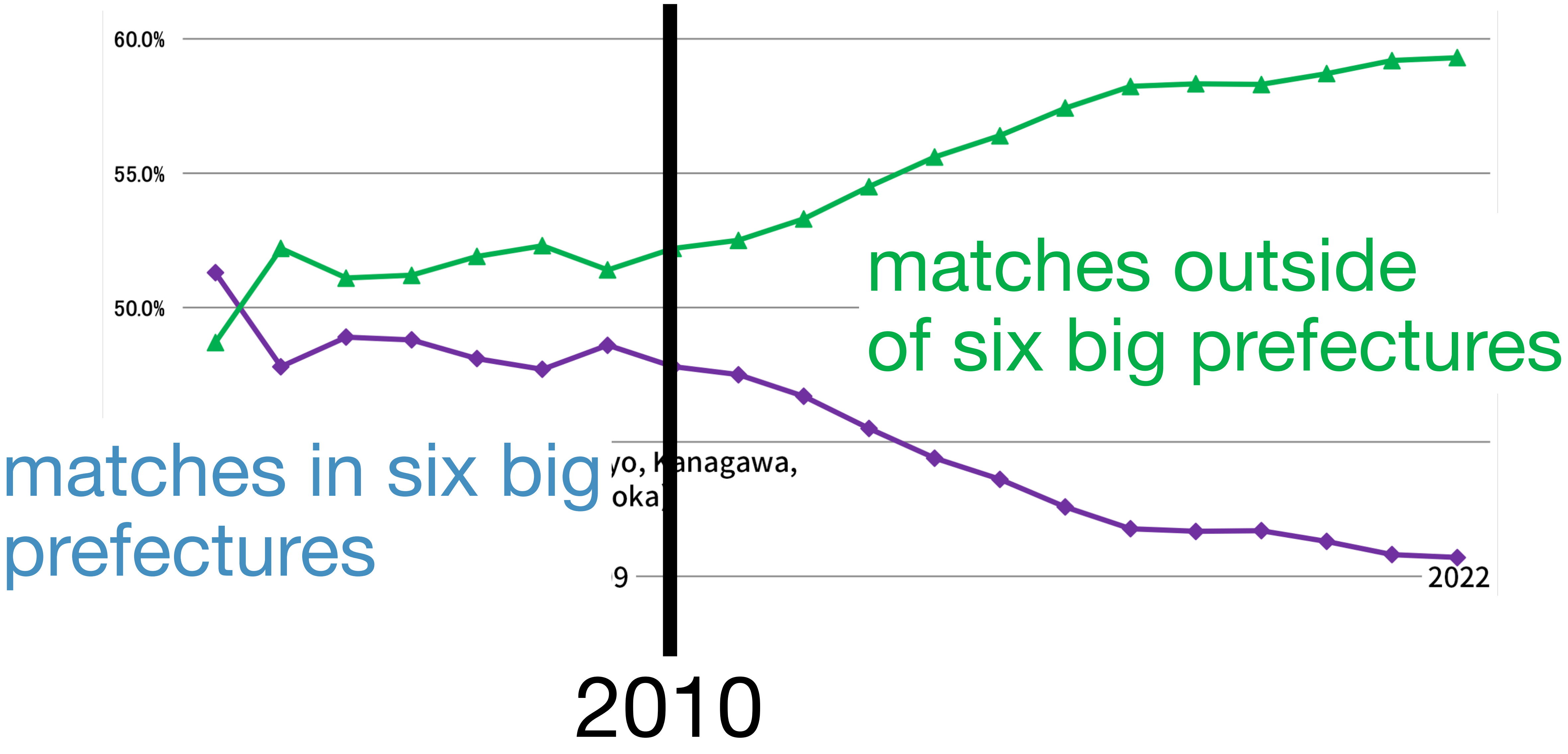
cap-based regulation in JRMP

- government has imposed cap-based regulation **since 2010**



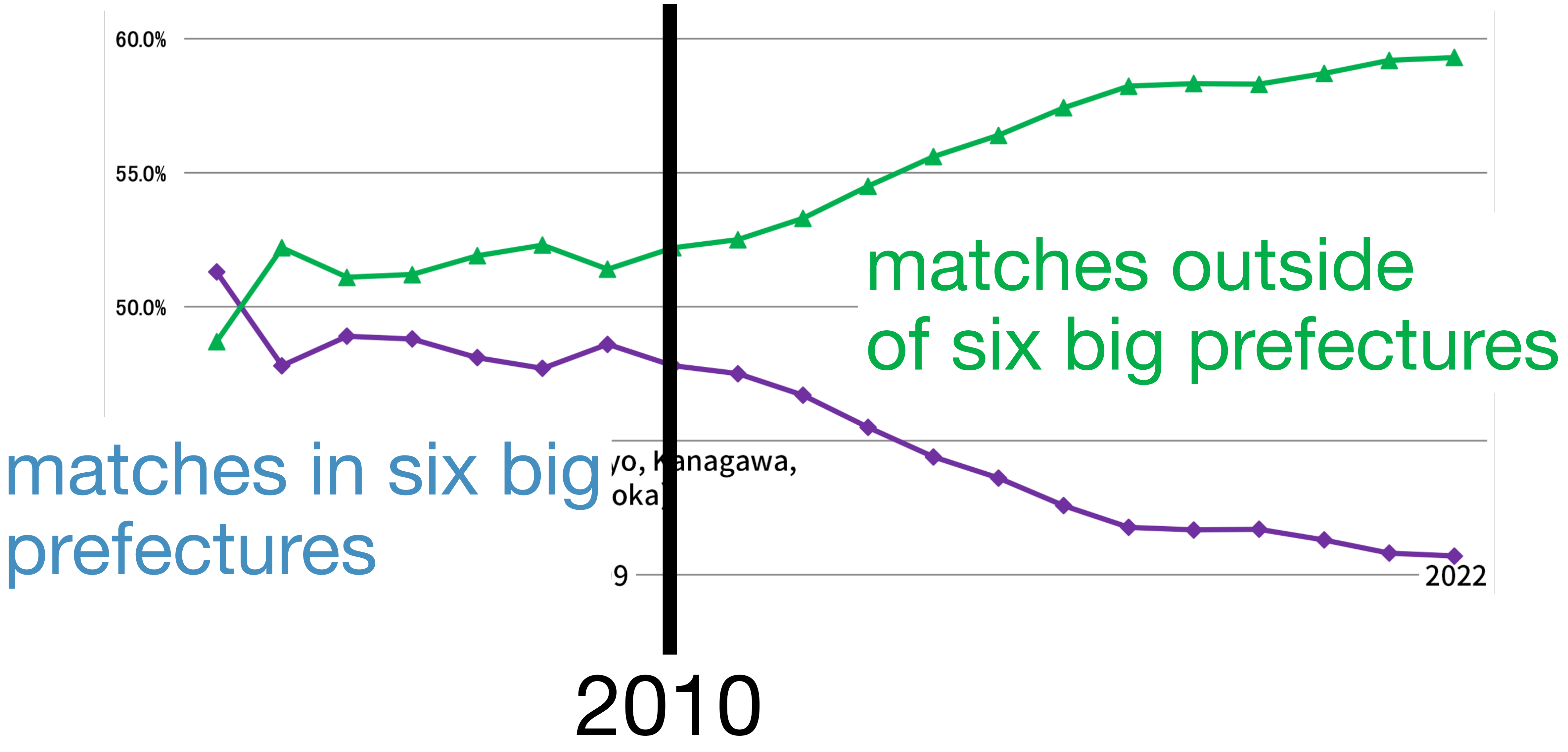
- **ratio of positions to residents** has decreased (**1.35** in 2008 to **1.06** in 2023)
- **share of positions in urban areas** has decreased (**42.3%** in 2004 to **35.1%** in 2023)

results of the regulation



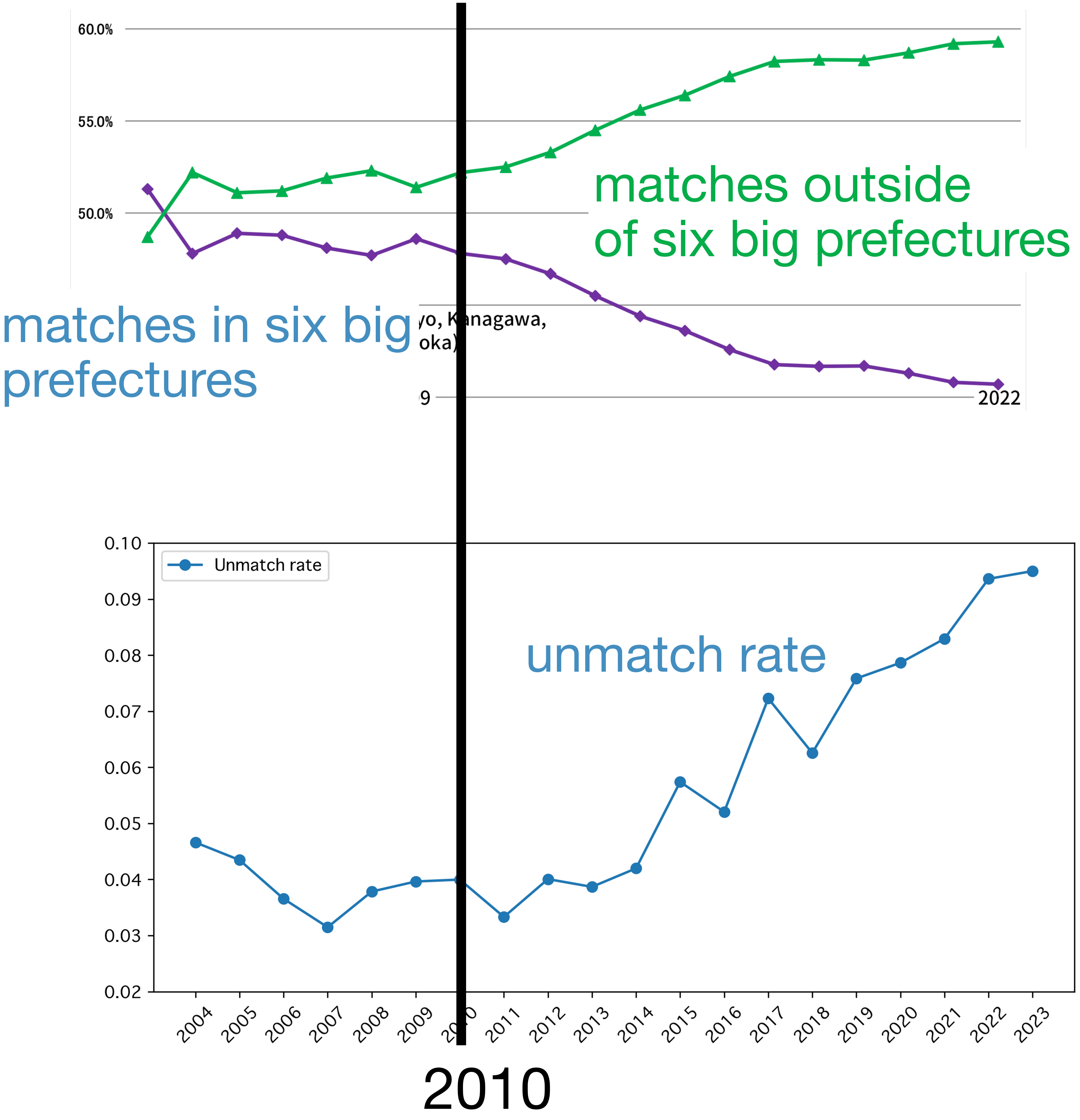
- share of matches in urban areas has decreased (+)

results of the regulation



- share of matches in urban areas has decreased (+)
- effects on rural areas are mixed (-)

results of the regulation



matches outside of six big prefectures

matches in six big prefectures

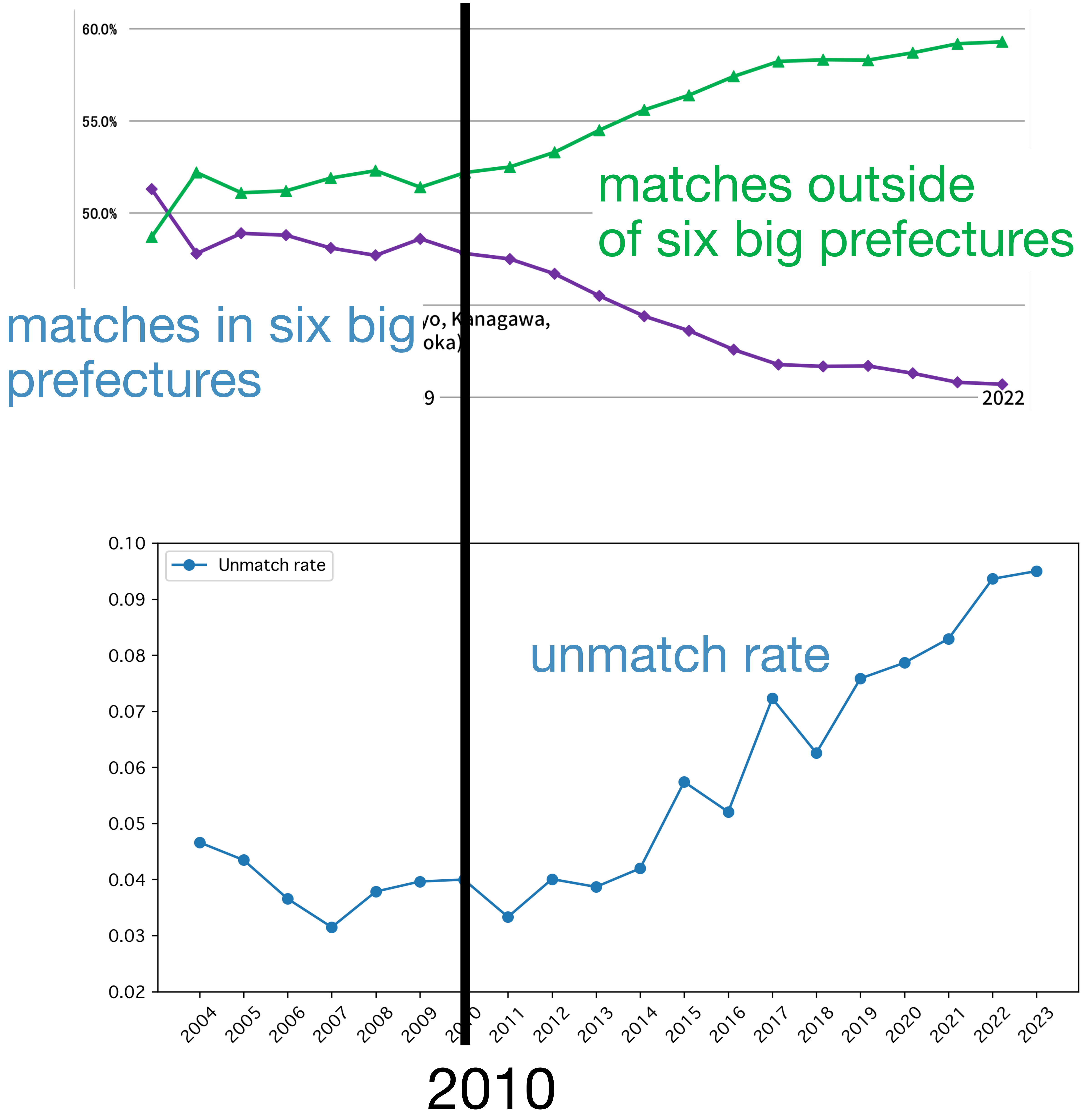
unmatch rate

2010

- share of matches in urban areas has decreased (+)
- effects on rural areas are mixed (-)

- unmatch rate has been increasing (-)
  - **4%** in 2010 → **9%+** in 2023

results of the regulation

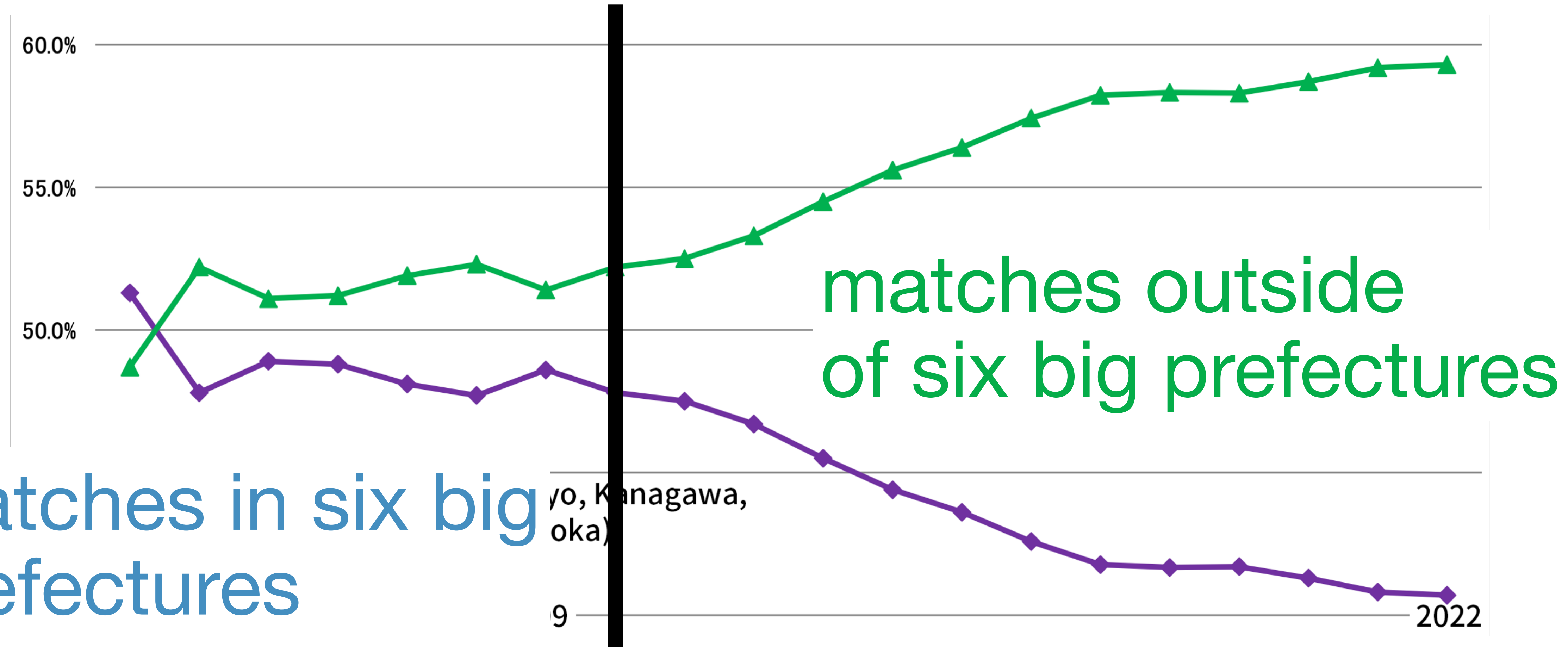


- share of matches in urban areas has decreased (+)
- effects on rural areas are mixed (-)

- unmatch rate has been increasing (-)
  - **4%** in 2010 → **9%+** in 2023

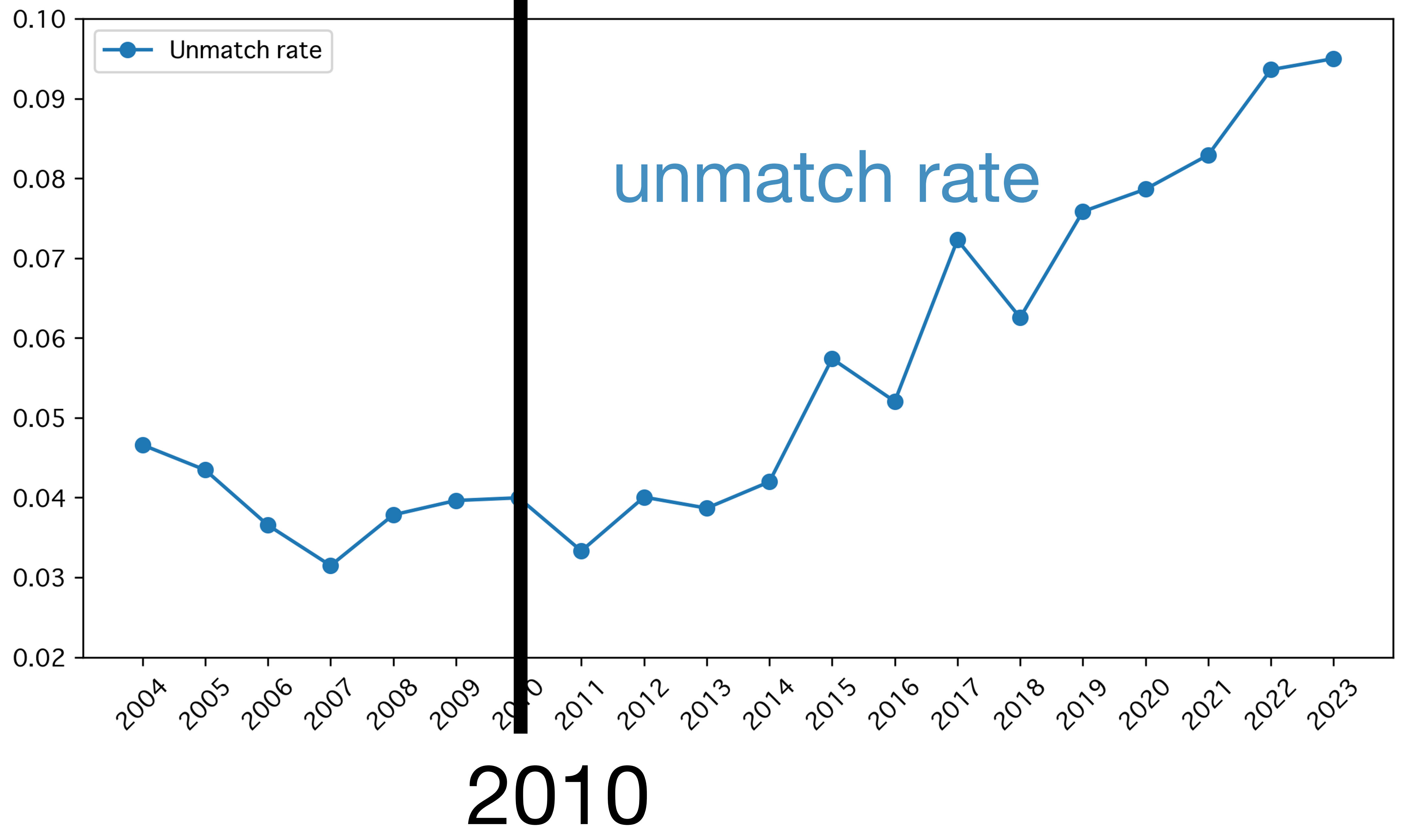
fail to find a position in the centralized process

results of the regulation



matches outside of six big prefectures

matches in six big prefectures



unmatch rate

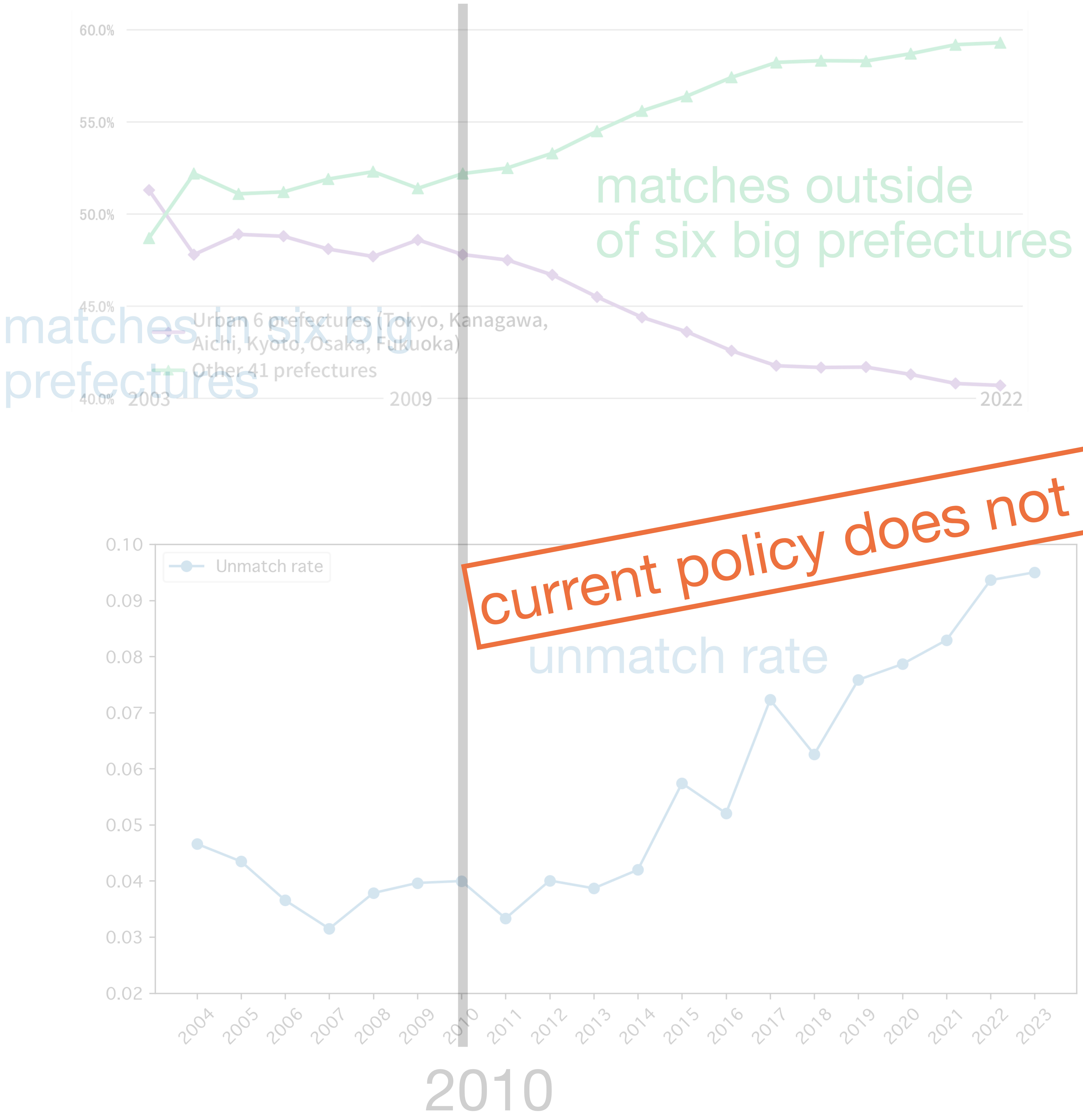
2010

- share of matches in urban areas has decreased (+)
- effects on rural areas are mixed (-)
- unmatch rate has been increasing (-)
  - **4%** in 2010 → **9%+** in 2023

fail to find a position in the centralized process

- search for a position in a rapid, decentralized mkt
- wait and reapply next year
- leave the medical career path

results of the regulation



- share of matches in urban areas has decreased (+)
- effects on rural areas are mixed (-)

current policy does not appear to be very effective

- unmatch rate has been increasing (-)
  - **4%** in 2010 → **9%+** in 2023

questions

- how effective is cap-based regulation?
- how do alternative policies, such as monetary interventions, compare?
- can we quantify their performance?

contribution

this paper:

- develops a framework to evaluate policies in matching mkt with distributional constraints
  - transferable utility matching model with **regional constraints**
  - "optimal" taxation policy generates higher social surplus than cap-based policies
  - in large mkts, an approx. opt. taxation policy can be computed from observable data

- applies the framework to a novel dataset from the Japan Residency Matching Program:
  - status quo cap-based regulation generates a significant welfare loss
  - moderate subsidy can achieve the same distributional goal, improving welfare

some relevant literature

- **matching with distributional constraints**
  - **Kamada and Kojima (2015)**, Abdulkadiroglu and Sonmez (2003), Ehlers, Hafalir, Yenmez, and Yildirim (2014), Kojima (2012), Hafalir, Yenmez, and Yildirim (2013), Fragiadakis and Troyan (2017) …
  - this literature uses the non-transferable utility (NTU) matching model
  - primarily focuses on how to **set caps** by modifying the deferred acceptance algorithm

some relevant literature

- **matching with distributional constraints**
  - **Kamada and Kojima (2015)**, Abdulkadiroglu and Sonmez (2003), Ehlers, Hafalir, Yenmez, and Yildirim (2014), Kojima (2012), Hafalir, Yenmez, and Yildirim (2013), Fragiadakis and Troyan (2017) …
  - this literature uses the non-transferable utility (NTU) matching model
  - primarily focuses on how to **set caps** by modifying the deferred acceptance algorithm
- **our paper**: uses the transferable utility (TU) matching model
  - accommodates **a broader class of policies,** including monetary interventions, with a clear welfare benchmark
  - accounts for **endogenous transfers** (e.g., salary adjustment in response to intervention)

model

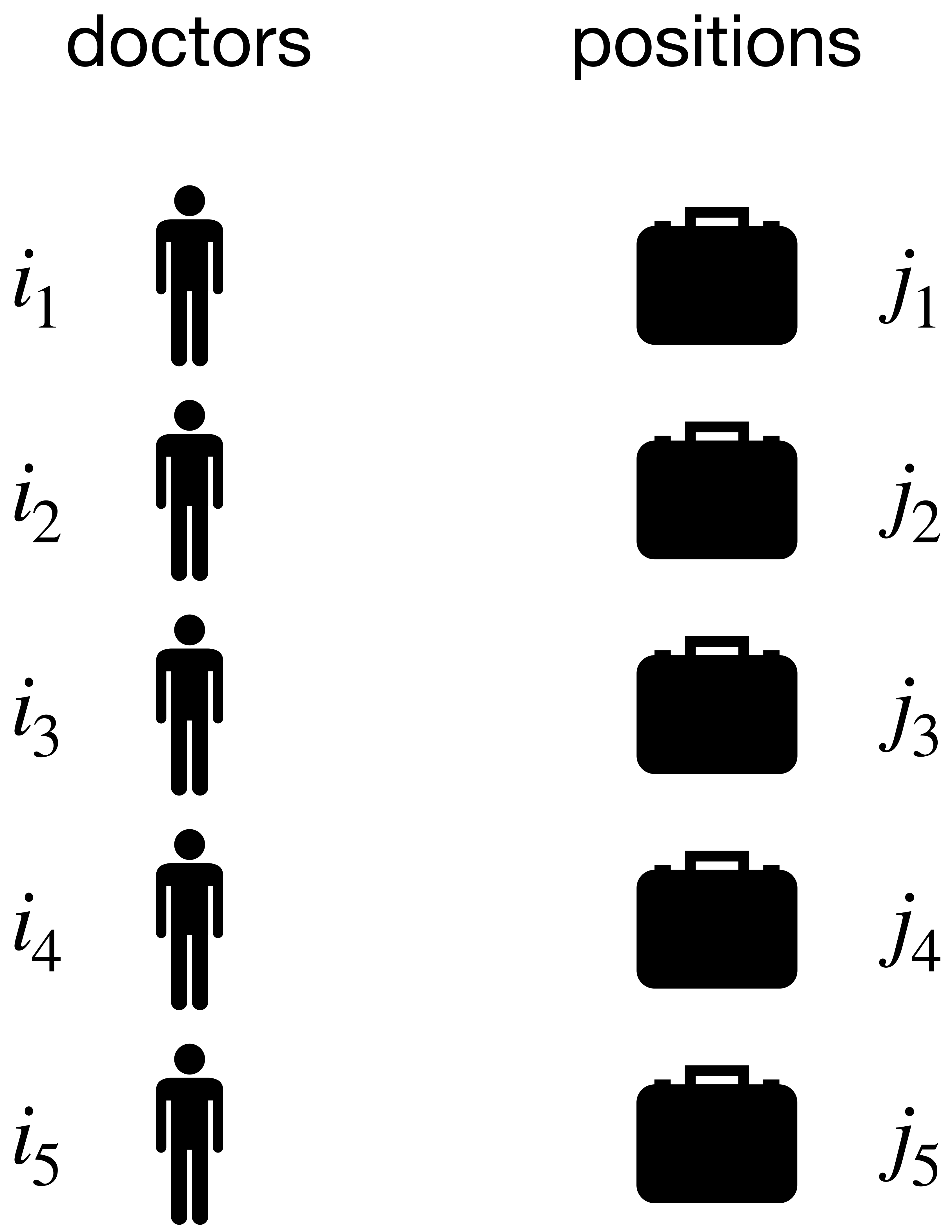
theoretical results

estimation

simulation

agents' problem

- **doctors** $i \in I$ and **positions** $j \in J$



doctors positions

i_1 j_1

i_2 j_2

i_3 j_3

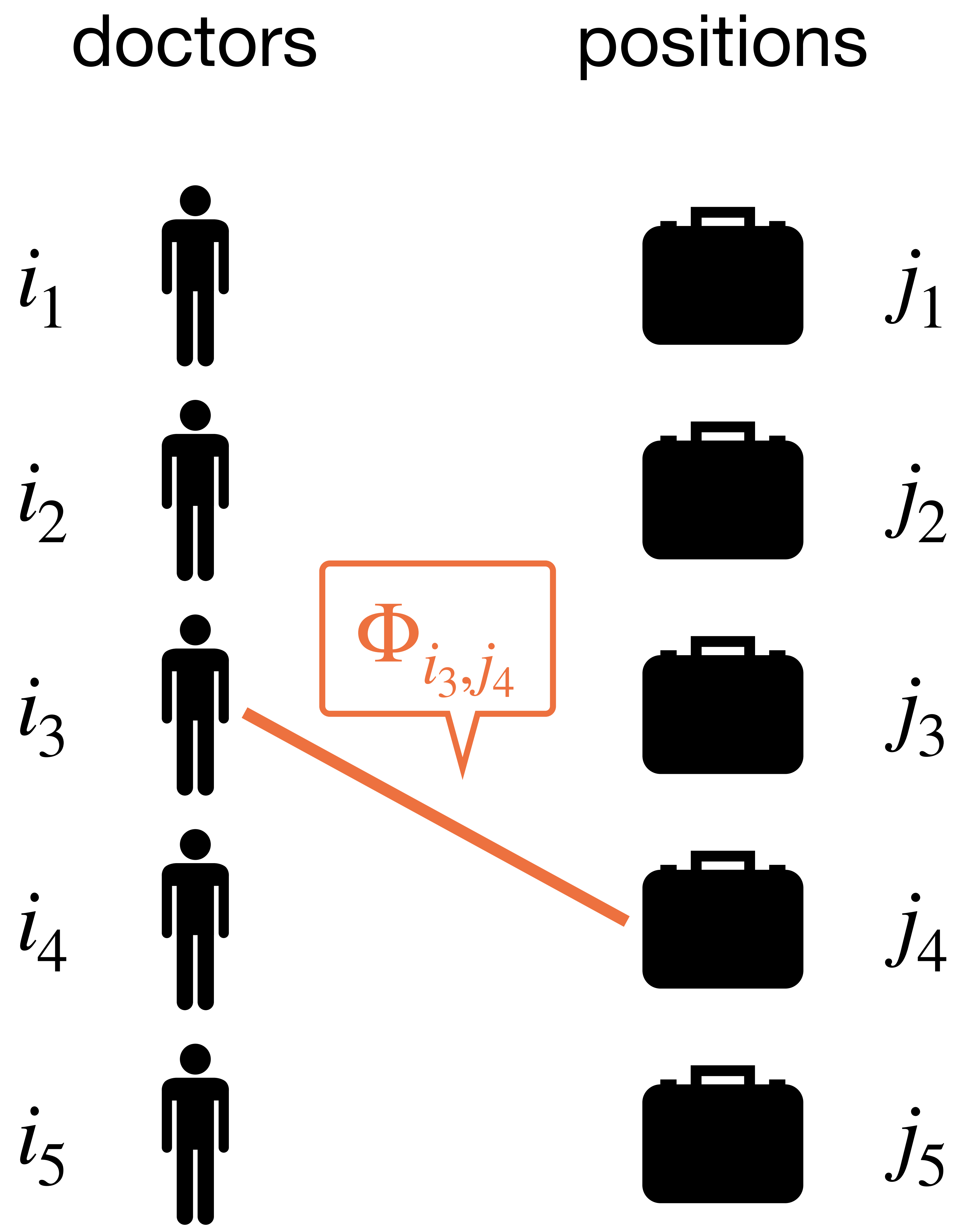
i_4 j_4

i_5 j_5

agents' problem

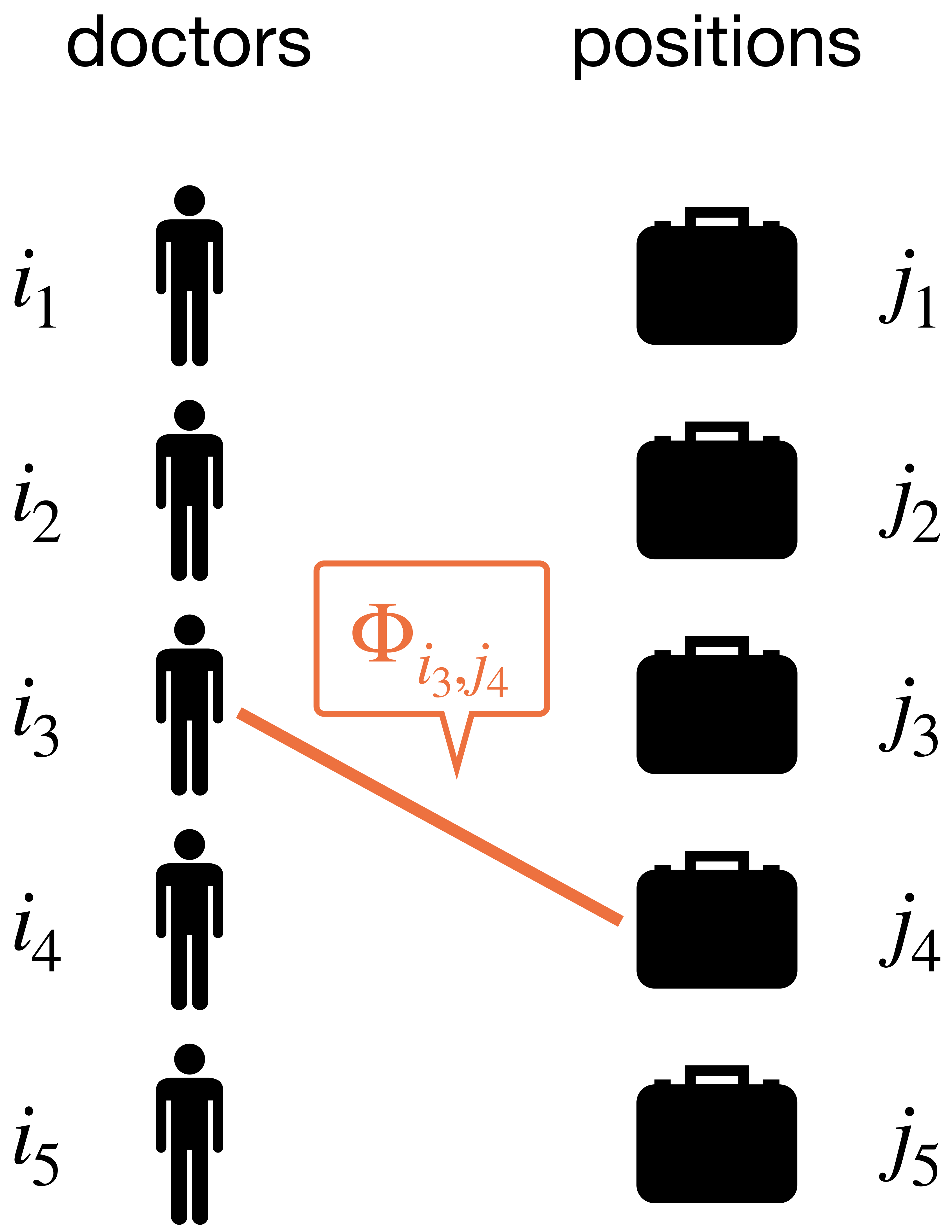
- **doctors** $i \in I$ and **positions** $j \in J$

- if i and j are matched, the pair generates a **joint surplus** $\Phi_{ij} \in \mathbb{R}$ and splits it



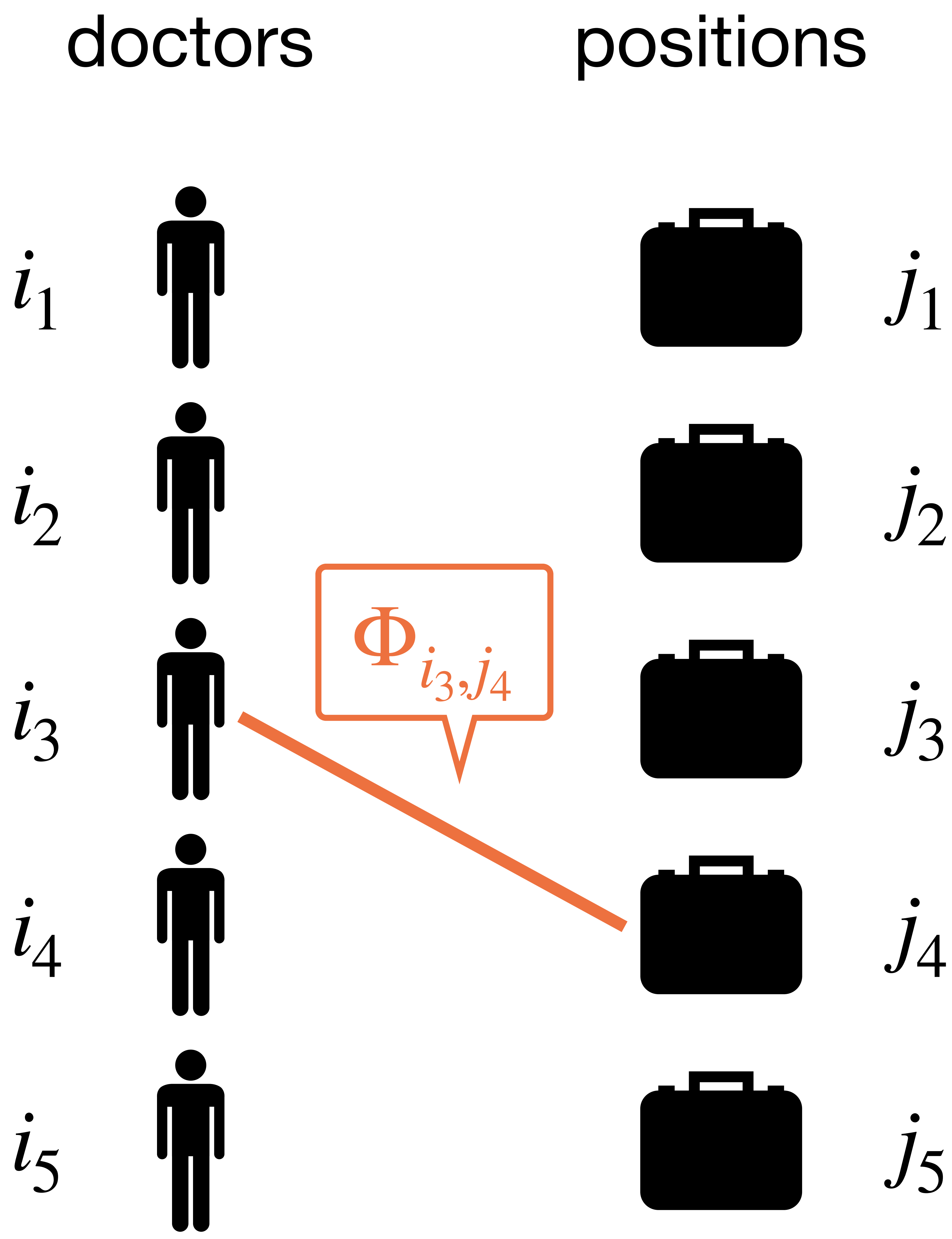
agents' problem

- **doctors** $i \in I$ and **positions** $j \in J$
- if i and j are matched, the pair generates a **joint surplus** $\Phi_{ij} \in \mathbb{R}$ and splits it
- a matching is a binary-valued matrix $d = (d_{ij})_{i,j}$
  - $d_{ij} = 1$ iff i and j are matched



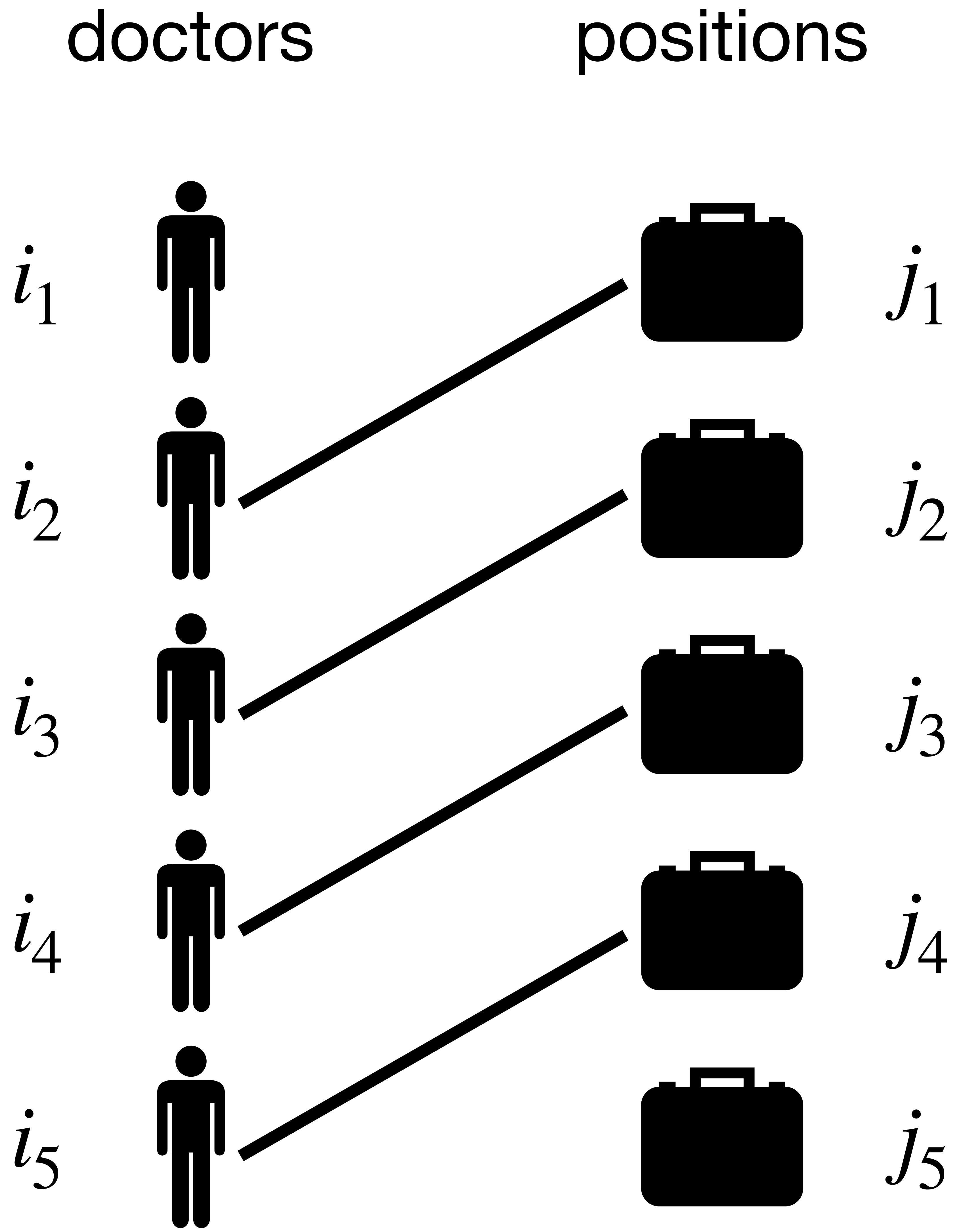
agents' problem

- **doctors** $i \in I$ and **positions** $j \in J$
- if i and j are matched, the pair generates a **joint surplus** $\Phi_{ij} \in \mathbb{R}$ and splits it
- a matching is a binary-valued matrix $d = (d_{ij})_{i,j}$
  - $d_{ij} = 1$ iff i and j are matched
- agents know $(\Phi_{ij})_{i,j}$ and form a matching, which yields a stable outcome



agents' problem

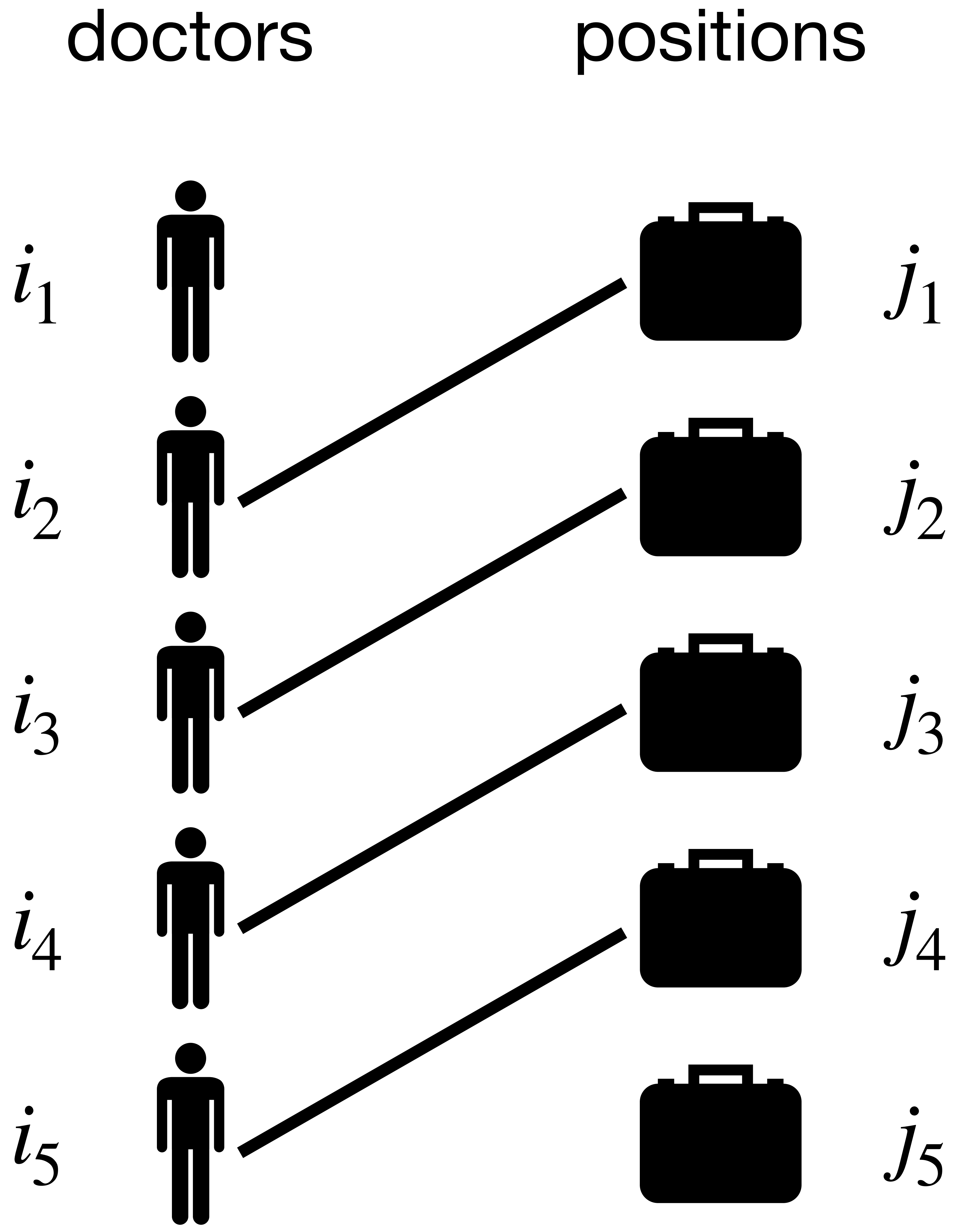
def. (stable outcome): given (Φ_{ij}), (d, u, v) is a stable outcome if it is **individually rational** and has **no blocking pair**



agents' problem

def. (stable outcome): given (Φ_{ij}),
(d, u, v) is a stable outcome if it is
individually rational and has **no blocking pair**

$$u = (u_i)_{i \in I},\ v = (v_j)_{j \in J}$$

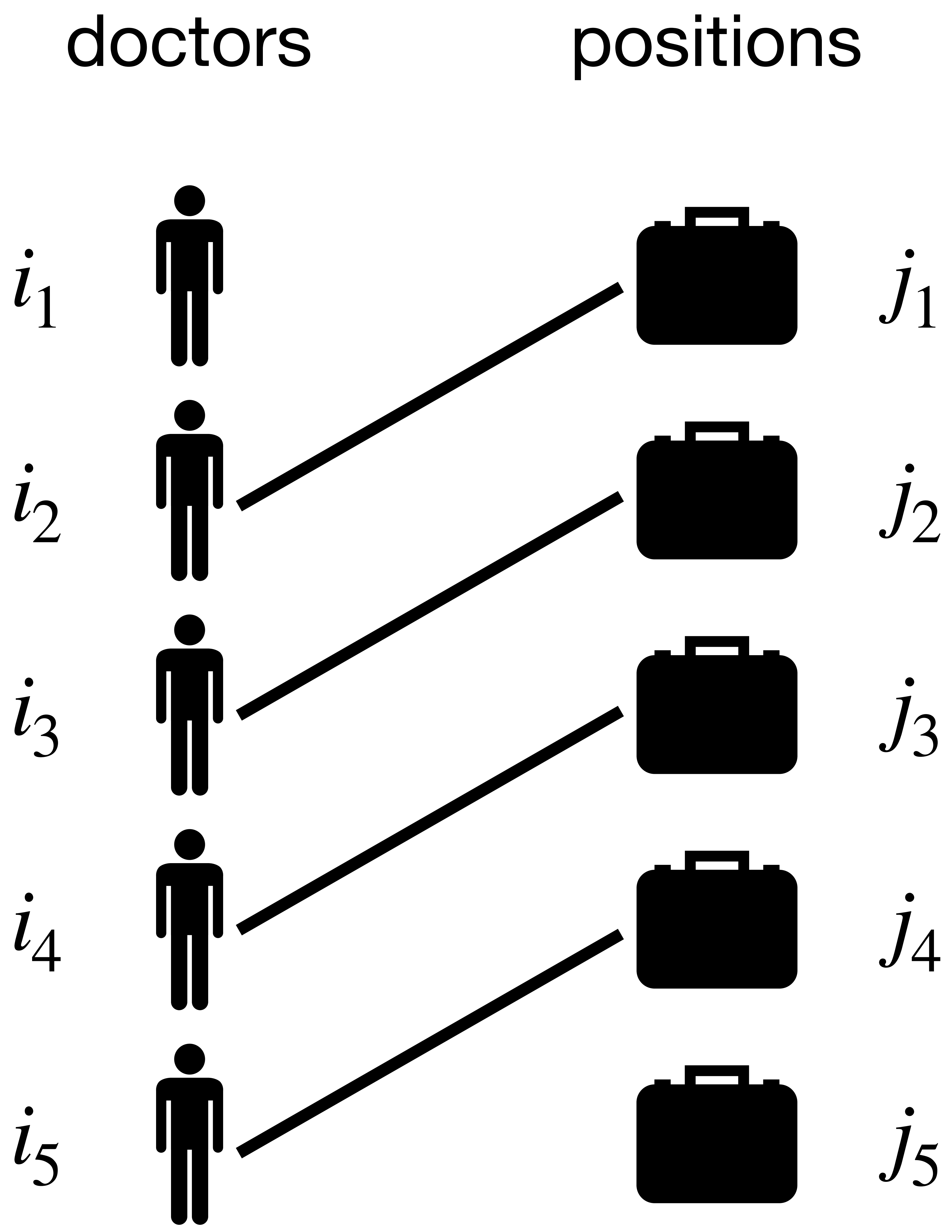


agents' problem

def. (stable outcome): given (Φ_{ij}), (d, u, v) is a stable outcome if it is **individually rational** and has **no blocking pair**

$$u = (u_i)_{i \in I},\ v = (v_j)_{j \in J}$$

doctor i's equilibrium payoff

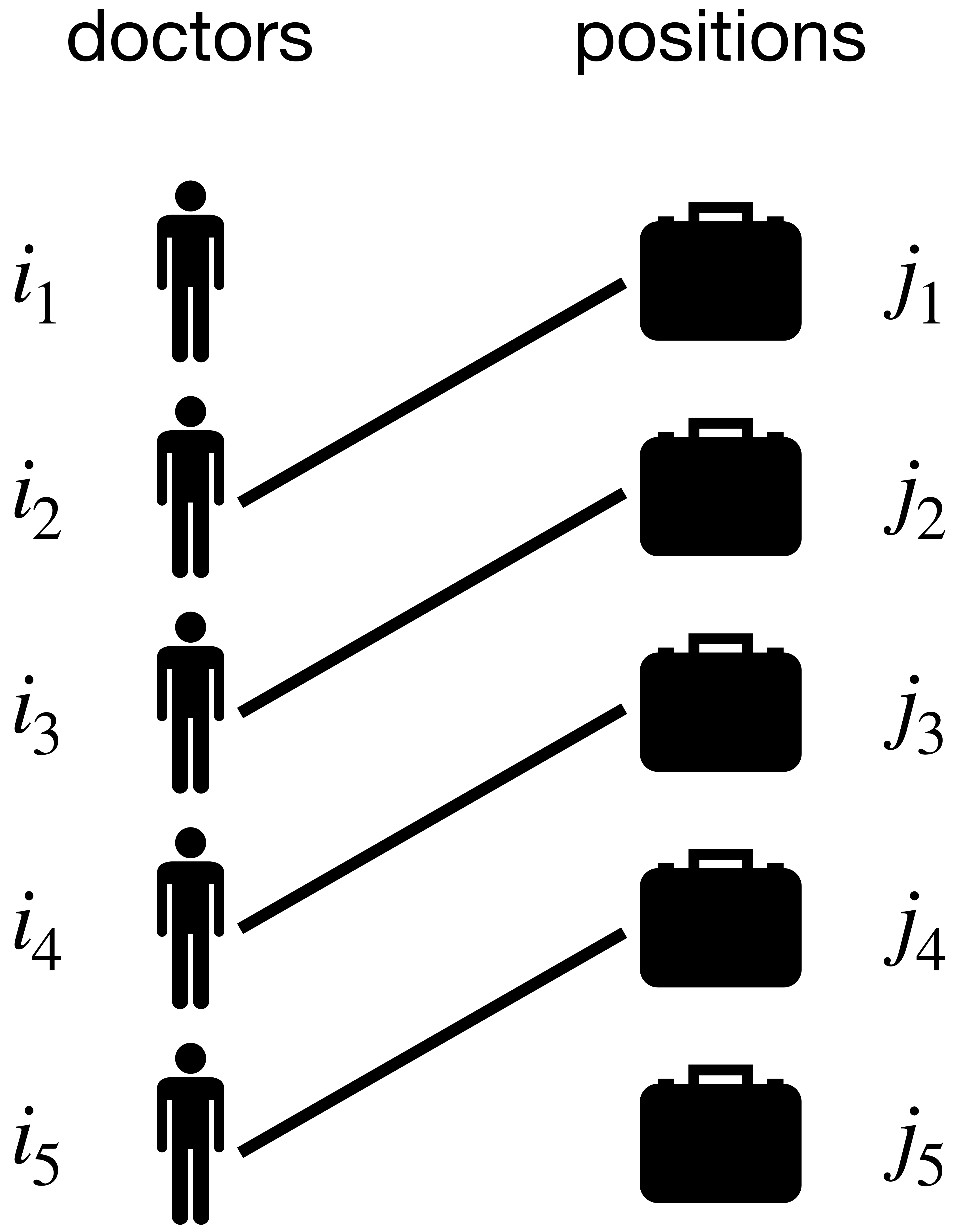


agents' problem

def. (stable outcome): given (Φ_{ij}), (d, u, v) is a stable outcome if it is **individually rational** and has **no blocking pair**

$$u = (u_i)_{i \in I},\ v = (v_j)_{j \in J}$$

position j's equilibrium payoff



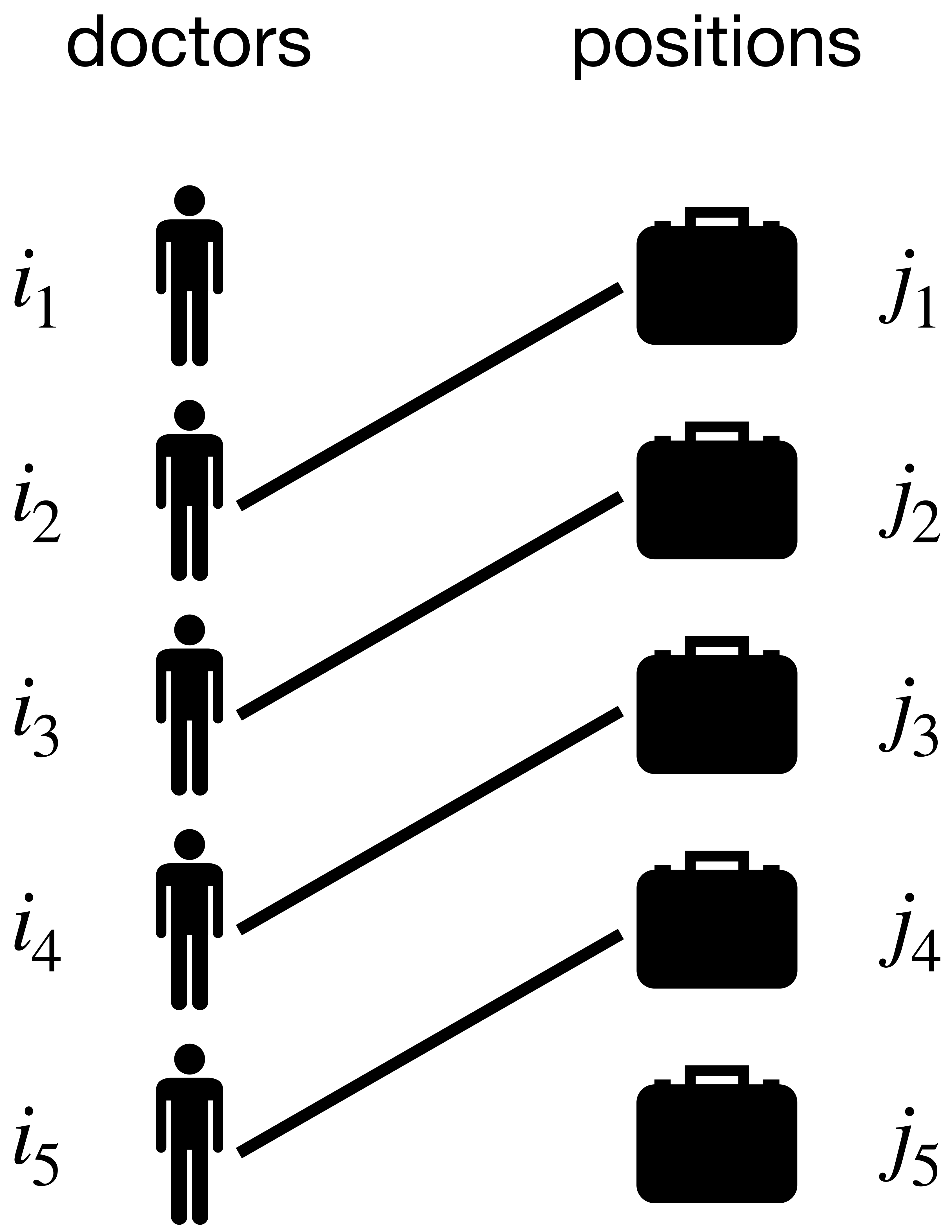
agents' problem

def. (stable outcome): given (Φ_{ij}),
(d, u, v) is a stable outcome if it is
individually rational and has **no blocking pair**

individual rationality:

$$u_i \geq \Phi_{i,0},\ v_j \geq \Phi_{0,j}$$

payoffs when unmatched



agents' problem

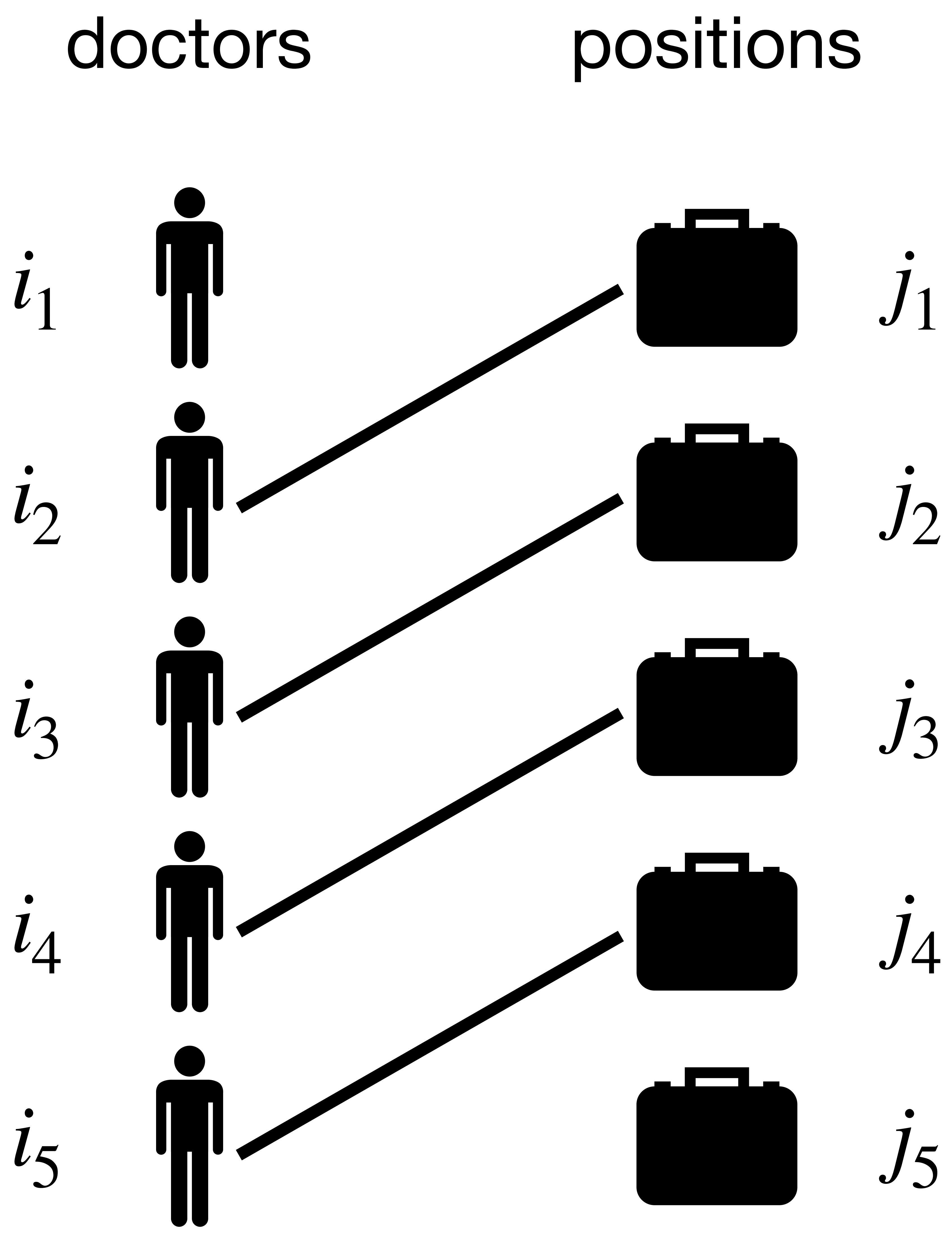
def. (stable outcome): given (Φ_{ij}), (d, u, v) is a stable outcome if it is **individually rational** and has **no blocking pair**

individual rationality:

$$u_i \geq \Phi_{i,0},\ v_j \geq \Phi_{0,j}$$

payoffs when unmatched

$$u_i = \Phi_{i,0} \quad \text{if } d_{ij} = 0 \text{ for all } j$$
$$v_j = \Phi_{0,j} \quad \text{if } d_{ij} = 0 \text{ for all } i$$



agents' problem

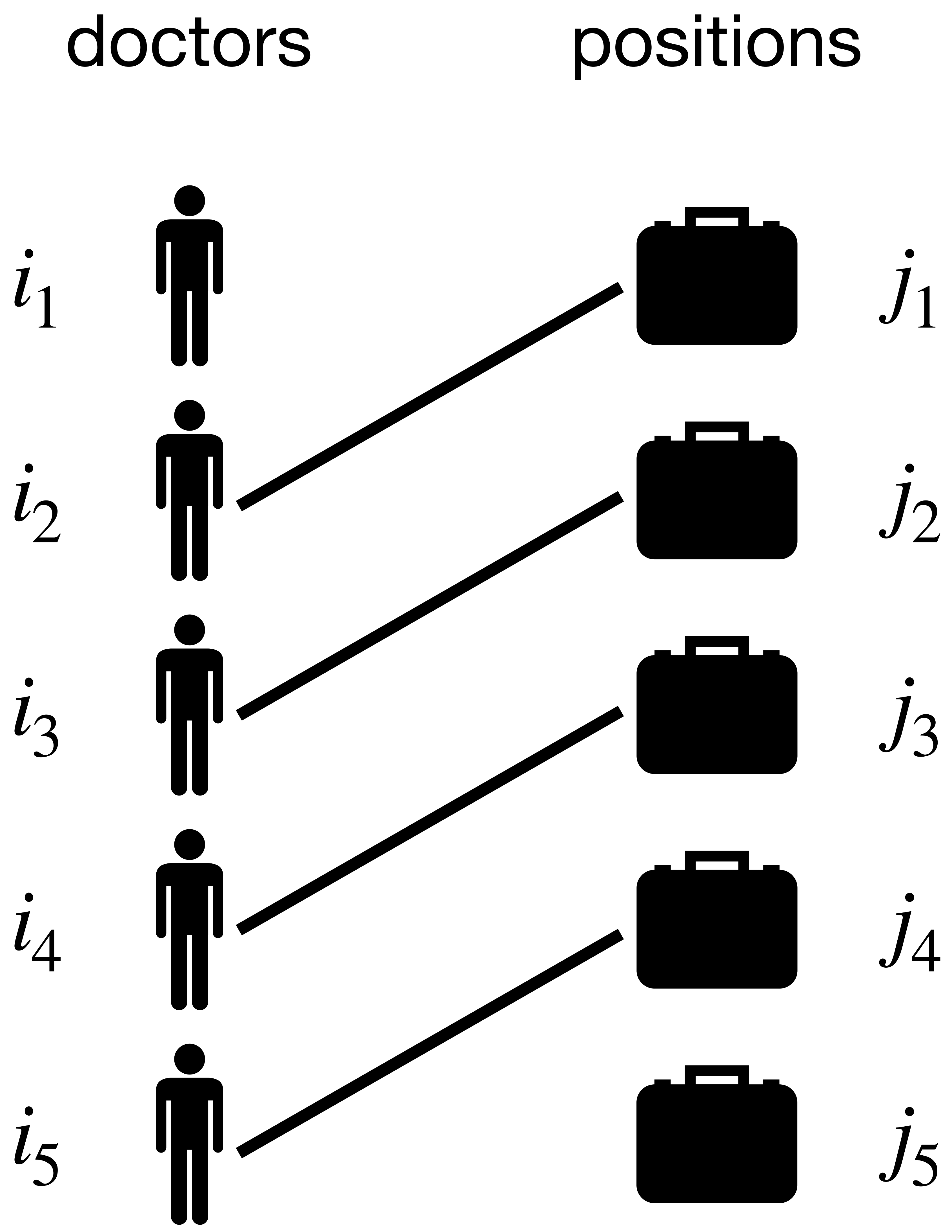
def. (stable outcome): given (Φ_{ij}), (d, u, v) is a stable outcome if it is **individually rational** and has **no blocking pair**

no blocking pair:

doctor i and position j **block** the matching if

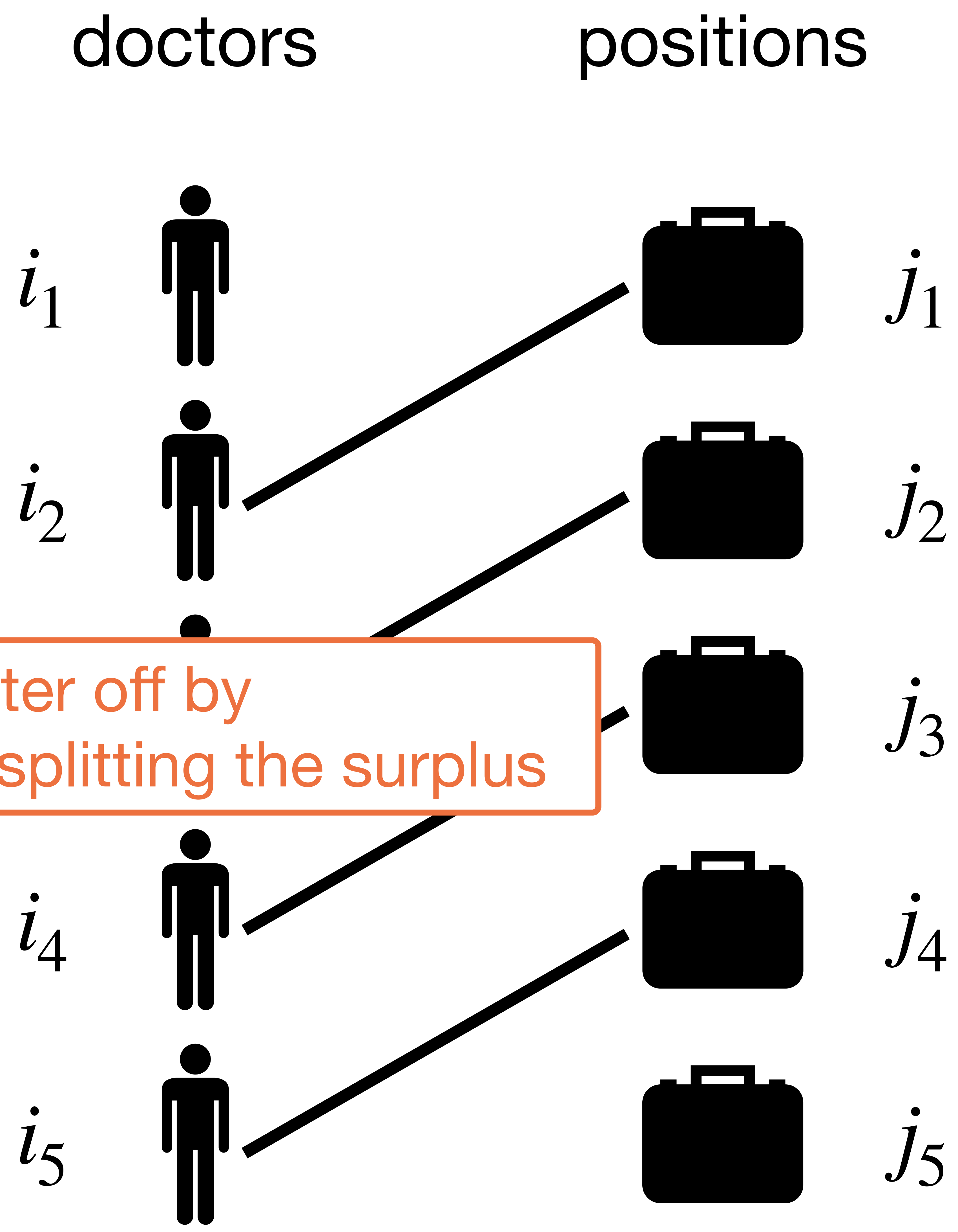
$$u_i + v_j < \Phi_{ij}$$

i's current payoff j's current payoff



agents' problem

def. (stable outcome): given (Φ_{ij}), (d, u, v) is a stable outcome if it is **individually rational** and has **no blocking pair**



no blocking pair:

doctor i and position j **block** the matching if

$$\boxed{u_i} + \boxed{v_j} < \Phi_{ij}$$

i's current payoff j's current payoff

both can be better off by forming a new match and splitting the surplus

agents' problem

def. (stable outcome): given (Φ_{ij}), (d, u, v) is a stable outcome if it is **individually rational** and has **no blocking pair**

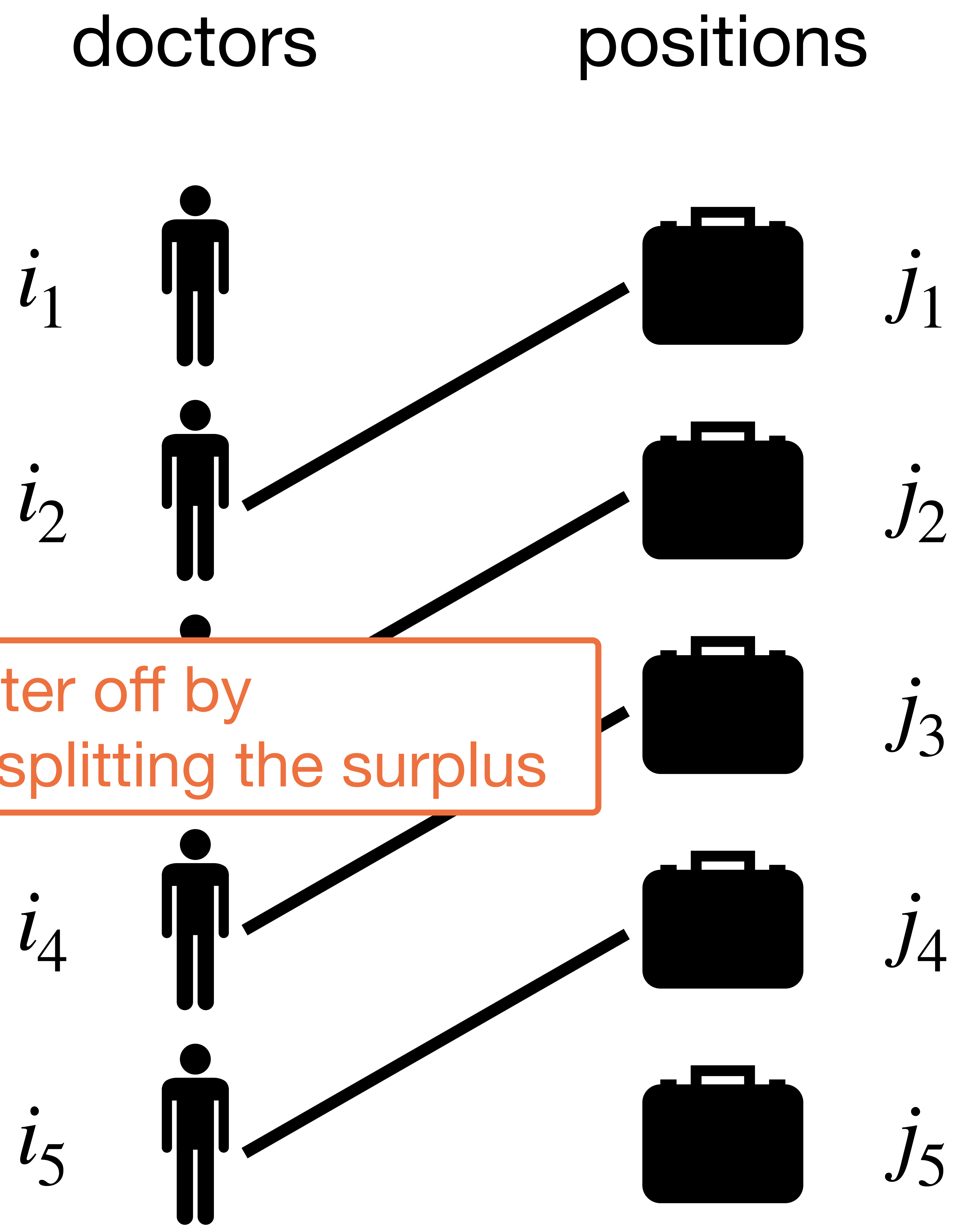
no blocking pair:

both can be better off by forming a new match and splitting the surplus

doctor i and position j **block** the matching if

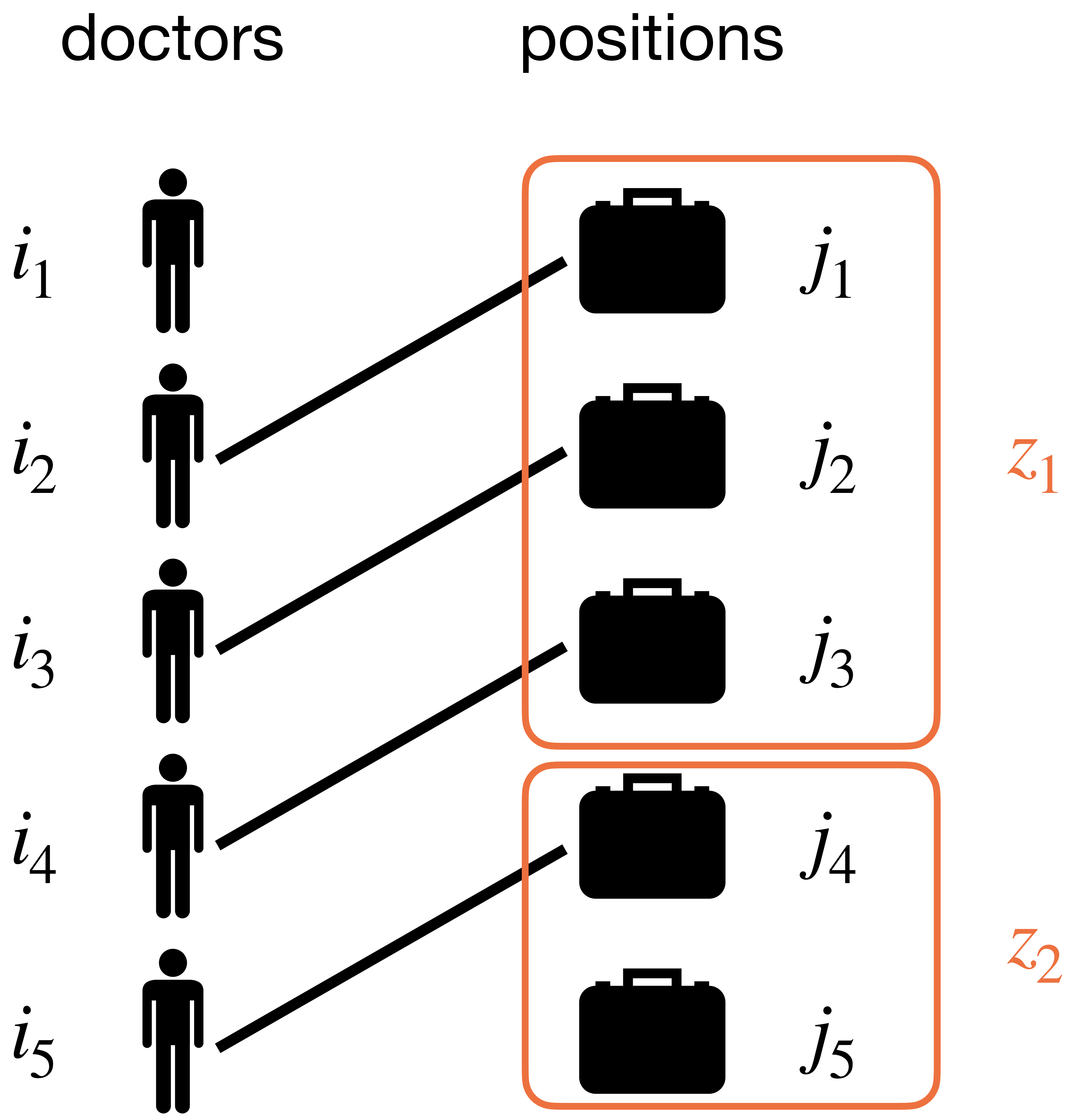
$$\underbrace{u_i}_{i\text{'s current payoff}} + \underbrace{v_j}_{j\text{'s current payoff}} < \Phi_{ij}$$

$$u_i + v_j = \Phi_{ij} \quad \text{if } d_{ij} = 1$$



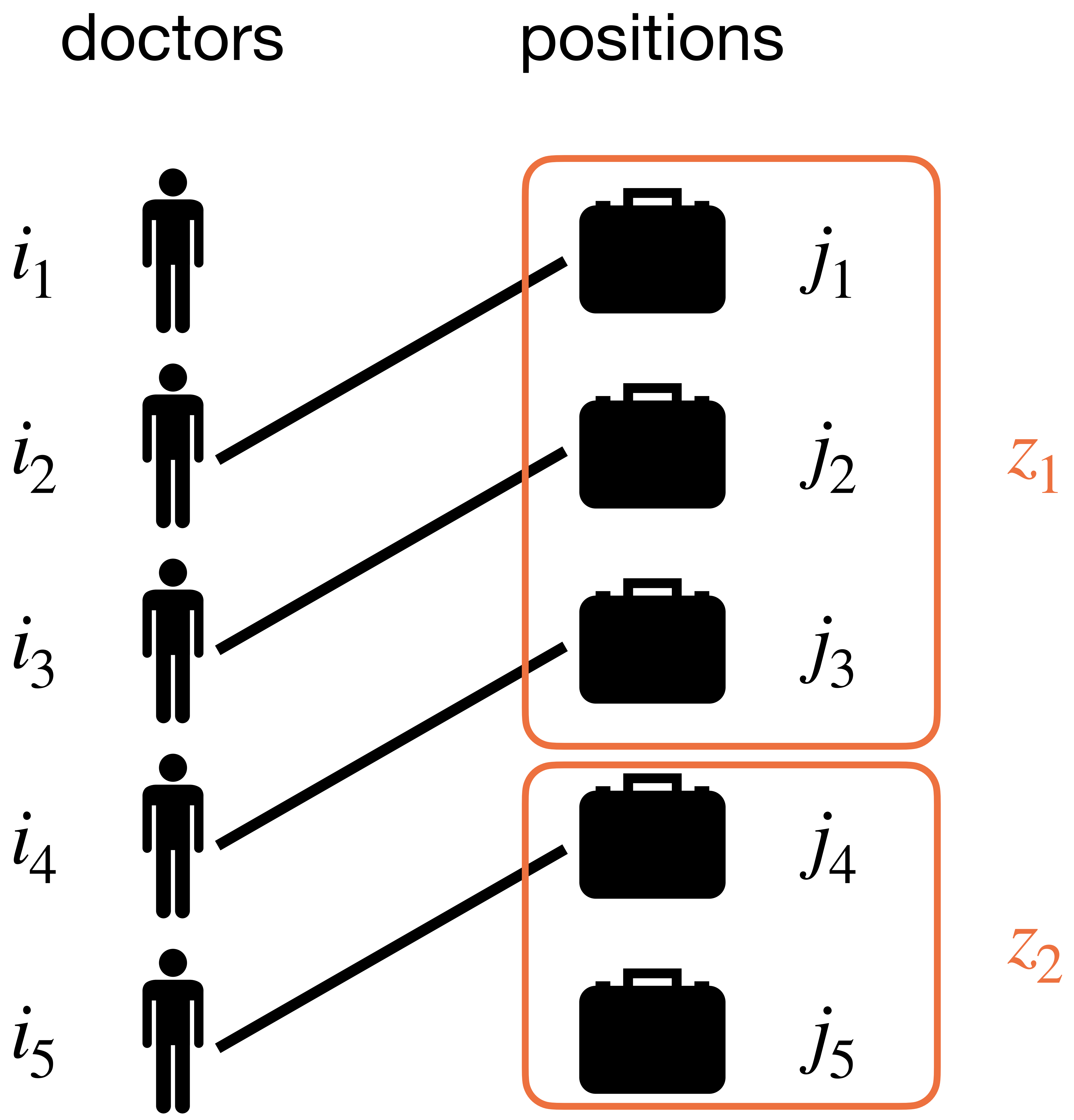
policymaker's problem

- each position belongs to a region $z \in Z$



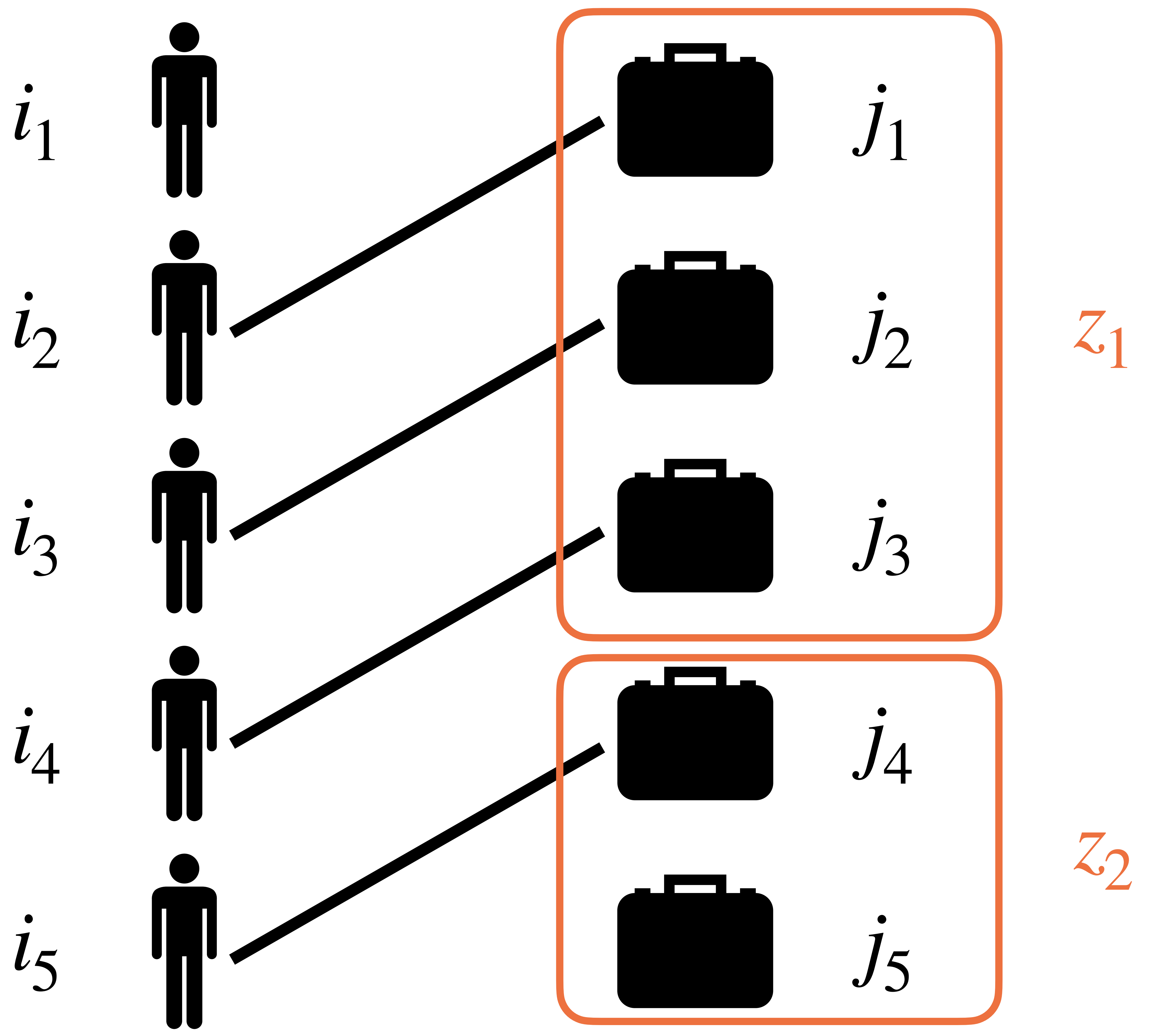
policymaker's problem

- each position belongs to a region $z \in Z$
- policymaker (PM) faces exogenous regional constraints
  - **lower and upper bounds** on # of matches in each region



policymaker's problem

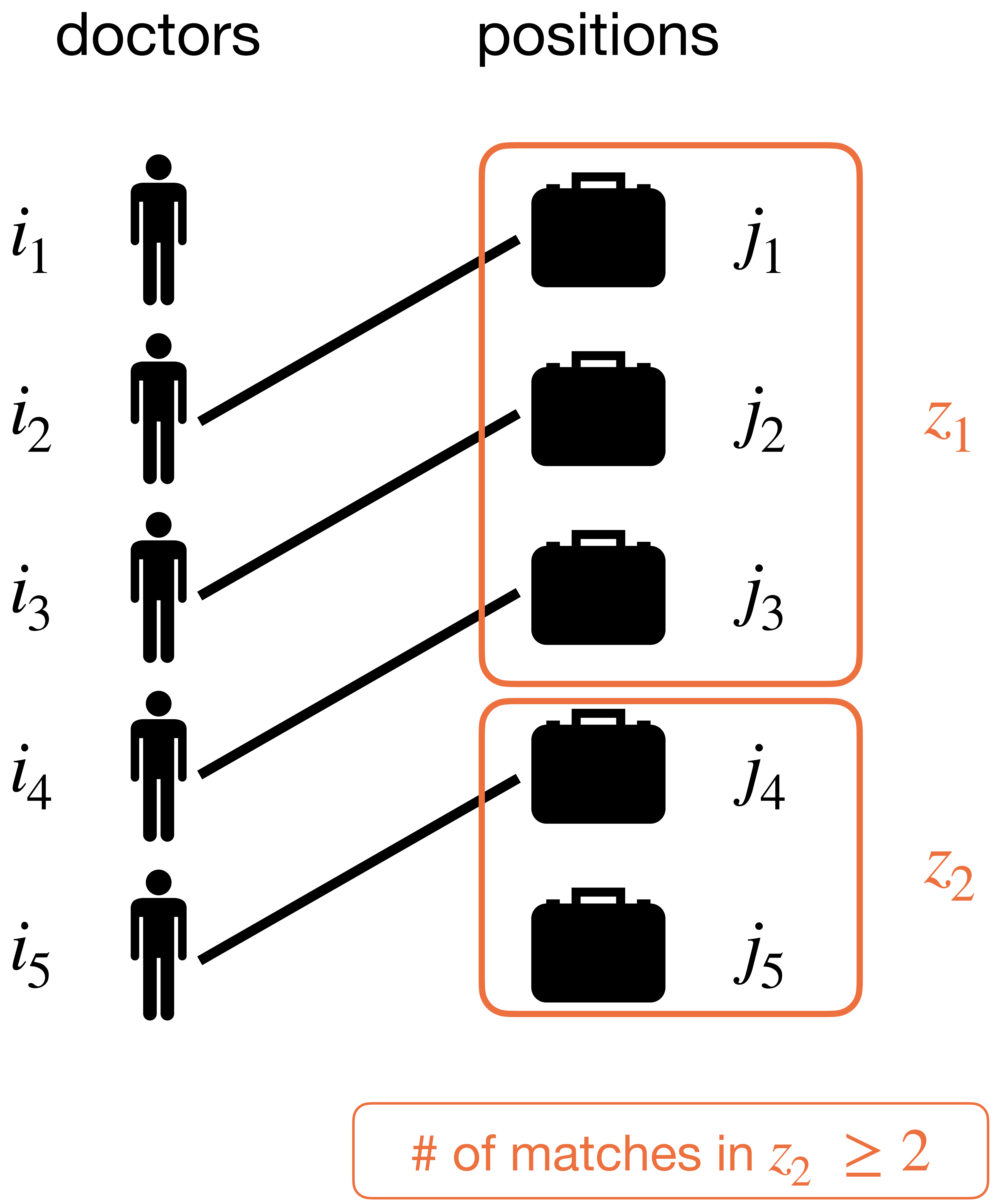
- each position belongs to a region $z \in Z$
- policymaker (PM) faces exogenous regional constraints
  - **lower and upper bounds** on # of matches in each region



# of matches in z_2 ≥ 2

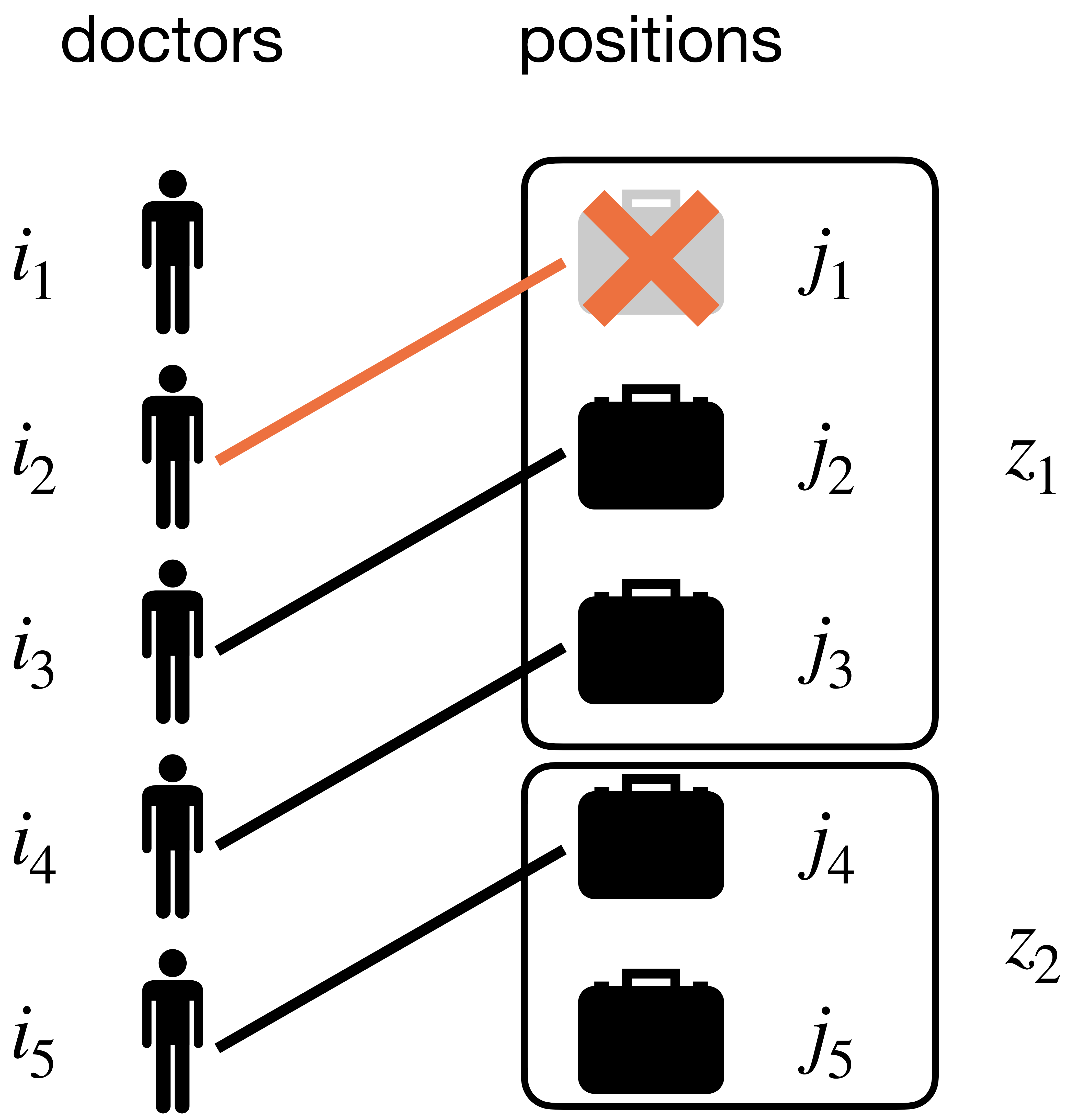
policymaker's problem

- each position belongs to a region $z \in Z$
- policymaker (PM) faces exogenous regional constraints
  - **lower and upper bounds** on # of matches in each region
- without intervention, the matching formed by the agents may not satisfy regional constraints



cap-based policy

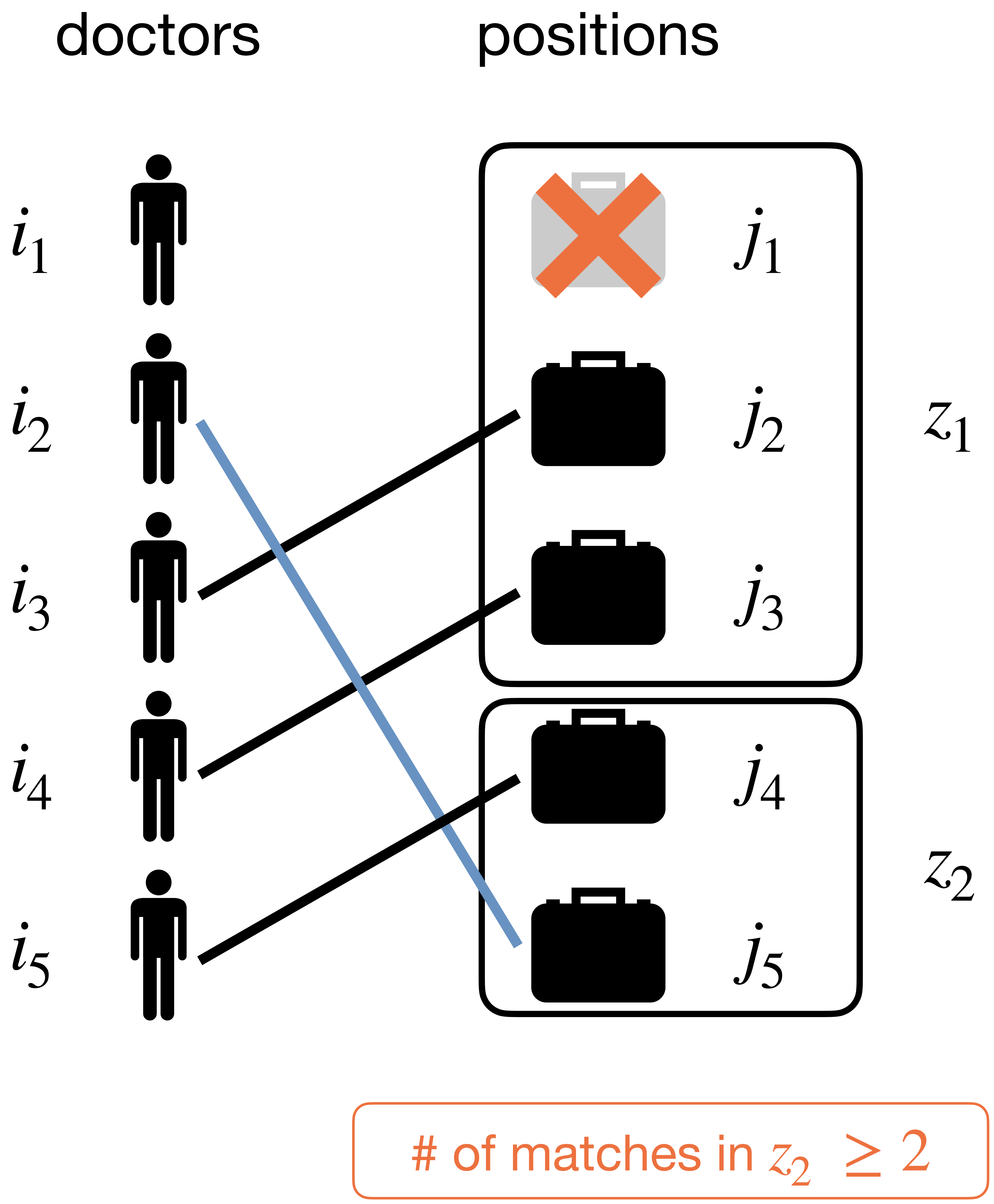
- cap-based policy removes positions from high-demand regions



# of matches in $z_2 \geq 2$

cap-based policy

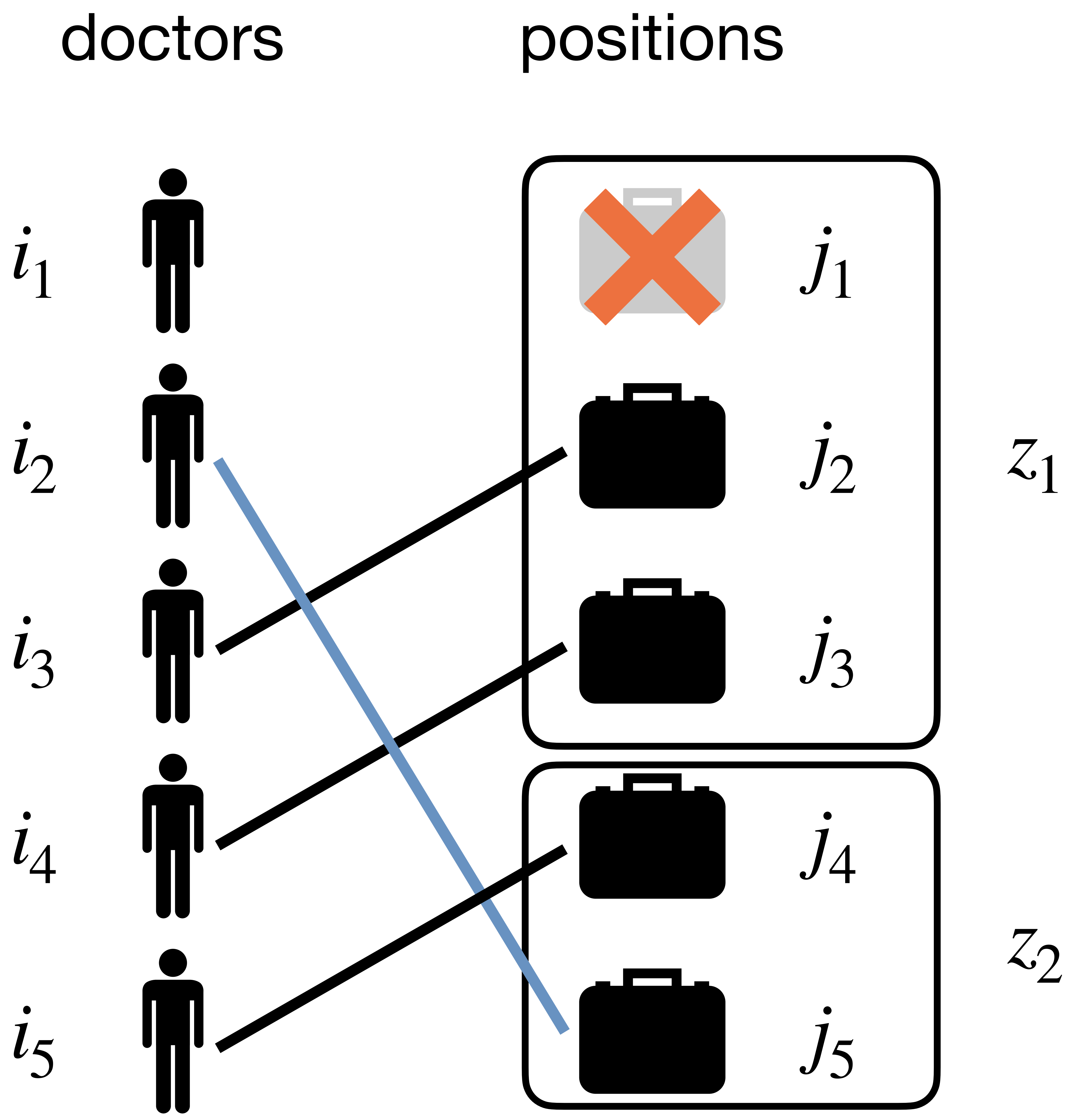
- cap-based policy removes positions from high-demand regions to encourage inflow into low-demand regions



cap-based policy

- cap-based policy removes positions from high-demand regions to encourage inflow into low-demand regions

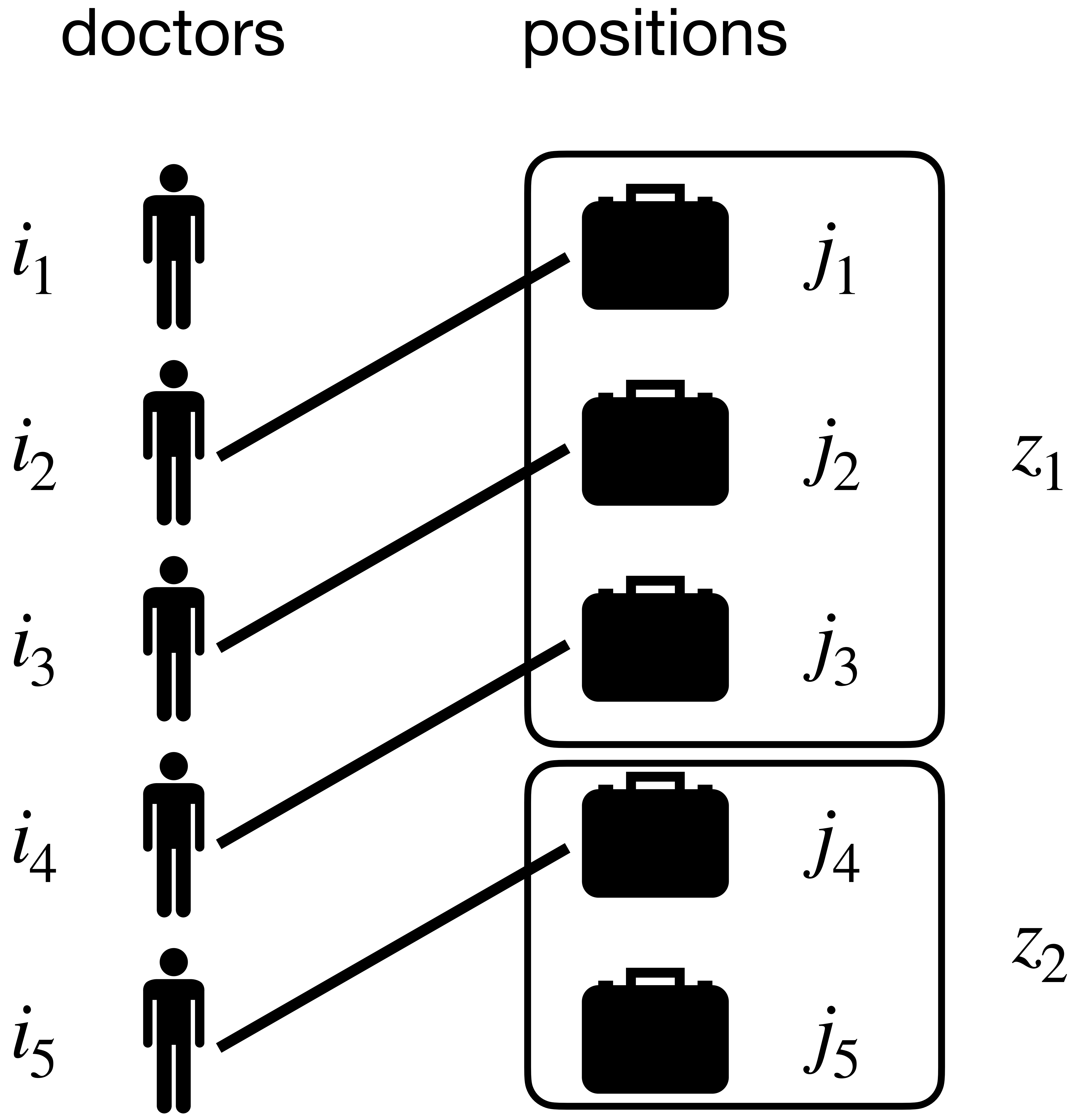
- agents form a stable outcome over the remaining slots



# of matches in z_2 ≥ 2

taxation policy

- alternative: taxation policy
  - a tax $w_z \in \mathbb{R}$ is applied uniformly to all matches in region z

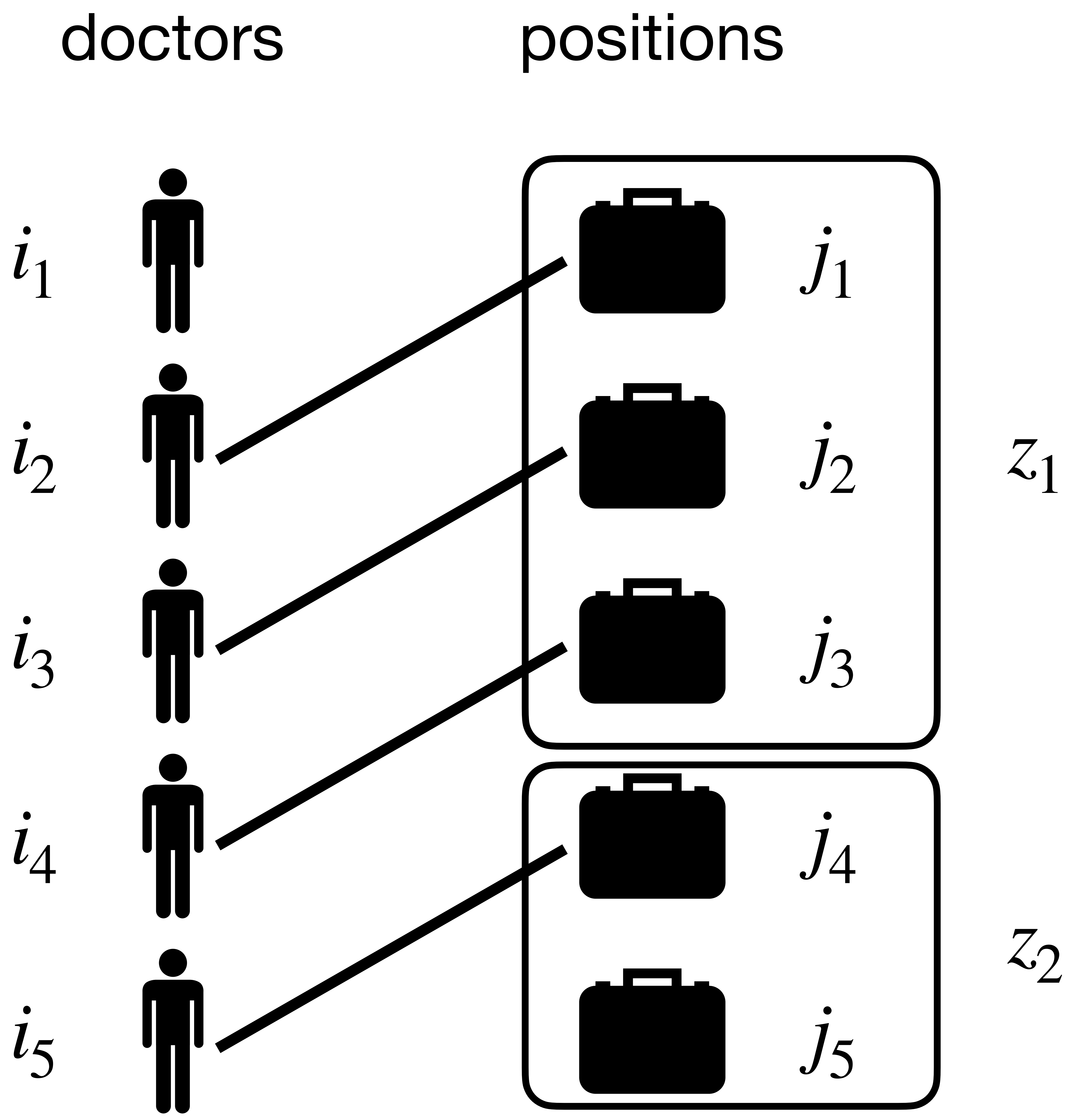


# of matches in $z_2 \geq 2$

taxation policy

- alternative: taxation policy
  - a tax $w_z \in \mathbb{R}$ is applied uniformly to all matches in region z

negative tax = subsidy



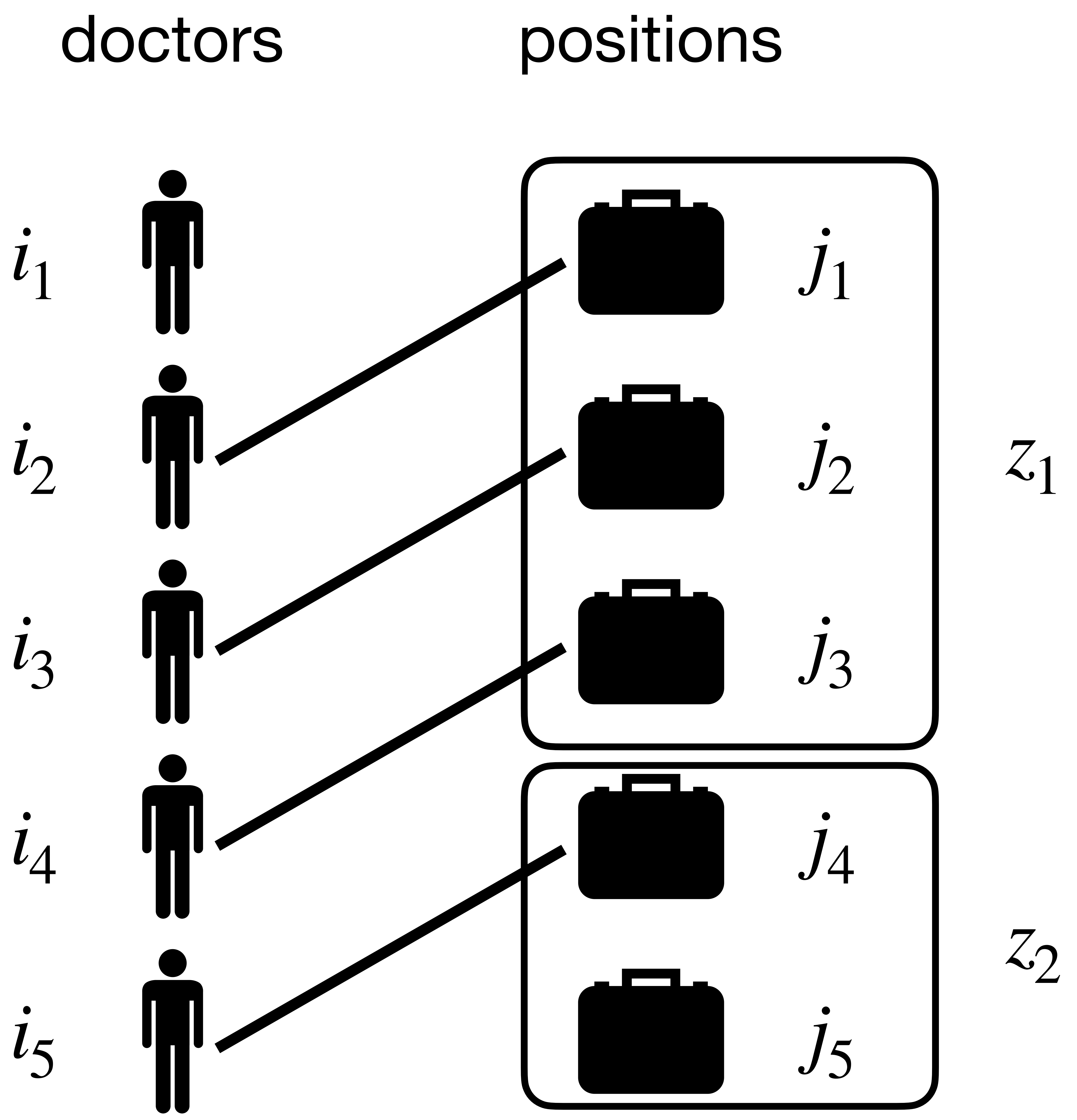
# of matches in z_2 ≥ 2

taxation policy

- alternative: taxation policy
  - a tax $w_z \in \mathbb{R}$ is applied uniformly to all matches in region z

negative tax = subsidy

  - Given $(w_z)_z$, agents form a stable outcome as if the adjusted joint surplus $\Phi_{ij} - w_z$ were the joint surplus



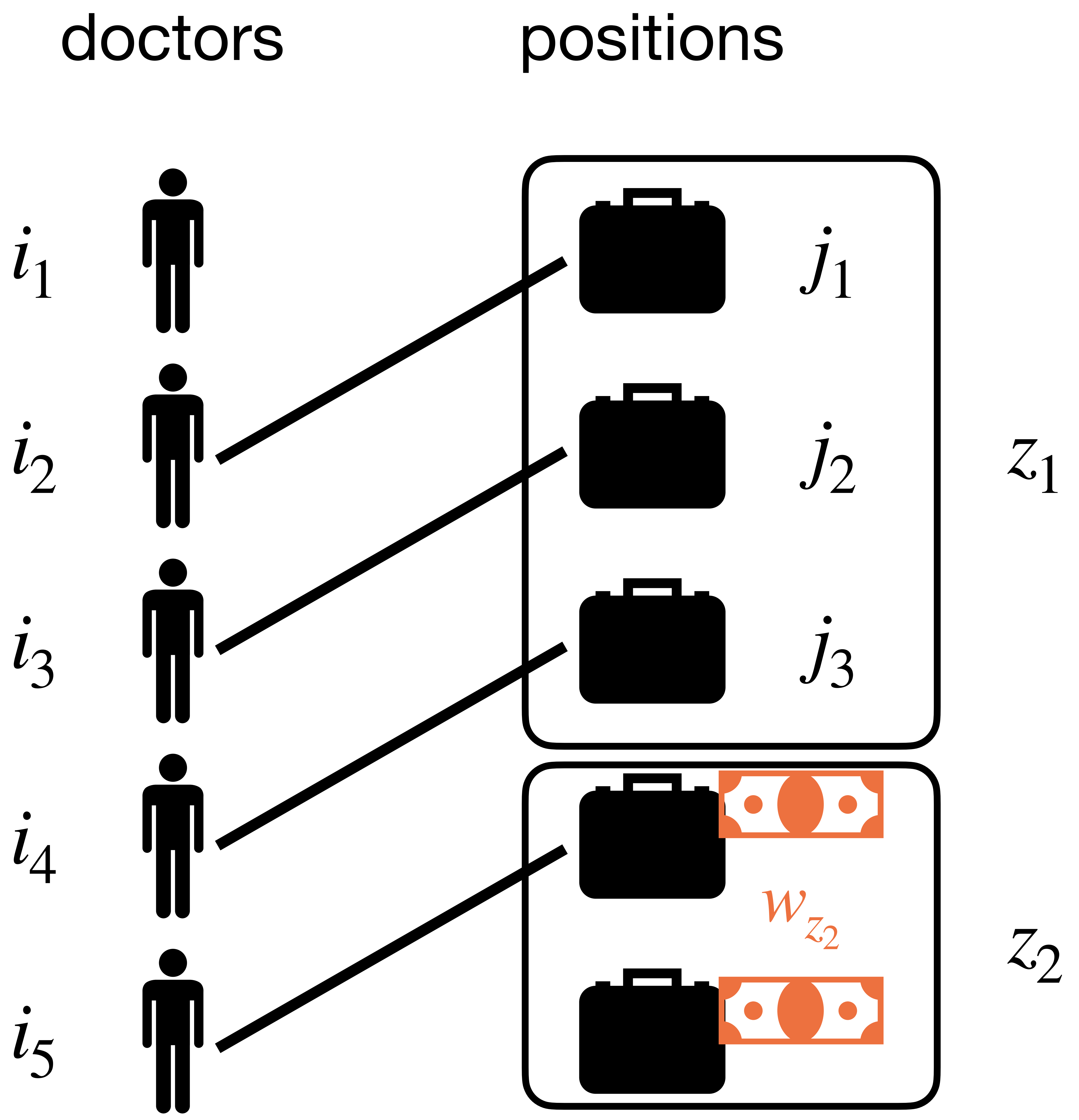
# of matches in z_2 ≥ 2

taxation policy

- alternative: taxation policy
  - a tax $w_z \in \mathbb{R}$ is applied uniformly to all matches in region z

negative tax = subsidy

  - Given $(w_z)_z$, agents form a stable outcome as if the adjusted joint surplus $\Phi_{ij} - w_z$ were the joint surplus



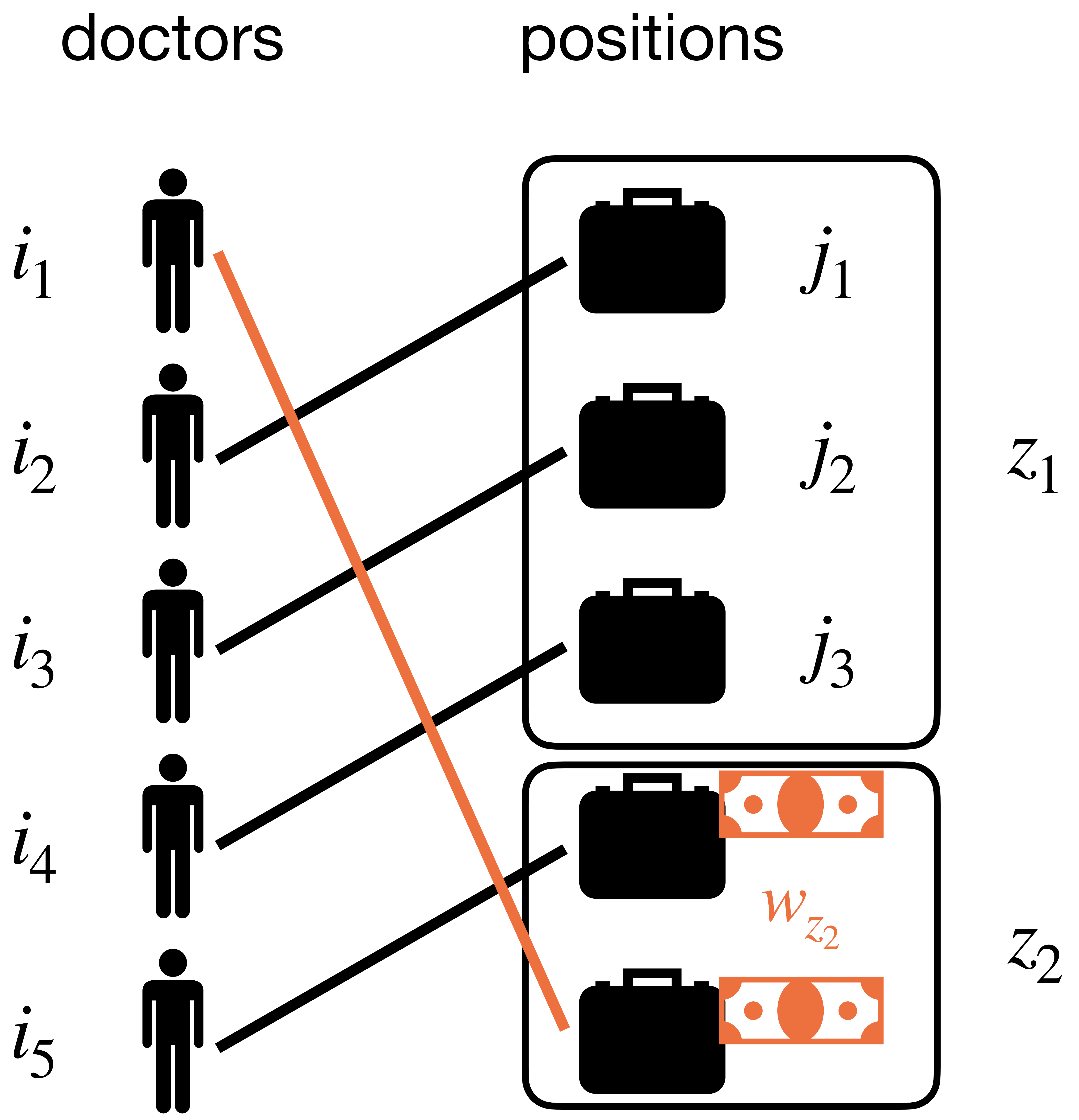
# of matches in z_2 ≥ 2

taxation policy

- alternative: taxation policy
  - a tax $w_z \in \mathbb{R}$ is applied uniformly to all matches in region z

negative tax = subsidy

  - Given $(w_z)_z$, agents form a stable outcome as if the adjusted joint surplus $\Phi_{ij} - w_z$ were the joint surplus



\# of matches in z_2 ≥ 2

policymaker's goal

design a policy that achieves high social surplus while satisfying regional constraints

policymaker's goal

design a policy that achieves high social surplus while satisfying regional constraints

social surplus := doctor surplus + position surplus - government expenditure

= the sum of the generated (non-adjusted) joint surplus $\sum_{i,j} \Phi_{ij} d_{ij}$

model

theoretical results

estimation

simulation

outline

1. **ideal world**: PM perfectly knows agents' preferences
   PM can compute the optimal taxation policy w^*.

2. **reality**: PM cannot compute w^*.
   alternative taxation policy w^A under more realistic assumptions on PM's knowledge

3. **large market approximation**:
   w^* and w^A are approximately the same in large markets

warm-up: property of stable outcomes

Fact (Shapley and Shubik, 1972): (d, u, v) is a stable outcome if and only if d is an integral solution to (P), and (u, v) is a solution to (D).

primal (P)

$$\left[\begin{array}{lll} \max\limits_{d \geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \end{array}\right.$$

dual (D)

$$\left[\begin{array}{lll} \min\limits_{u,v \geq 0} & \sum\limits_{i} u_i + \sum\limits_{j} v_j & \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} & (i \in I, j \in J) \end{array}\right.$$

warm-up: property of stable outcomes

Fact (Shapley and Shubik, 1972): (d, u, v) is a stable outcome if and only if d is an integral solution to (P), and (u, v) is a solution to (D).

primal (P)

$$\max_{d \geq 0} \quad \sum_{i,j} \Phi_{ij} d_{ij}$$

there exists an integral optimal solution d^*

$$\sum_i d_{ij} \leq 1 \quad (j \in J)$$

dual (D)

$$\left[\begin{array}{ll} \min\limits_{u,v \geq 0} & \sum\limits_i u_i + \sum\limits_j v_j \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} \quad (i \in I, j \in J) \end{array}\right.$$

warm-up: property of stable outcomes

Fact (Shapley and Shubik, 1972): (d, u, v) is a stable outcome if and only if d is an integral solution to (P), and (u, v) is a solution to (D).

primal (P)

$$\left[\begin{array}{lll} \max\limits_{d\geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \end{array}\right.$$

social surplus

dual (D)

$$\left[\begin{array}{lll} \min\limits_{u,v\geq 0} & \sum\limits_{i} u_i + \sum\limits_{j} v_j & \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} & (i \in I, j \in J) \end{array}\right.$$

warm-up: property of stable outcomes

Fact (Shapley and Shubik, 1972): (d, u, v) is a stable outcome if and only if d is an integral solution to (P), and (u, v) is a solution to (D).

primal (P)

$$\begin{aligned} \max_{d \geq 0} \quad & \sum_{i,j} \Phi_{ij} d_{ij} \\ \text{s.t.} \quad & \textstyle\sum_j d_{ij} \leq 1 \quad (i \in I) \\ & \textstyle\sum_i d_{ij} \leq 1 \quad (j \in J) \end{aligned}$$

social surplus

feasibility constraints

dual (D)

$$\begin{aligned} \min_{u,v \geq 0} \quad & \sum_i u_i + \sum_j v_j \\ \text{s.t.} \quad & u_i + v_j \geq \Phi_{ij} \quad (i \in I, j \in J) \end{aligned}$$

warm-up: property of stable outcomes

Fact (Shapley and Shubik, 1972): (d, u, v) is a stable outcome if and only if d is an integral solution to (P), and (u, v) is a solution to (D).

primal (P)

$$\left[\begin{array}{lll} \max\limits_{d \geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \text{social surplus} \\ \text{s.t.} & \sum_j d_{ij} \leq 1 \quad (i \in I) & \\ & \sum_i d_{ij} \leq 1 \quad (j \in J) & \text{feasibility constraints} \end{array}\right.$$

dual (D)

$$\left[\begin{array}{ll} \min\limits_{u,v \geq 0} & \sum\limits_{i} u_i + \sum\limits_{j} v_j \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} \quad (i \in I, j \in J) \end{array}\right.$$

Corollary: matching d under a stable outcome maximizes social surplus among all feasible matchings

optimal taxation policy: benchmark case

Thm 1: Fix any regional constraints.

If PM knows (Φ_{ij}), she can compute a taxation policy $w^* = (w_z^*)_z$ such that

- w^* implements a matching d^*
- d^* maximizes social surplus among all matchings that satisfy the regional constraints

optimal taxation policy: benchmark case

Thm 1: Fix any regional constraints.

If PM knows (Φ_{ij}), she can compute a taxation policy $w^* = (w_z^*)_z$ such that

- w^* implements a matching d^*
- d^* maximizes social surplus among all matchings that satisfy the regional constraints

d^* is part of a stable outcome (d^*, u^*, v^*) under w^*

optimal taxation policy: benchmark case

Thm 1: Fix any regional constraints.

If PM knows (Φ_{ij}), she can compute a taxation policy $w^* = (w_z^*)_z$ such that

- w^* implements a matching d^*
- d^* maximizes social surplus among all matchings that satisfy the regional constraints

- w^* is the optimal taxation policy

proof sketch

$\Phi_{i,0} = \Phi_{0,j} = 0$

proof sketch

$$\Phi_{i,0} = \Phi_{0,j} = 0$$

$$\left[\begin{array}{lll} \max\limits_{d \geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \\ & \sum_{j \in z} \sum_i d_{ij} \leq \bar{o}_z & (z \in Z) \\ & \sum_{j \in z} \sum_i d_{ij} \geq \underline{o}_z & (z \in Z) \end{array}\right.$$

proof sketch

$$\Phi_{i,0} = \Phi_{0,j} = 0$$

$$\left[\begin{array}{lll} \max\limits_{d \geq 0} & \boxed{\sum\limits_{i,j} \Phi_{ij} d_{ij}} & \text{social surplus} \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \\ & \sum_{j \in z} \sum_i d_{ij} \leq \bar{o}_z & (z \in Z) \\ & \sum_{j \in z} \sum_i d_{ij} \geq \underline{o}_z & (z \in Z) \end{array}\right.$$

proof sketch

$$\Phi_{i,0} = \Phi_{0,j} = 0$$

$$\left[\begin{array}{lll} \max\limits_{d\geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \\ & \sum_{j\in z} \sum_i d_{ij} \leq \bar{o}_z & (z \in Z) \\ & \sum_{j\in z} \sum_i d_{ij} \geq \underline{o}_z & (z \in Z) \end{array}\right.$$

social surplus

regional constraints

proof sketch

$$\left[\begin{array}{lll} \max\limits_{d\geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \\ & \sum_{j\in z} \sum_i d_{ij} \leq \overline{o}_z & (z \in Z) \\ & \sum_{j\in z} \sum_i d_{ij} \geq \underline{o}_z & (z \in Z) \end{array}\right.$$

proof sketch

$$\left[\begin{array}{lll}
\max\limits_{d \geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \\
\text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\
& \sum_i d_{ij} \leq 1 & (j \in J) \\
& \sum_{j \in z} \sum_i d_{ij} \leq \bar{o}_z & (z \in Z) \\
& \sum_{j \in z} \sum_i d_{ij} \geq \underline{o}_z & (z \in Z)
\end{array}\right.$$

$$\left[\begin{array}{lll}
\min\limits_{u,v,\bar{w},\underline{w} \geq 0} & \sum\limits_{i} u_i + \sum\limits_{j} v_j + \sum\limits_{z \in Z} \bar{o}_z \bar{w}_z - \sum\limits_{z \in Z} \underline{o}_z \underline{w}_z & \\
\text{s.t.} & u_i + v_j \geq \Phi_{ij} - \bar{w}_{z(j)} + \underline{w}_{z(j)} & (i \in I, j \in J)
\end{array}\right.$$

proof sketch

$$\left[\begin{array}{lll} \max\limits_{d\geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \\ & \sum_{j\in z} \sum_i d_{ij} \leq \bar{o}_z & (z \in Z) \\ & \sum_{j\in z} \sum_i d_{ij} \geq \underline{o}_z & (z \in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{u,v,\bar{w},\underline{w}\geq 0} & \sum\limits_i u_i + \sum\limits_j v_j + \sum\limits_{z\in Z} \bar{o}_z \bar{w}_z - \sum\limits_{z\in Z} \underline{o}_z \underline{w}_z & \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} - \bar{w}_{z(j)} + \underline{w}_{z(j)} & (i \in I, j \in J) \end{array}\right.$$

- solutions d^* and $(u^*, v^*, \bar{w}, \underline{w})$

proof sketch

$$\left[\begin{array}{lll} \max\limits_{d\geq 0} & \sum\limits_{i,j}\Phi_{ij}d_{ij} & \\ \text{s.t.} & \sum_j d_{ij}\leq 1 & (i\in I) \\ & \sum_i d_{ij}\leq 1 & (j\in J) \\ & \sum_{j\in z}\sum_i d_{ij}\leq \bar{o}_z & (z\in Z) \\ & \sum_{j\in z}\sum_i d_{ij}\geq \underline{o}_z & (z\in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{u,v,\bar{w},\underline{w}\geq 0} & \sum\limits_i u_i+\sum\limits_j v_j+\sum\limits_{z\in Z}\bar{o}_z\bar{w}_z-\sum\limits_{z\in Z}\underline{o}_z\underline{w}_z & \\ \text{s.t.} & u_i+v_j\geq \Phi_{ij}-\bar{w}_{z(j)}+\underline{w}_{z(j)} & (i\in I, j\in J) \end{array}\right.$$

- solutions $d*$ and $(u*, v*, \bar{w}, \underline{w})$

there exists an integral optimal solution $d*$

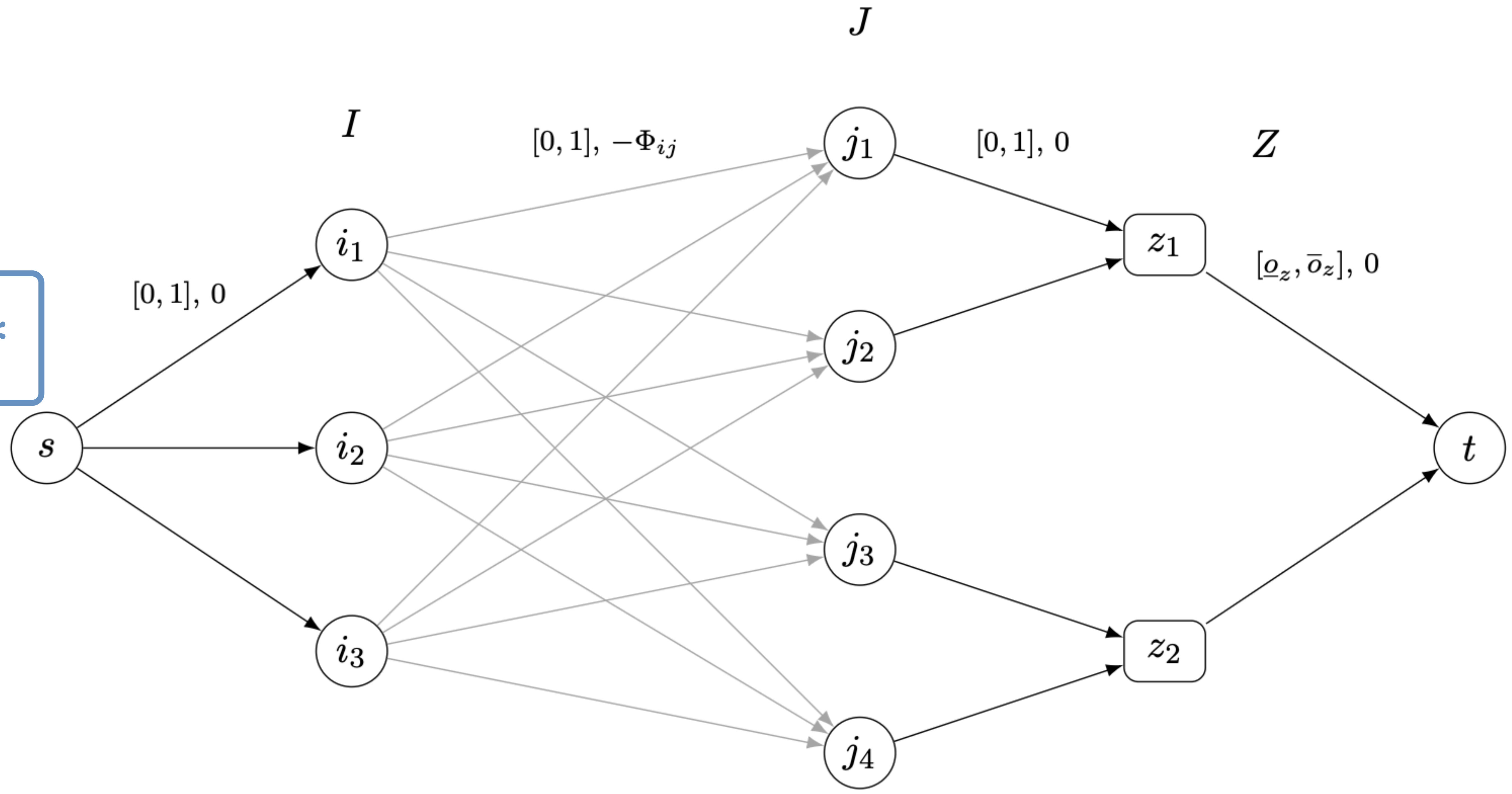
proof sketch

$$\left[\begin{array}{lll} \max\limits_{d\geq 0} & \sum\limits_{i,j}\Phi_{ij}d_{ij} & \\ \text{s.t.} & \sum_j d_{ij}\leq 1 & (i\in I) \\ & \sum_i d_{ij}\leq 1 & (j\in J) \\ & \sum_{j\in z}\sum_i d_{ij}\leq \bar{o}_z & (z\in Z) \\ & \sum_{j\in z}\sum_i d_{ij}\geq \underline{o}_z & (z\in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{u,v,\bar{w},\underline{w}\geq 0} & \sum\limits_i u_i+\sum\limits_j v_j+\sum\limits_{z\in Z}\bar{o}_z\bar{w}_z-\sum\limits_{z\in Z}\underline{o}_z\underline{w}_z & \\ \text{s.t.} & u_i+v_j\geq \Phi_{ij}-\bar{w}_{z(j)}+\underline{w}_{z(j)} & (i\in I, j\in J) \end{array}\right.$$

- solutions d^* and $(u^*, v^*, \bar{w}, \underline{w})$

there exists an integral optimal solution d^*



proof sketch

$$\left[\begin{array}{lll} \max\limits_{d\geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \\ & \sum_{j\in z} \sum_i d_{ij} \leq \bar{o}_z & (z \in Z) \\ & \sum_{j\in z} \sum_i d_{ij} \geq \underline{o}_z & (z \in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{u,v,\bar{w},\underline{w}\geq 0} & \sum\limits_{i} u_i + \sum\limits_{j} v_j + \sum\limits_{z\in Z} \bar{o}_z \bar{w}_z - \sum\limits_{z\in Z} \underline{o}_z \underline{w}_z & \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} - \bar{w}_{z(j)} + \underline{w}_{z(j)} & (i \in I, j \in J) \end{array}\right.$$

- solutions d^* and $(u^*, v^*, \bar{w}, \underline{w})$
  - $\bar{w}_z, \underline{w}_z \geq 0$: Lagrange multipliers for the regional constraints
  - by complementary slackness, at most one of them is strictly positive

proof sketch

$$\left[\begin{array}{lll} \max\limits_{d\geq 0} & \sum\limits_{i,j}\Phi_{ij}d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \\ & \sum_{j\in z}\sum_i d_{ij} \leq \bar{o}_z & (z \in Z) \\ & \sum_{j\in z}\sum_i d_{ij} \geq \underline{o}_z & (z \in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{u,v,\bar{w},\underline{w}\geq 0} & \sum\limits_i u_i + \sum\limits_j v_j + \sum\limits_{z\in Z}\bar{o}_z\bar{w}_z - \sum\limits_{z\in Z}\underline{o}_z\underline{w}_z & \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} - \bar{w}_{z(j)} + \underline{w}_{z(j)} & (i \in I, j \in J) \end{array}\right.$$

$$\Phi_{ij} - w^*_{z(j)}$$

- solutions d^* and $(u^*, v^*, \bar{w}^*, \underline{w}^*)$
- define the taxation policy $w^*_z := \bar{w}^*_z - \underline{w}^*_z$
    - $w^*_z = \bar{w}^*_z > 0$ if the cap binds (tax)
    - $w^*_z = -\underline{w}^*_z < 0$ if the floor binds (subsidy)

proof sketch

- WTS: (d^*, u^*, v^*) forms a stable outcome under taxation policy w^*

proof sketch

- WTS: (d^*, u^*, v^*) forms a stable outcome under w^*
- a stable outcome under w^* is characterized by solutions to the following LPs:

adjusted joint surplus

$$\left[\begin{array}{lll} \max\limits_{d \geq 0} & \sum\limits_{i,j} (\Phi_{ij} - w^*_{z(j)}) d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \end{array}\right. \qquad \left[\begin{array}{lll} \min\limits_{u,v \geq 0} & \sum\limits_{i \in I} u_i + \sum\limits_{j \in J} v_j & \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} - w^*_{z(j)} & (i \in I, j \in J) \end{array}\right.$$

proof sketch

- WTS: (d^*, u^*, v^*) forms a stable outcome under w^*
- a stable outcome under w^* is characterized by solutions to the following LPs:

adjusted joint surplus

$$\left[\begin{array}{lll} \max\limits_{d \geq 0} & \sum\limits_{i,j} (\Phi_{ij} - w^*_{z(j)}) d_{ij} & \\ \text{s.t.} & \quad \sum_j d_{ij} \leq 1 & (i \in I) \\ & \quad \sum_i d_{ij} \leq 1 & (j \in J) \end{array}\right. \qquad \left[\begin{array}{lll} \min\limits_{u,v \geq 0} & \sum\limits_{i \in I} u_i + \sum\limits_{j \in J} v_j & \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} - w^*_{z(j)} & (i \in I, j \in J) \end{array}\right.$$

- (d^*, u^*, v^*) satisfies the optimality conditions for these problems
  - they are part of the optimality conditions for the previous LPs

proof sketch

- WTS: (d^*, u^*, v^*) forms a stable outcome under w^*
- a stable outcome under w^* is characterized by solutions to the following LPs:

adjusted joint surplus

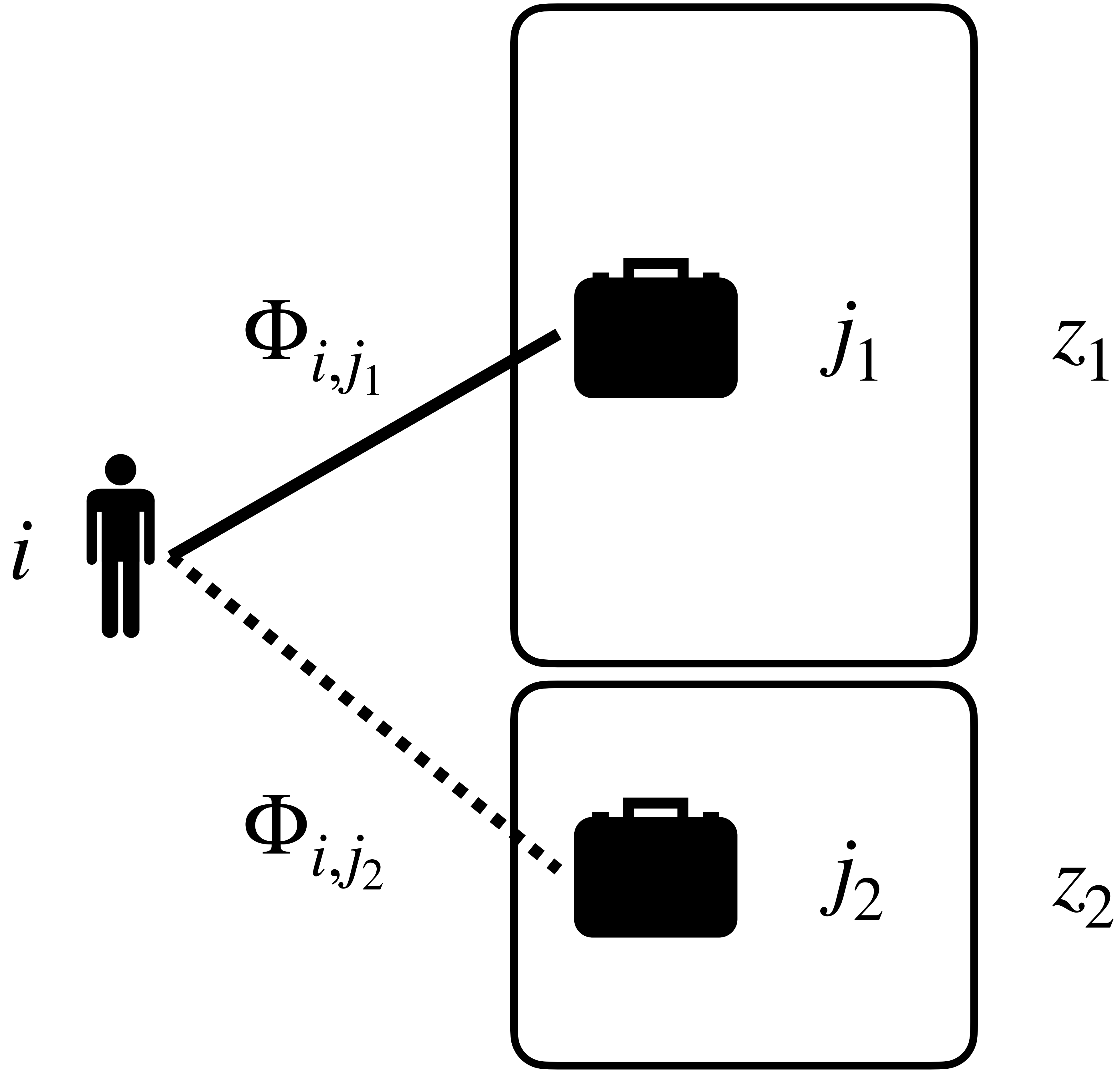
$$\left[\begin{array}{lll} \max\limits_{d\geq 0} & \sum\limits_{i,j}(\Phi_{ij} - w^*_{z(j)})d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{u,v\geq 0} & \sum\limits_{i\in I} u_i + \sum\limits_{j\in J} v_j & \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} - w^*_{z(j)} & (i \in I, j \in J) \end{array}\right.$$

- (d^*, u^*, v^*) satisfies the optimality conditions for these problems
  - they are part of the optimality conditions for the previous LPs
- taxation policy w^* implements the planner's optimal solution d^*

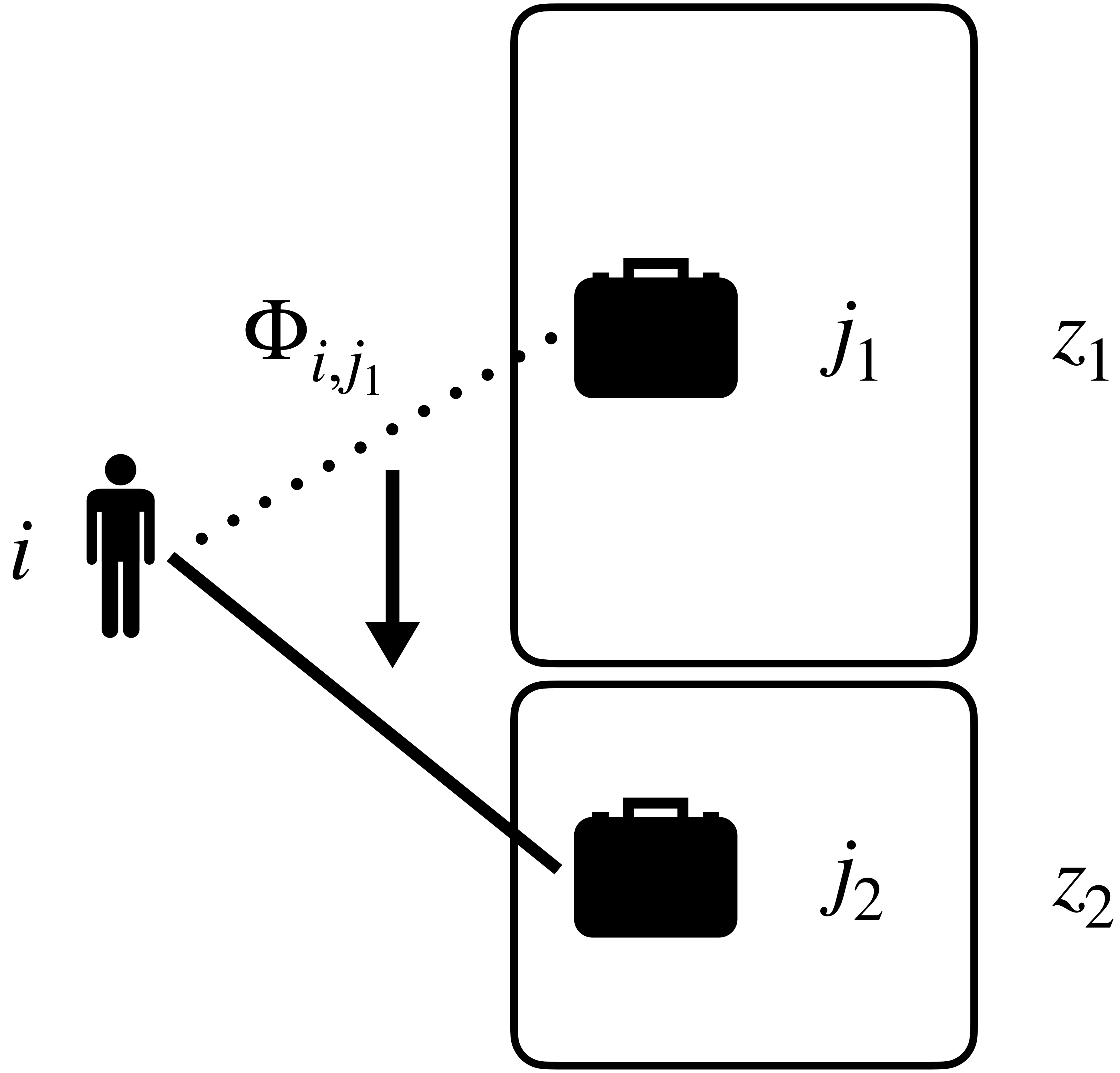
properties of optimal taxation policy

- w_z^*: shadow prices of the regional constraints



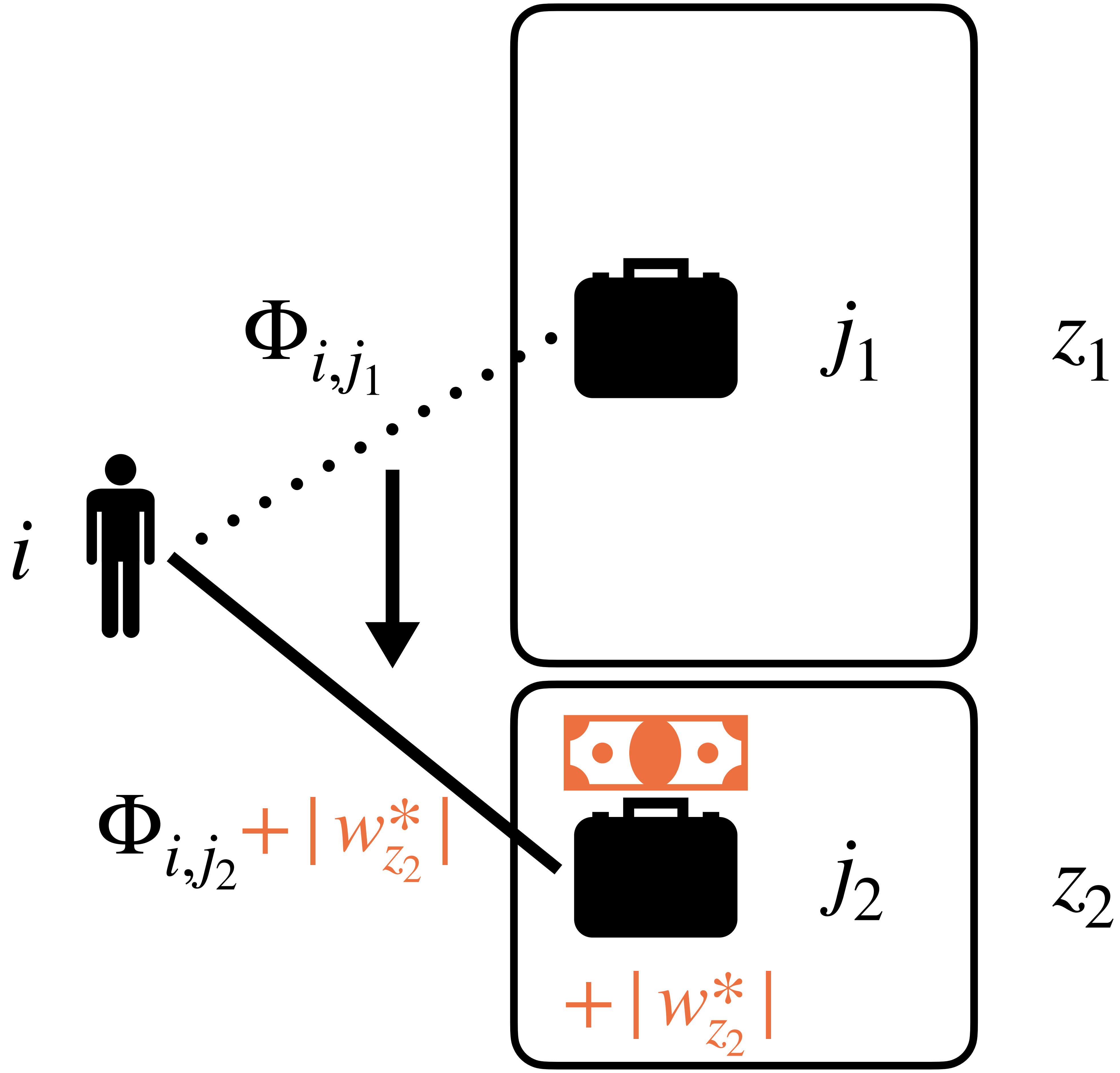
properties of optimal taxation policy

- w_z^*: shadow prices of the regional constraints
  - suppose that $w_{z_2}^* < 0$ (i.e., subsidize $|w_{z_2^*}|$)
  - recruit an extra doctor for region z_2



properties of optimal taxation policy

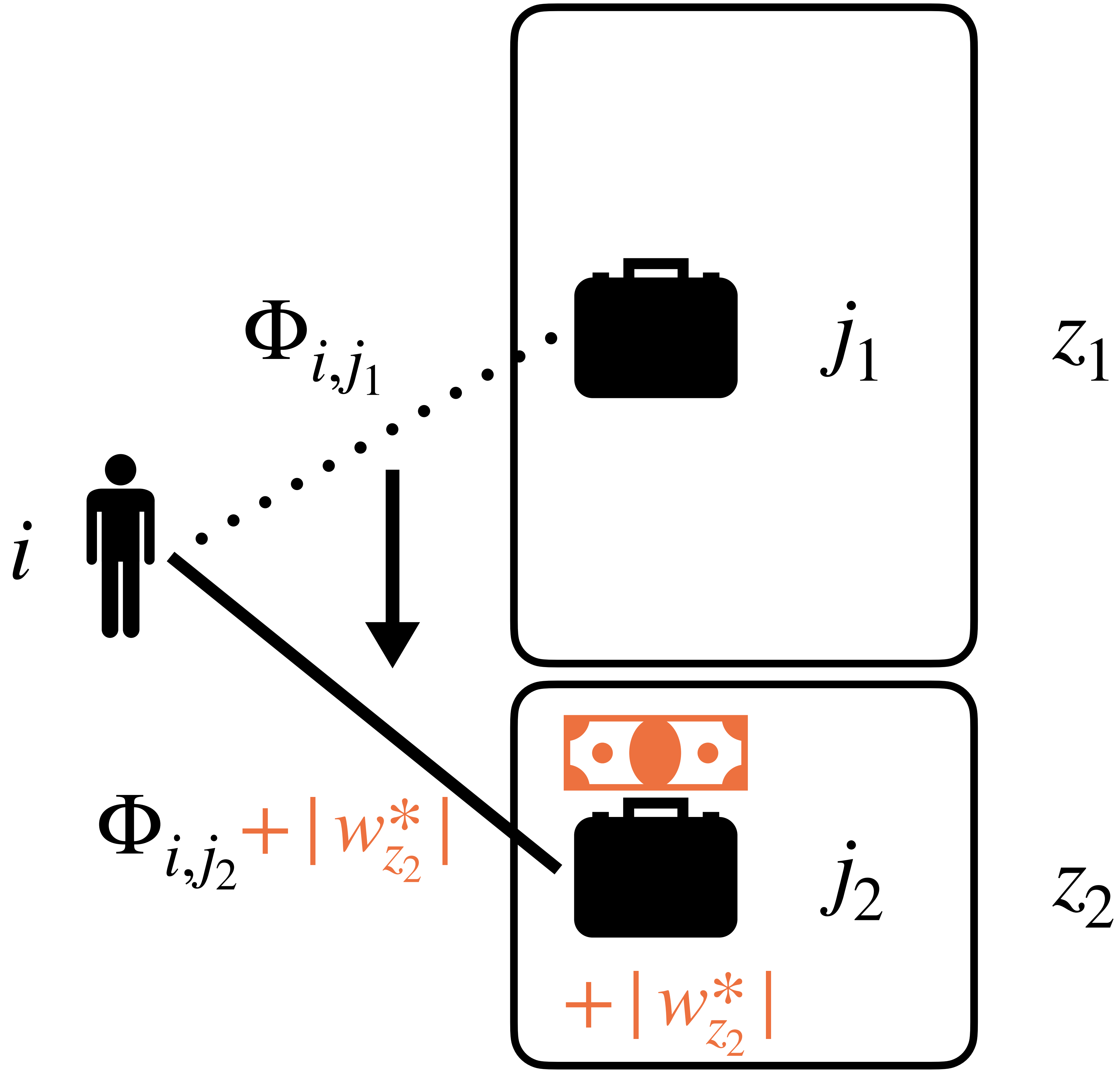
- w_z^*: shadow prices of the regional constraints
  - suppose that $w_{z_2}^* < 0$ (i.e., subsidize $|w_{z_2^*}|$)
  - recruit an extra doctor for region z_2



properties of optimal taxation policy

- w_z^*: shadow prices of the regional constraints
  - suppose that $w_{z_2}^* < 0$ (i.e., subsidize $|w_{z_2^*}|$)
  - recruit an extra doctor for region z_2

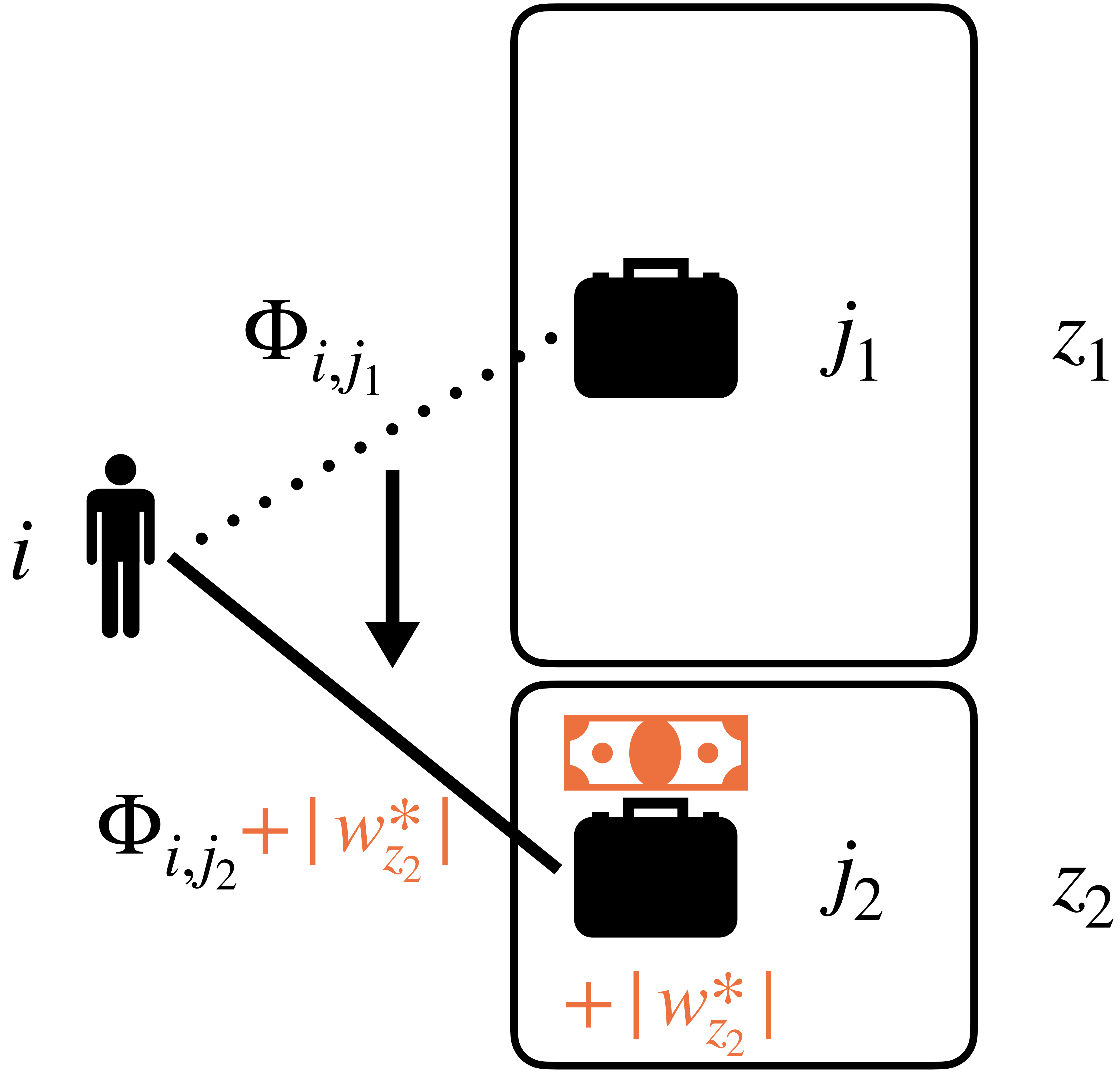
$$\Phi_{i,j_2} + |w_{z_2}^*| = \Phi_{i,j_1}$$



properties of optimal taxation policy

- w_z^*: shadow prices of the regional constraints
  - suppose that $w_{z_2}^* < 0$ (i.e., subsidize $|w_{z_2^*}|$)
  - recruit an extra doctor for region z_2

$$\Phi_{i,j_2} = \Phi_{i,j_1} - |w_{z_2}^*|$$

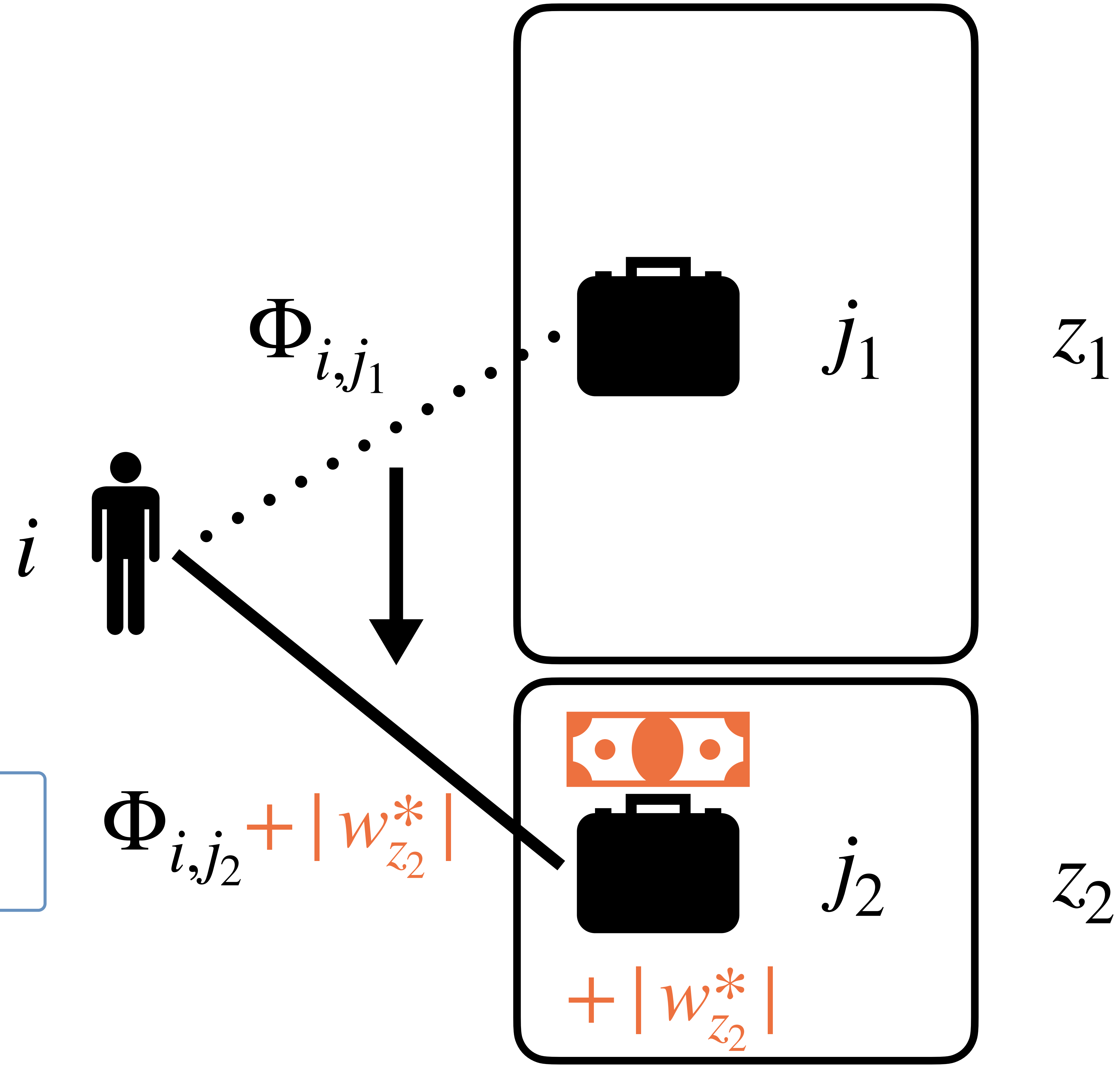


properties of optimal taxation policy

- w^*_z: shadow prices of the regional constraints
  - suppose that $w^*_{z_2} < 0$ (i.e., subsidize $|w_{z_2^*}|$)
  - recruit an extra doctor for region z_2

$$\Phi_{i,j_2} = \Phi_{i,j_1} - |w^*_{z_2}|$$

generated (non-adjusted) joint surplus is reduced by $|w^*_{z_2}|$

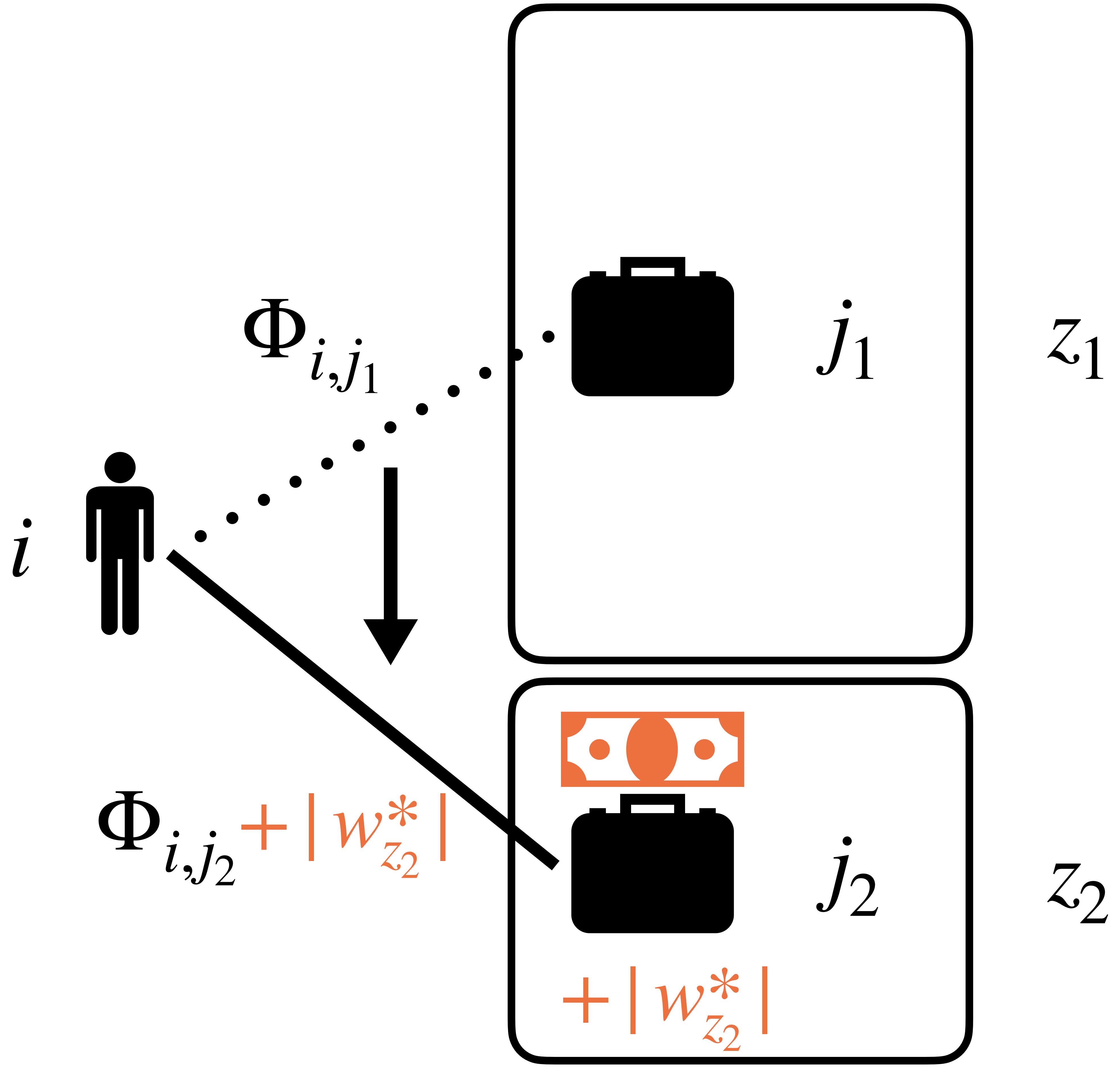


properties of optimal taxation policy

- w^*_z: shadow prices of the regional constraints
  - suppose that $w^*_{z_2} < 0$ (i.e., subsidize $|w_{z_2^*}|$)
  - recruit an extra doctor for region z_2

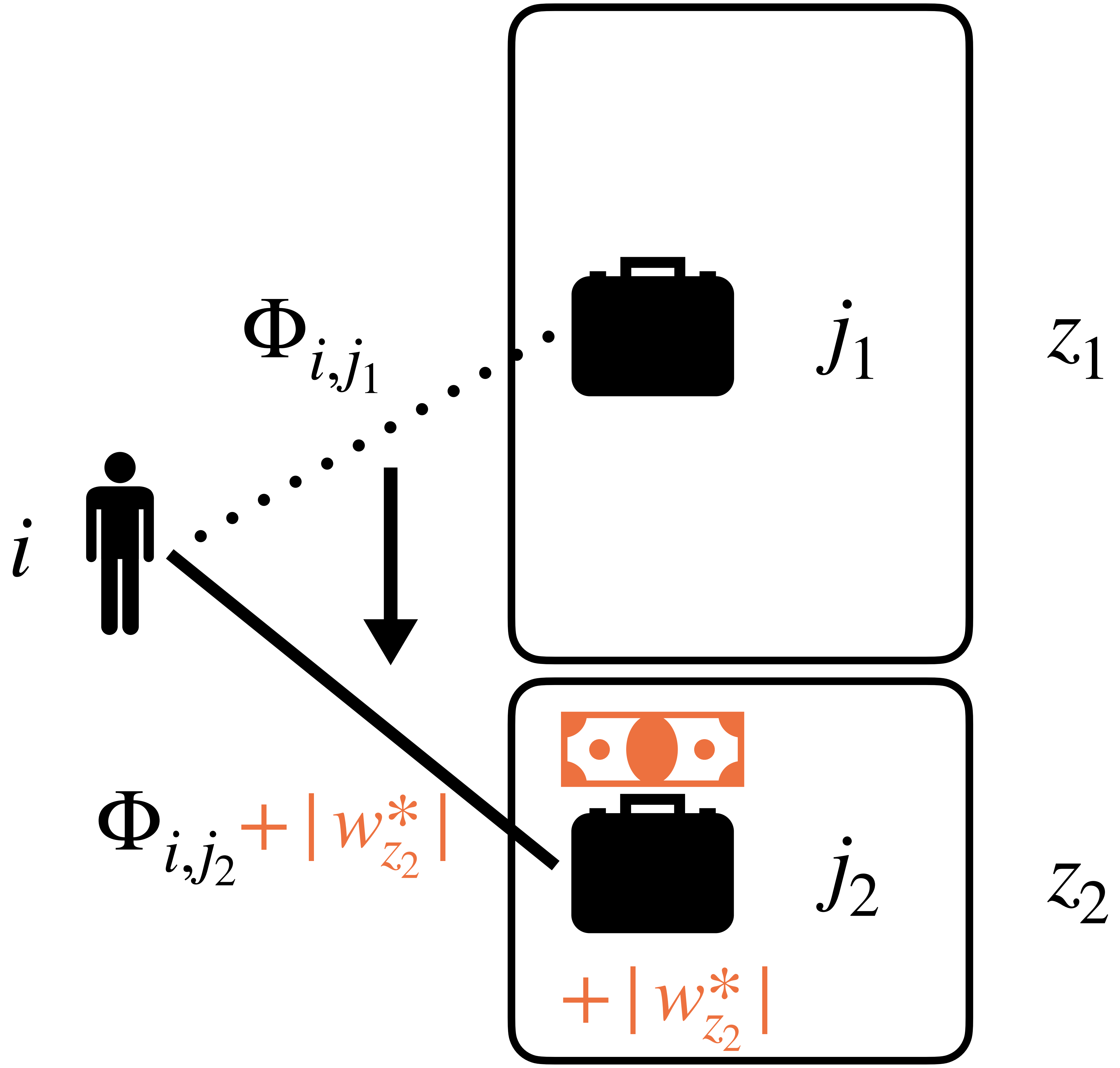
$$|w^*_{z_2}| = \Phi_{i,j_1} - \Phi_{i,j_2}$$

amount of subsidy = difference in joint surplus for the marginal doctor



properties of optimal taxation policy

- w_z^*: shadow prices of the regional constraints
  - suppose that $w_{z_2}^* < 0$ (i.e., subsidize $|w_{z_2^*}|$)
  - recruit an extra doctor for region z_2
- assuming a uniform tax within the same region is WLOG
  - having $w_{ij}^* \neq w_{ij'}^*$ for $j, j' \in z$ cannot increase social surplus
  - useful for practical implementation



PM's challenge: unobserved heterogeneity

- **problem:** the exact joint surplus Φ_{ij} is difficult to identify
    - we cannot directly apply the previous result

PM's challenge: unobserved heterogeneity

- **problem:** the exact joint surplus Φ_{ij} is difficult to identify
  - we cannot directly apply the previous result

- PM often faces unobserved heterogeneity
  - PM observes a finite set of agent types (observable characteristics)
  - agents with the same type are indistinguishable to PM

PM's challenge: unobserved heterogeneity

- **problem:** the exact joint surplus Φ_{ij} is difficult to identify
  - we cannot directly apply the previous result

- PM often faces unobserved heterogeneity
  - PM observes a finite set of agent types (observable characteristics)
  - agents with the same type are indistinguishable to PM
  - each doctor has type $s \in S$ (school)
  - each position has type $h \in H$ (hospital)

PM's challenge: unobserved heterogeneity

- **problem:** the exact joint surplus Φ_{ij} is difficult to identify
  - we cannot directly apply the previous result

- PM often faces unobserved heterogeneity
  - PM observes a finite set of agent types (observable characteristics)
  - agents with the same type are indistinguishable to PM
  - each doctor has type $s \in S$ (school)
  - each position has type $h \in H$ (hospital)

- can we still implement the optimal taxation policy? under what conditions?

(Galichon and Salanié, 2021)

key structural assumption

- **additive separability**: joint surplus Φ_{ij} is generated as follows: there exists a matrix $(\Phi_{sh})_{s,h}$ such that for all $i \in s, j \in h$,

$$\Phi_{ij} = \Phi_{sh} + \varepsilon_{ih} + \eta_{sj}$$

(Galichon and Salanié, 2021)

key structural assumption

- **additive separability**: joint surplus Φ_{ij} is generated as follows: there exists a matrix $(\Phi_{sh})_{s,h}$ such that for all $i \in s, j \in h$,

$$\Phi_{ij} = \boxed{\Phi_{sh}} + \varepsilon_{ih} + \eta_{sj}$$

aggregate-level
joint surplus

key structural assumption

(Galichon and Salanié, 2021)

- **additive separability**: joint surplus Φ_{ij} is generated as follows: there exists a matrix $(\Phi_{sh})_{s,h}$ such that for all $i \in s, j \in h$,

$$\Phi_{ij} = \underbrace{\boxed{\Phi_{sh}}}_{\text{aggregate-level joint surplus}} + \boxed{\varepsilon_{ih} + \eta_{sj}} \quad \text{error terms}$$

(Galichon and Salanié, 2021)

key structural assumption

- **additive separability**: joint surplus Φ_{ij} is generated as follows: there exists a matrix $(\Phi_{sh})_{s,h}$ such that for all $i \in s, j \in h$,

$$\Phi_{ij} = \boxed{\Phi_{sh}} + \boxed{\varepsilon_{ih} + \eta_{sj}}$$

error terms

aggregate-level
joint surplus

- for each $i \in s$, errors $\varepsilon_i = (\varepsilon_{ih})_h \in \mathbb{R}^{|H|+1}$ is drawn from P_s
- for each $j \in h$, errors $\eta_j = (\eta_{sj})_s \in \mathbb{R}^{|S|+1}$ is drawn from Q_h

key structural assumption

(Galichon and Salanié, 2021)

- **additive separability**: joint surplus Φ_{ij} is generated as follows: there exists a matrix $(\Phi_{sh})_{s,h}$ such that for all $i \in s, j \in h$,

$$\Phi_{ij} = \underbrace{\Phi_{sh}}_{\text{aggregate-level joint surplus}} + \overbrace{\varepsilon_{ih} + \eta_{sj}}^{\text{error terms}}$$

- for each $i \in s$, errors $\varepsilon_i = (\varepsilon_{ih})_h \in \mathbb{R}^{|H|+1}$ is drawn from P_s
- for each $j \in h$, errors $\eta_j = (\eta_{sj})_s \in \mathbb{R}^{|S|+1}$ is drawn from Q_h
- $P := (P_s)_{s \in S}$, $Q := (Q_h)_{h \in H}$

(Galichon and Salanié, 2021)

key structural assumption

- **additive separability**: joint surplus Φ_{ij} is generated as follows: there exists a matrix $(\Phi_{sh})_{s,h}$ such that for all $i \in s, j \in h$,

$$\Phi_{ij} = \Phi_{sh} + \varepsilon_{ih} + \eta_{sj}$$

error terms

aggregate-level joint surplus

  - for each $i \in s$, errors $\varepsilon_i = (\varepsilon_{ih})_h \in \mathbb{R}^{|H|+1}$ is drawn from P_s
  - for each $j \in h$, errors $\eta_j = (\eta_{sj})_s \in \mathbb{R}^{|S|+1}$ is drawn from Q_h
  - $P := (P_s)_{s \in S}$, $Q := (Q_h)_{h \in H}$
  - error vectors $((\varepsilon_i)_i, (\eta_j)_j)$ are independent

assumptions on PM’s knowledge

$$\Phi_{ij} = \boxed{\Phi_{sh}} + \boxed{\varepsilon_{ih} + \eta_{sj}}$$

aggregate-level
joint surplus

error terms

assumptions on PM's knowledge

$$\Phi_{ij} = \Phi_{sh} + \varepsilon_{ih} + \eta_{sj}$$

aggregate-level
joint surplus

error terms

- we assume that PM knows $(\Phi_{sh})_{s,h}$, P, and Q

assumptions on PM's knowledge

$$\Phi_{ij} = \boxed{\Phi_{sh}} + \boxed{\varepsilon_{ih} + \eta_{sj}}$$

aggregate-level joint surplus

error terms

- we assume that PM knows $(\Phi_{sh})_{s,h}$, P, and Q
- GS2021: $(\Phi_{sh})_{s,h}$ can be estimated from aggregate-level match data $\mu = (\mu_{sh})_{s,h}$
  - μ_{sh} denotes # of matches between school s and hospital h
  - P and Q are chosen by PM (e.g., Gumbel distribution)

assumptions on PM's knowledge

$$\Phi_{ij} = \boxed{\Phi_{sh}} + \boxed{\varepsilon_{ih} + \eta_{sj}}$$

aggregate-level
joint surplus

error terms

- we assume that PM knows $(\Phi_{sh})_{s,h}$, P, and Q
- GS2021: $(\Phi_{sh})_{s,h}$ can be estimated from aggregate-level match data $\mu = (\mu_{sh})_{s,h}$
  - μ_{sh} denotes # of matches between school s and hospital h
  - P and Q are chosen by PM (e.g., Gumbel distribution)
- **next goal:** design a taxation policy computable with this knowledge

discrete choice representation

(Galichon and Salanié, 2021)

Lemma (GS2021)

- suppose that (u, v) is an equilibrium payoff profile under a stable outcome
- let

$$U_{sh} := \min_{i \in s} \{u_i - \varepsilon_{ih}\}$$

- then, under additive separability, for any $i \in s$,

$$u_i = \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

(Galichon and Salanié, 2021)

discrete choice representation

Lemma (GS2021)

- suppose that (u, v) is an equilibrium payoff profile under a stable outcome
- let

$$U_{sh} := \min_{i \in s}\{u_i - \varepsilon_{ih}\}$$

- then, under additive separability, for any $i \in s$,

$$u_i = \max_{h \in H \cup \{0\}} \{U_{sh} + \varepsilon_{ih}\}$$

(Galichon and Salanié, 2021)

discrete choice representation

Lemma (GS2021)

- suppose that (u, v) is an equilibrium payoff profile under a stable outcome
- let

$$U_{sh} :=$$

doctor i behaves as if he chooses
his favorite hospital from choice set H

- then, under additive separability, for any $i \in s$,

$$u_i = \max_{h \in H \cup \{0\}} \{U_{sh} + \varepsilon_{ih}\}$$

discrete choice representation

(Galichon and Salanié, 2021)

Lemma (GS2021)

- suppose that (u, v) is an equilibrium payoff profile under a stable outcome
- let

$$U_{sh} :=$$

doctor i behaves as if he chooses
his favorite hospital from choice set H

- then, under additive separability, for any $i \in s$,

$$u_i = \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

(analogous result for positions also holds)

toward large market approximation

$$u_i = \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

toward large market approximation

$$u_i = \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

- rewrite doctors' surplus:

toward large market approximation

$$u_i = \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

- rewrite doctors' surplus:

$$\sum_i u_i$$

toward large market approximation

$$u_i = \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

- rewrite doctors' surplus:

$$\sum_i u_i = \sum_{s \in S} \sum_{i \in s} u_i$$

toward large market approximation

$$u_i = \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

- rewrite doctors' surplus:

$$\sum_i u_i = \sum_{s \in S} \sum_{i \in s} u_i = \sum_{s \in S} \sum_{i \in s} \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

toward large market approximation

$$u_i = \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

- rewrite doctors' surplus:

$$\sum_i u_i = \sum_{s \in S} \sum_{i \in s} u_i = \sum_{s \in S} \sum_{i \in s} \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

$$= \sum_{s \in S} n_s \frac{1}{n_s} \sum_{i \in s} \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

toward large market approximation

$$u_i = \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

- rewrite doctors' surplus:

$$\sum_i u_i = \sum_{s \in S} \sum_{i \in s} u_i = \sum_{s \in S} \sum_{i \in s} \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

$$= \sum_{s \in S} n_s \frac{1}{n_s} \sum_{i \in s} \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

# of type-s doctors

toward large market approximation

$$u_i = \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

- rewrite doctors' surplus:

$$\sum_i u_i = \sum_{s \in S} \sum_{i \in s} u_i = \sum_{s \in S} \sum_{i \in s} \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

$$= \sum_{s \in S} n_s \frac{1}{n_s} \sum_{i \in s} \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

toward large market approximation

$$u_i = \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

- rewrite doctors' surplus:

$$\sum_i u_i = \sum_{s \in S} \sum_{i \in s} u_i = \sum_{s \in S} \sum_{i \in s} \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\}$$

$$= \sum_{s \in S} n_s \frac{1}{n_s} \sum_{i \in s} \max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\} \overset{?}{\approx} \underset{\varepsilon_i \sim P_s}{E} \left[\max_{h \in H \cup \{0\}} \left\{ U_{sh} + \varepsilon_{ih} \right\} \right]$$

toward large market approximation

$$u_i = \max_{h \in H \cup \{0\}} \{U_{sh} + \varepsilon_{ih}\}$$

- rewrite doctors' surplus:

$$\sum_i u_i = \sum_{s \in S} \sum_{i \in s} u_i = \sum_{s \in S} \sum_{i \in s} \max_{h \in H \cup \{0\}} \{U_{sh} + \varepsilon_{ih}\}$$

$$= \sum_{s \in S} n_s \frac{1}{n_s} \sum_{i \in s} \max_{h \in H \cup \{0\}} \{U_{sh} + \varepsilon_{ih}\} \overset{?}{\approx} \underset{\varepsilon_i \sim P_s}{E} \left[\max_{h \in H \cup \{0\}} \{U_{sh} + \varepsilon_{ih}\} \right]$$

$$G(U) := \sum_{s \in S} n_s \underset{\varepsilon_i \sim P_s}{E} \left[\max_{h \in H \cup \{0\}} \{U_{sh} + \varepsilon_{ih}\} \right]$$

taxation policy with unobs. heterogeneity

- to define a taxation policy, we construct an optimization problem that corresponds to the previous LP

taxation policy with unobs. heterogeneity

- to define a taxation policy, we construct an optimization problem that corresponds to the previous LP

primal

dual

ideal

(Φ_{ij}): given

$$\left[\begin{array}{lll} \max\limits_{d\geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \\ & \sum_{j\in z} \sum_i d_{ij} \leq \overline{o}_z & (z \in Z) \\ & \sum_{j\in z} \sum_i d_{ij} \geq \underline{o}_z & (z \in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{u,v,\overline{w},\underline{w}\geq 0} & \sum\limits_i u_i + \sum\limits_j v_j + \sum\limits_{z\in Z} \overline{o}_z \overline{w}_z - \sum\limits_{z\in Z_L} \underline{o}_z \underline{w}_z & \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} - \overline{w}_{z(j)} + \underline{w}_{z(j)} & (i \in I, j \in J) \end{array}\right.$$

taxation policy with unobs. heterogeneity

- to define a taxation policy, we construct an optimization problem that corresponds to the previous LP

primal

dual

ideal
(Φ_{ij}): given

$$\left[\begin{array}{lll} \max\limits_{d \geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \\ & \sum_{j \in z} \sum_i d_{ij} \leq \bar{o}_z & (z \in Z) \\ & \sum_{j \in z} \sum_i d_{ij} \geq \underline{o}_z & (z \in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{u,v,\overline{w},\underline{w} \geq 0} & \sum\limits_i u_i + \sum\limits_j v_j + \sum\limits_{z \in Z} \bar{o}_z \overline{w}_z - \sum\limits_{z \in Z_L} \underline{o}_z \underline{w}_z & \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} - \overline{w}_{z(j)} + \underline{w}_{z(j)} & (i \in I, j \in J) \end{array}\right.$$

---

w/ unobs. heterogeneity
(Φ_{sh}), P, Q: given

taxation policy with unobs. heterogeneity

- to define a taxation policy, we construct an optimization problem that corresponds to the previous LP

primal

dual

ideal

(Φ_{ij}): given

$$\left[\begin{array}{lll} \max_{d \geq 0} & \sum_{i,j} \Phi_{ij} d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \\ & \sum_{j \in z} \sum_i d_{ij} \leq \bar{o}_z & (z \in Z) \\ & \sum_{j \in z} \sum_i d_{ij} \geq \underline{o}_z & (z \in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min_{u,v,\overline{w},\underline{w} \geq 0} & \sum_i u_i + \sum_j v_j + \sum_{z \in Z} \bar{o}_z \overline{w}_z - \sum_{z \in Z_L} \underline{o}_z \underline{w}_z & \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} - \overline{w}_{z(j)} + \underline{w}_{z(j)} & (i \in I, j \in J) \end{array}\right.$$

---

w/ unobs. heterogeneity

(Φ_{sh})**,** P**,** Q**:** given

$$\left[\begin{array}{lll} \min_{U,V,\underline{w},\overline{w} \geq 0} & G(U) + H(V) + \sum_z \bar{o}_z \overline{w}_z - \sum_z \underline{o}_z \underline{w}_z & \\ \text{s.t.} & U_{sh} + V_{sh} \geq \Phi_{sh} + \underline{w}_z - \overline{w}_z & (s \in S, h \in z, z \in Z) \end{array}\right.$$

taxation policy with unobs. heterogeneity

- to define a taxation policy, we construct an optimization problem that corresponds to the previous LP

primal

dual

ideal

(Φ_{ij}): given

$$\left[\begin{array}{lll} \max\limits_{d\geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \\ & \sum_{j\in z} \sum_i d_{ij} \leq \overline{o}_z & (z \in Z) \\ & \sum_{j\in z} \sum_i d_{ij} \geq \underline{o}_z & (z \in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{u,v,\overline{w},\underline{w}\geq 0} & \sum\limits_i u_i + \sum\limits_j v_j + \sum\limits_{z\in Z} \overline{o}_z \overline{w}_z - \sum\limits_{z\in Z_L} \underline{o}_z \underline{w}_z & \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} - \overline{w}_{z(j)} + \underline{w}_{z(j)} & (i \in I, j \in J) \end{array}\right.$$

---

w/ unobs. heterogeneity

(Φ_{sh}), P, Q: given

$$\left[\begin{array}{lll} \min\limits_{U,V,\underline{w},\overline{w}\geq 0} & G(U) + H(V) + \sum\limits_z \overline{o}_z \overline{w}_z - \sum\limits_z \underline{o}_z \underline{w}_z & \\ \text{s.t.} & U_{sh} + V_{sh} \geq \Phi_{sh} + \underline{w}_z - \overline{w}_z & (s \in S, h \in z, z \in Z) \end{array}\right.$$

taxation policy with unobs. heterogeneity

- to define a taxation policy, we construct an optimization problem that corresponds to the previous LP

primal | **dual**

ideal

(Φ_{ij}): given

primal

$$\left[\begin{array}{lll} \max\limits_{d\geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \\ & \sum_{j\in z} \sum_i d_{ij} \leq \bar{o}_z & (z \in Z) \\ & \sum_{j\in z} \sum_i d_{ij} \geq \underline{o}_z & (z \in Z) \end{array}\right.$$

dual

$$\left[\begin{array}{lll} \min\limits_{u,v,\overline{w},\underline{w}\geq 0} & \sum\limits_i u_i + \sum\limits_j v_j + \sum\limits_{z\in Z} \bar{o}_z \overline{w}_z - \sum\limits_{z\in Z_L} \underline{o}_z \underline{w}_z & \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} - \overline{w}_{z(j)} + \underline{w}_{z(j)} & (i \in I, j \in J) \end{array}\right.$$

---

w/ unobs. heterogeneity

(Φ_{sh}), P, Q: given

$$\left[\begin{array}{lll} \min\limits_{U,V,\underline{w},\overline{w}\geq 0} & G(U) + H(V) + \sum\limits_z \bar{o}_z \overline{w}_z - \sum\limits_z \underline{o}_z \underline{w}_z & \\ \text{s.t.} & U_{sh} + V_{sh} \geq \Phi_{sh} + \underline{w}_z - \overline{w}_z & (s \in S, h \in z, z \in Z) \end{array}\right.$$

taxation policy with unobs. heterogeneity

- to define a taxation policy, we construct an optimization problem that corresponds to the previous LP

primal

dual

ideal

(Φ_{ij}): given

$$\left[\begin{array}{lll} \max\limits_{d\geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \\ & \sum_{j\in z} \sum_i d_{ij} \leq \overline{o}_z & (z \in Z) \\ & \sum_{j\in z} \sum_i d_{ij} \geq \underline{o}_z & (z \in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{u,v,\overline{w},\underline{w}\geq 0} & \sum\limits_i u_i + \sum\limits_j v_j + \sum\limits_{z\in Z} \overline{o}_z \overline{w}_z - \sum\limits_{z\in Z_L} \underline{o}_z \underline{w}_z & \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} - \overline{w}_{z(j)} + \underline{w}_{z(j)} & (i \in I, j \in J) \end{array}\right.$$

---

w/ unobs. heterogeneity

(Φ_{sh})**,** P, Q**:** given

$$\left[\begin{array}{lll} \min\limits_{U,V,\underline{w},\overline{w}\geq 0} & G(U) + H(V) + \sum\limits_z \overline{o}_z \overline{w}_z - \sum\limits_z \underline{o}_z \underline{w}_z & \\ \text{s.t.} & U_{sh} + V_{sh} \geq \Phi_{sh} + \underline{w}_z - \overline{w}_z & (s \in S, h \in z, z \in Z) \end{array}\right.$$

taxation policy with unobs. heterogeneity

- to define a taxation policy, we construct an optimization problem that corresponds to the previous LP

primal

dual

ideal

(Φ_{ij}): given

$$\left[\begin{array}{lll} \max\limits_{d \geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \\ & \sum_{j \in z} \sum_i d_{ij} \leq \bar{o}_z & (z \in Z) \\ & \sum_{j \in z} \sum_i d_{ij} \geq \underline{o}_z & (z \in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{u,v,\overline{w},\underline{w} \geq 0} & \sum\limits_i u_i + \sum\limits_j v_j + \sum\limits_{z \in Z} \bar{o}_z \overline{w}_z - \sum\limits_{z \in Z_L} \underline{o}_z \underline{w}_z & \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} - \overline{w}_{z(j)} + \underline{w}_{z(j)} & (i \in I, j \in J) \end{array}\right.$$

---

w/ unobs. heterogeneity

(Φ_{sh}), P, Q: given

$$\left[\begin{array}{lll} \max\limits_{\mu \geq 0} & \sum\limits_{s,h} \Phi_{sh} \mu_{sh} + \mathcal{E}(\mu) & \\ \text{s.t.} & \sum_{h \in H_0} \mu_{sh} = n_s & (s \in S) \\ & \sum_{s \in S_0} \mu_{sh} = m_h & (h \in H) \\ & \sum_{s \in S_0} \sum_{h \in z} \mu_{sh} \in [\underline{o}_z, \bar{o}_z] & (z \in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{U,V,\underline{w},\overline{w} \geq 0} & G(U) + H(V) + \sum\limits_z \bar{o}_z \overline{w}_z - \sum\limits_z \underline{o}_z \underline{w}_z & \\ \text{s.t.} & U_{sh} + V_{sh} \geq \Phi_{sh} + \underline{w}_z - \overline{w}_z & (s \in S, h \in z, z \in Z) \end{array}\right.$$

taxation policy with unobs. heterogeneity

$$\left[\begin{array}{lll} \max\limits_{\mu \geq 0} & \sum\limits_{s,h} \Phi_{sh} \mu_{sh} + \mathcal{E}(\mu) & \\ \text{s.t.} & \sum_{h \in H_0} \mu_{sh} = n_s & (s \in S) \\ & \sum_{s \in S_0} \mu_{sh} = m_h & (h \in H) \\ & \sum_{s \in S_0} \sum_{h \in z} \mu_{sh} \in [\underline{o}_z, \overline{o}_z] & (z \in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{U, V, \underline{w}, \overline{w} \geq 0} & G(U) + H(V) + \sum\limits_{z} \overline{o}_z \overline{w}_z - \sum\limits_{z} \underline{o}_z \underline{w}_z & \\ \text{s.t.} & U_{sh} + V_{sh} \geq \Phi_{sh} + \underline{w}_z - \overline{w}_z & (s \in S, h \in z, z \in Z) \end{array}\right.$$

taxation policy with unobs. heterogeneity

$$\left[\begin{array}{lll} \max\limits_{\mu \geq 0} & \sum\limits_{s,h} \Phi_{sh}\mu_{sh} + \mathcal{E}(\mu) & \\ \text{s.t.} & \sum_{h \in H_0} \mu_{sh} = n_s & (s \in S) \\ & \sum_{s \in S_0} \mu_{sh} = m_h & (h \in H) \\ & \sum_{s \in S_0} \sum_{h \in z} \mu_{sh} \in [\underline{o}_z, \overline{o}_z] & (z \in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{U,V,\underline{w},\overline{w} \geq 0} & G(U) + H(V) + \sum\limits_{z} \overline{o}_z \overline{w}_z - \sum\limits_{z} \underline{o}_z \underline{w}_z & \\ \text{s.t.} & U_{sh} + V_{sh} \geq \Phi_{sh} + \underline{w}_z - \overline{w}_z & (s \in S, h \in z, z \in Z) \end{array}\right.$$

- PM can formulate and solve these problems given $(\Phi_{sh})_{s,h}$, P, and Q

taxation policy with unobs. heterogeneity

$$\left[\begin{array}{lll} \max\limits_{\mu \geq 0} & \sum\limits_{s,h} \Phi_{sh}\mu_{sh} + \mathcal{E}(\mu) & \\ \text{s.t.} & \sum_{h \in H_0} \mu_{sh} = n_s & (s \in S) \\ & \sum_{s \in S_0} \mu_{sh} = m_h & (h \in H) \\ & \sum_{s \in S_0} \sum_{h \in z} \mu_{sh} \in [\underline{o}_z, \overline{o}_z] & (z \in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{U,V,\underline{w},\overline{w} \geq 0} & G(U) + H(V) + \sum\limits_{z} \overline{o}_z \overline{w}_z - \sum\limits_{z} \underline{o}_z \underline{w}_z & \\ \text{s.t.} & U_{sh} + V_{sh} \geq \Phi_{sh} + \underline{w}_z - \overline{w}_z & (s \in S, h \in z, z \in Z) \end{array}\right.$$

- PM can formulate and solve these problems given $(\Phi_{sh})_{s,h}$, P, and Q
- $\mathscr{E}(\mu) := -G^*(\mu) - H^*(\mu)$,
  where G^* and H^* are the Legendre-Fenchel transforms of G and H, resp.
  - captures the contribution of unobserved heterogeneity (error terms) to social surplus

taxation policy with unobs. heterogeneity

$$\left[\begin{array}{lll} \max\limits_{\mu \geq 0} & \sum\limits_{s,h} \Phi_{sh} \mu_{sh} + \mathcal{E}(\mu) & \\ \text{s.t.} & \sum_{h \in H_0} \mu_{sh} = n_s & (s \in S) \\ & \sum_{s \in S_0} \mu_{sh} = m_h & (h \in H) \\ & \sum_{s \in S_0} \sum_{h \in z} \mu_{sh} \in [\underline{o}_z, \overline{o}_z] & (z \in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{U,V,\underline{w},\overline{w} \geq 0} & G(U) + H(V) + \sum\limits_{z} \overline{o}_z \overline{w}_z - \sum\limits_{z} \underline{o}_z \underline{w}_z & \\ \text{s.t.} & U_{sh} + V_{sh} \geq \Phi_{sh} + \underline{w}_z - \overline{w}_z & (s \in S, h \in z, z \in Z) \end{array}\right.$$

Thm 2: under certain regularity conditions on the error distributions, these optimization problems have a unique solution μ^A and $(U, V, \bar{w}^A, \underline{w}^A)$

taxation policy with unobs. heterogeneity

$$\left[\begin{array}{lll} \max\limits_{\mu\geq 0} & \sum\limits_{s,h}\Phi_{sh}\mu_{sh}+\mathcal{E}(\mu) & \\ \text{s.t.} & \sum_{h\in H_0}\mu_{sh}=n_s & (s\in S) \\ & \sum_{s\in S_0}\mu_{sh}=m_h & (h\in H) \\ & \sum_{s\in S_0}\sum_{h\in z}\mu_{sh}\in[\underline{o}_z,\overline{o}_z] & (z\in Z) \end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{U,V,\underline{w},\overline{w}\geq 0} & G(U)+H(V)+\sum\limits_{z}\overline{o}_z\overline{w}_z-\sum\limits_{z}\underline{o}_z\underline{w}_z & \\ \text{s.t.} & U_{sh}+V_{sh}\geq\Phi_{sh}+\underline{w}_z-\overline{w}_z & (s\in S,h\in z,z\in Z) \end{array}\right.$$

full support + abs. conti. wrt Lebesgue measure

Thm 2: under certain regularity conditions on the error distributions, these optimization problems have a unique solution μ^A and $(U, V, \bar{w}^A, \underline{w}^A)$

taxation policy with unobs. heterogeneity

$$\left[\begin{array}{lll} \max\limits_{\mu\geq 0} & \sum\limits_{s,h}\Phi_{sh}\mu_{sh}+\mathcal{E}(\mu) & \\ \text{s.t.} & \sum_{h\in H_0}\mu_{sh}=n_s & (s\in S)\\ & \sum_{s\in S_0}\mu_{sh}=m_h & (h\in H)\\ & \sum_{s\in S_0}\sum_{h\in z}\mu_{sh}\in[\underline{o}_z,\overline{o}_z] & (z\in Z)\end{array}\right.$$

$$\left[\begin{array}{lll} \min\limits_{U,V,\underline{w},\overline{w}\geq 0} & G(U)+H(V)+\sum\limits_{z}\overline{o}_z\overline{w}_z-\sum\limits_{z}\underline{o}_z\underline{w}_z & \\ \text{s.t.} & U_{sh}+V_{sh}\geq\Phi_{sh}+\underline{w}_z-\overline{w}_z & (s\in S,h\in z,z\in Z)\end{array}\right.$$

full support + abs. conti. wrt Lebesgue measure

Thm 2: under certain regularity conditions on the error distributions, these optimization problems have a unique solution μ^A and $(U,V,\bar{w}^A,\underline{w}^A)$

- define the aggregate-level taxation policy by

$$w_z^A := \bar{w}_z^A - \underline{w}_z^A$$

large market approximation: setup

large market approximation: setup

- consider a sequence of markets indexed by market size $N = |I^N| + |J^N|$

large market approximation: setup

- consider a sequence of markets indexed by market size $N = |I^N| + |J^N|$
- for each market N,
  - denote the optimal policy by w_N^* and the aggregate-level policy by w_N^A

large market approximation: setup

- consider a sequence of markets indexed by market size $N = |I^N| + |J^N|$
- for each market N,
  - denote the optimal policy by w_N^* and the aggregate-level policy by w_N^A
- assume the following objects converge as $N \to \infty$:

$$\frac{n_s^N}{N}, \frac{m_h^N}{N}, \frac{\bar{o}_z^N}{N}, \frac{\underline{o}_z^N}{N}$$

large market approximation: setup

- consider a sequence of markets indexed by market size $N = |I^N| + |J^N|$
- for each market N,
  - denote the optimal policy by w_N^* and the aggregate-level policy by w_N^A
- assume the following objects converge as $N \to \infty$:

$$\frac{n_s^N}{N}, \frac{m_h^N}{N}, \frac{\bar{o}_z^N}{N}, \frac{\underline{o}_z^N}{N}$$

share of type-s doctors

large market approximation: setup

- consider a sequence of markets indexed by market size $N = |I^N| + |J^N|$
- for each market N,
  - denote the optimal policy by w_N^* and the aggregate-level policy by w_N^A
- assume the following objects converge as $N \to \infty$:

$$\frac{n_s^N}{N}, \frac{m_h^N}{N}, \frac{\bar{o}_z^N}{N}, \frac{\underline{o}_z^N}{N}$$

share of type-h positions

large market approximation: setup

- consider a sequence of markets indexed by market size $N = |I^N| + |J^N|$
- for each market N,
  - denote the optimal policy by w_N^* and the aggregate-level policy by w_N^A
- assume the following objects converge as $N \to \infty$:

$$\frac{n_s^N}{N}, \frac{m_h^N}{N}, \boxed{\frac{\bar{o}_z^N}{N}, \frac{\underline{o}_z^N}{N}}$$

normalized regional constraints

large market approximation: result

Thm. 3 (informal): under regularity conditions,

- the two policies asymptotically coincide:

$$\|w_N^A - w_N^*\|_\infty \to 0, \quad \text{a.s.} \quad (N \to \infty)$$

- main outcomes under these policies (e.g., per-capita social surplus and the share of matches for each type pair) also asymptotically coincide.

large market approximation: result

Thm. 3 (informal): under regularity conditions,

- the two policies asymptotically coincide:

$$\|w_N^A - w_N^*\|_\infty \to 0, \quad \text{a.s.} \quad (N \to \infty)$$

- main outcomes under these policies (e.g., per-capita social surplus and the share of matches for each type pair) also asymptotically coincide.

- in large markets, with the structural assumptions,
  the PM can approximately implement the optimal taxation policy
  if she knows $(\Phi_{sh})_{s,h}$, P, and Q

applicability of our framework

- applicable to other problems by reinterpreting types and regions
- in our problem:
  - type of a position = hospital
  - region: prefecture (one geographic area)
- in other settings:
  - type: occupation subcategory (e.g., registered nurse, physician assistant)
  - region: broader occupational category (e.g., healthcare)
  - PM's goal: achieve a balanced distribution across occupations

model

theoretical results

estimation

simulation

available data

from 2016 to 2019

- **aggregate-level matching outcomes**
  - # of matches for each school-hospital pair
  - # of positions offered in each hospital
  - # of applicants from each school

- **salary**
  - monthly salary paid to residents by each program

- **characteristics of hospitals and schools**
  - **hospital:** hospital type (e.g., university-affiliated/government), # of beds, location
  - **school:** public/private, T-score (a measure of entrance exam difficulty) national-exam pass rate, location
  - **pair:** distance, # of previous matches, affiliation

hospital size

students' quality

two-step procedure

- **goal:** to estimate agents' preferences and evaluate them in monetary terms

two-step procedure

- **goal:** to estimate agents' preferences and evaluate them in monetary terms

$$u_i = U_{ij}^{\text{base}} + \tau_i$$

two-step procedure

- **goal:** to estimate agents' preferences and evaluate them in monetary terms

$$u_i = U_{ij}^{\text{base}} + \tau_i$$

base utility

utility that doctor i derives from the match with position j without transfer

two-step procedure

- **goal:** to estimate agents' preferences and evaluate them in monetary terms

$$u_i = U_{ij}^{\text{base}} + \tau_i$$

equilibrium payoff
(actual payoff)

base utility

transfer

utility that doctor i derives from the match with position j without transfer

two-step procedure

- **goal:** to estimate agents' preferences and evaluate them in monetary terms

$$u_i = U_{ij}^{\text{base}} + \tau_i$$

equilibrium payoff (actual payoff) base utility transfer

- we conduct estimation in two steps:

1. estimate equilibrium payoffs
2. estimate the parameters in the base utilities and transfers

why two steps?

1. estimate equilibrium payoff
2. estimate the parameters in the base utilities and transfers

why two steps?

1. estimate equilibrium payoff
2. estimate the parameters in the base utilities and transfers

- **step 1** estimates aggregate-level objects $(U_{sh}, V_{sh}, \Phi_{sh})$ using observed match data
  - we can apply the method of Galichon and Salanié (2021)
  - this is sufficient to set up the optimization problem and compute $(w_z^*)_z$

why two steps?

1. estimate equilibrium payoff
2. estimate the parameters in the base utilities and transfers

- **step 1** estimates aggregate-level objects $(U_{sh}, V_{sh}, \Phi_{sh})$ using observed match data
  - we can apply the method of Galichon and Salanié (2021)
  - this is sufficient to set up the optimization problem and compute $(w_z^*)_z$
- **problem**: the scale of the solution is not determined solely by step 1

why two steps?

1. estimate equilibrium payoff
2. estimate the parameters in the base utilities and transfers

- **step 1** estimates aggregate-level objects $(U_{sh}, V_{sh}, \Phi_{sh})$ using observed match data
  - we can apply the method of Galichon and Salanié (2021)
  - this is sufficient to set up the optimization problem and compute $(w_z^*)_z$

- **problem**: the scale of the solution is not determined solely by step 1
  - to evaluate social surplus and implement a taxation policy, we must express them in monetary terms, such as USD

why two steps?

1. estimate equilibrium payoff
2. estimate the parameters in the base utilities and transfers

- **step 1** estimates aggregate-level objects $(U_{sh}, V_{sh}, \Phi_{sh})$ using observed match data
  - we can apply the method of Galichon and Salanié (2021)
  - this is sufficient to set up the optimization problem and compute $(w_z^*)_z$

- **problem**: the scale of the solution is not determined solely by step 1
  - to evaluate social surplus and implement a taxation policy, we must express them in monetary terms, such as USD

- **step 2** deals with these issues

step 1: estimate equilibrium payoff

$$u_i = U_{ij}^{\text{base}} + \tau_i$$

equilibrium payoff
(actual payoff)

base utility

transfer

step 1: estimate equilibrium payoff

$$u_i = U_{ij}^{\text{base}} + \tau_i$$

equilibrium payoff (actual payoff)   base utility   transfer

- with unobserved heterogeneity, individual-level eqm payoff u_i is not identifiable
- we can still estimate **"aggregate-level equilibrium payoff"** U_{sh}

step 1: estimate equilibrium payoff

$$u_i = U_{ij}^{\text{base}} + \tau_i$$

equilibrium payoff (actual payoff) base utility transfer

- with unobserved heterogeneity, individual-level eqm payoff u_i is not identifiable
- we can still estimate **"aggregate-level equilibrium payoff"** U_{sh}
- Galichon and Salanié (2021) propose nonparametric identification results
  - eqm matching can be seen as a result of "two-sided discrete choice" problems
- we estimate a parametric version using polynomials

(see the paper for details)

step 2: estimate agents' preferences

$$u_i = U_{ij}^{\text{base}} + \tau_i$$

equilibrium payoff (actual payoff)

base utility

transfer

step 2: estimate agents' preferences

$$u_i = U_{ij}^{\text{base}} + \tau_i$$

equilibrium payoff
(actual payoff)

base utility

transfer

- impose two structural assumptions for base utilities:

step 2: estimate agents' preferences

$$u_i = U_{ij}^{\text{base}} + \tau_i$$

equilibrium payoff (actual payoff) base utility transfer

- impose two structural assumptions for base utilities:
  - **additive separability**: $U_{ij}^{\text{base}} = U_{sh}^{\text{base}} + \varepsilon_{ih} \qquad (j \in h)$

step 2: estimate agents' preferences

$$u_i = U_{ij}^{\text{base}} + \tau_i$$

equilibrium payoff (actual payoff)　base utility　transfer

- impose two structural assumptions for base utilities:
  - **additive separability:** $U_{ij}^{\text{base}} = U_{sh}^{\text{base}} + \varepsilon_{ih} \qquad (j \in h)$

aggregate-level base utility

step 2: estimate agents' preferences

$$u_i = U_{ij}^{\text{base}} + \tau_i$$

equilibrium payoff (actual payoff)

base utility

transfer

- impose two structural assumptions for base utilities:
  - **additive separability:** $U_{ij}^{\text{base}} = U_{sh}^{\text{base}} + \varepsilon_{ih} \qquad (j \in h)$
  - **linearity:** $U_{sh}^{\text{base}} = X_U^{\top} \beta_U$

parameters

aggregate-level base utility

step 2: measurement model for transfer

$$u_i = U_{ij}^{\text{base}} + \tau_i$$

equilibrium payoff (actual payoff)

base utility

transfer

- for transfers, we assume that agents' utilities are **quasi-linear** in monetary transfers

step 2: measurement model for transfer

$$u_i = U_{ij}^{\text{base}} + \tau_i$$

equilibrium payoff (actual payoff)

base utility

transfer

- for transfers, we assume that agents' utilities are **quasi-linear** in monetary transfers
- the transfer term may reflect multiple components:
  - salary, workload, experience, risk of medical accidents, etc.

step 2: measurement model for transfer

$$u_i = U_{ij}^{\text{base}} + \tau_i$$

equilibrium payoff (actual payoff)

base utility

transfer

- for transfers, we assume that agents' utilities are **quasi-linear** in monetary transfers
- the transfer term may reflect multiple components:
  - salary, workload, experience, risk of medical accidents, etc.

unobservable

step 2: estimating equation

$$\sum_s \omega_{sh}\hat{U}_{sh} = \gamma_{0,U} + \gamma_{1,U}S_h + \sum_s \omega_{sh}\left((X_{sh}^{U,\text{base}})^\top \beta_U\right) + \psi_h^U \qquad (h \in H)$$

step 2: estimating equation

$$\boxed{\sum_s \omega_{sh}\hat{U}_{sh}} = \gamma_{0,U} + \gamma_{1,U}S_h + \sum_s \omega_{sh}\left((X_{sh}^{U,\mathrm{base}})^\top \beta_U\right) + \psi_h^U \qquad (h \in H)$$

weighted sum of
equilibrium payoffs
(constructed in step 1)

step 2: estimating equation

$$\sum_{s} \omega_{sh}\hat{U}_{sh} = \gamma_{0,U} + \gamma_{1,U} S_h + \sum_{s} \omega_{sh}\left((X_{sh}^{U,\text{base}})^{\top}\beta_U\right) + \psi_h^U \qquad (h \in H)$$

weighted sum of
equilibrium payoffs
(constructed in step 1)

parameters of interest

step 2: estimating equation

marginal utility of money for doctors

$$\sum_{s} \omega_{sh}\hat{U}_{sh} = \gamma_{0,U} + \gamma_{1,U} S_h + \sum_{s} \omega_{sh}\left((X_{sh}^{U,\text{base}})^{\top}\beta_U\right) + \psi_h^U \qquad (h \in H)$$

weighted sum of
equilibrium payoffs
(constructed in step 1)

parameters of interest

step 2: estimating equation

marginal utility of money for doctors

$$\sum_s \omega_{sh}\hat{U}_{sh} = \gamma_{0,U} + \gamma_{1,U} S_h + \sum_s \omega_{sh}\left((X_{sh}^{U,\mathrm{base}})^{\top} \beta_U\right) + \psi_h^U \qquad (h \in H)$$

weighted sum of
equilibrium payoffs
(constructed in step 1)

parameters of interest

unobserved transfer
from hospital h
to matched doctors

step 2: estimating equation

marginal utility of money for doctors

$$\sum_s \omega_{sh}\hat{U}_{sh} = \gamma_{0,U} + \gamma_{1,U} S_h + \sum_s \omega_{sh}\left((X_{sh}^{U,\text{base}})^{\top}\beta_U\right) + \psi_h^U \qquad (h \in H)$$

weighted sum of
equilibrium payoffs
(constructed in step 1)

parameters of interest

unobserved transfer
from hospital h
to matched doctors

- we estimate coefficients $(\gamma_{1,U}, \beta_U)$ using IVs

step 2: estimating equation

marginal utility of money for doctors

$$\sum_s \omega_{sh}\hat{U}_{sh} = \gamma_{0,U} + \gamma_{1,U} S_h + \sum_s \omega_{sh}\left((X_{sh}^{U,\text{base}})^{\top} \beta_U\right) + \psi_h^U \qquad (h \in H)$$

weighted sum of equilibrium payoffs (constructed in step 1)

parameters of interest

unobserved transfer from hospital h to matched doctors

- we estimate coefficients $(\gamma_{1,U}, \beta_U)$ using IVs

(we conduct analogous estimation for hospitals)

estimation results

- geography matters
- students prefer closer and larger hospitals
  - 10% decrease in the distance ≈ gain of **$33/month**
  - 10% increase in hospital beds ≈ gain of **$53/month**
  - same prefecture match ≈ gain of **$1,655/month** (but the point estimate may be imprecise)
- hospitals favor closer, public-school, and higher-T-score students (but the point estimates may be imprecise)

higher quality

model

theoretical results

estimation

simulation

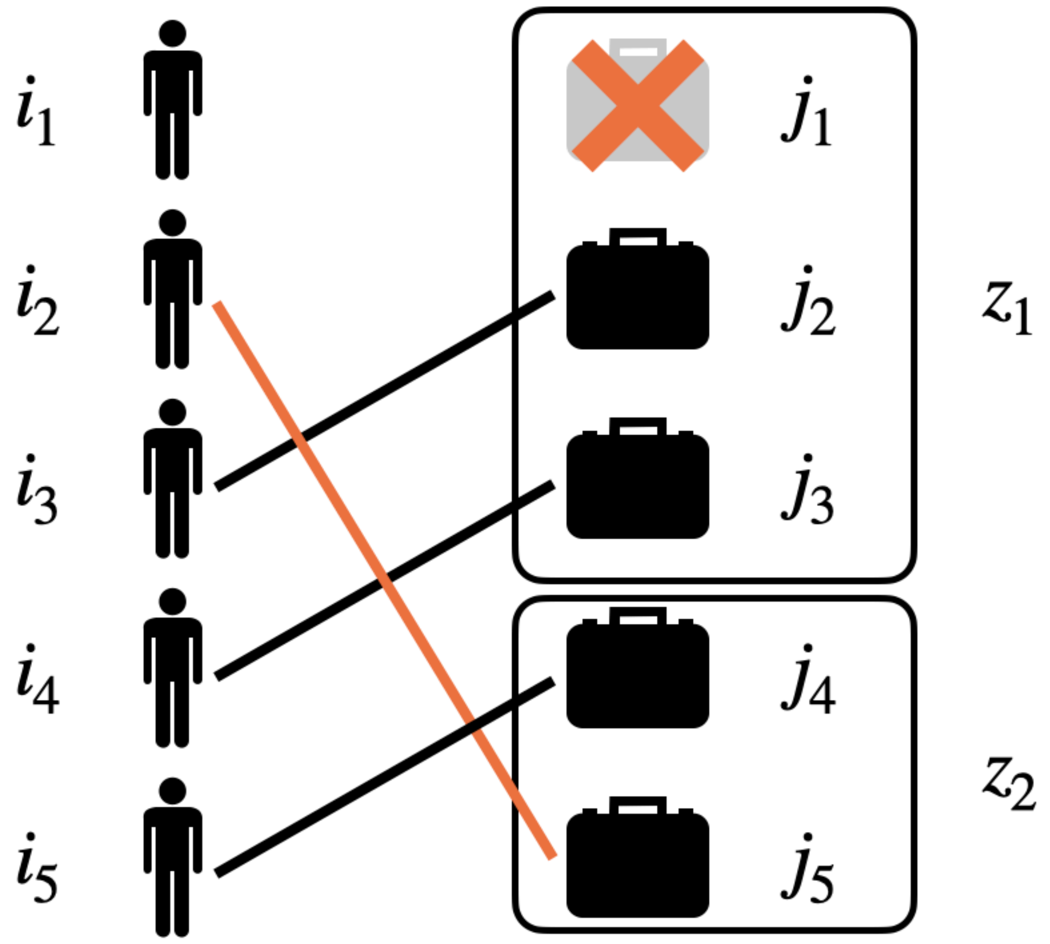
simulation setup

- we simulate the JRMP market in 2019 using the estimated model
- compare three different policies:

simulation setup

- we simulate the JRMP market in 2019 using the estimated model
- compare three different policies:

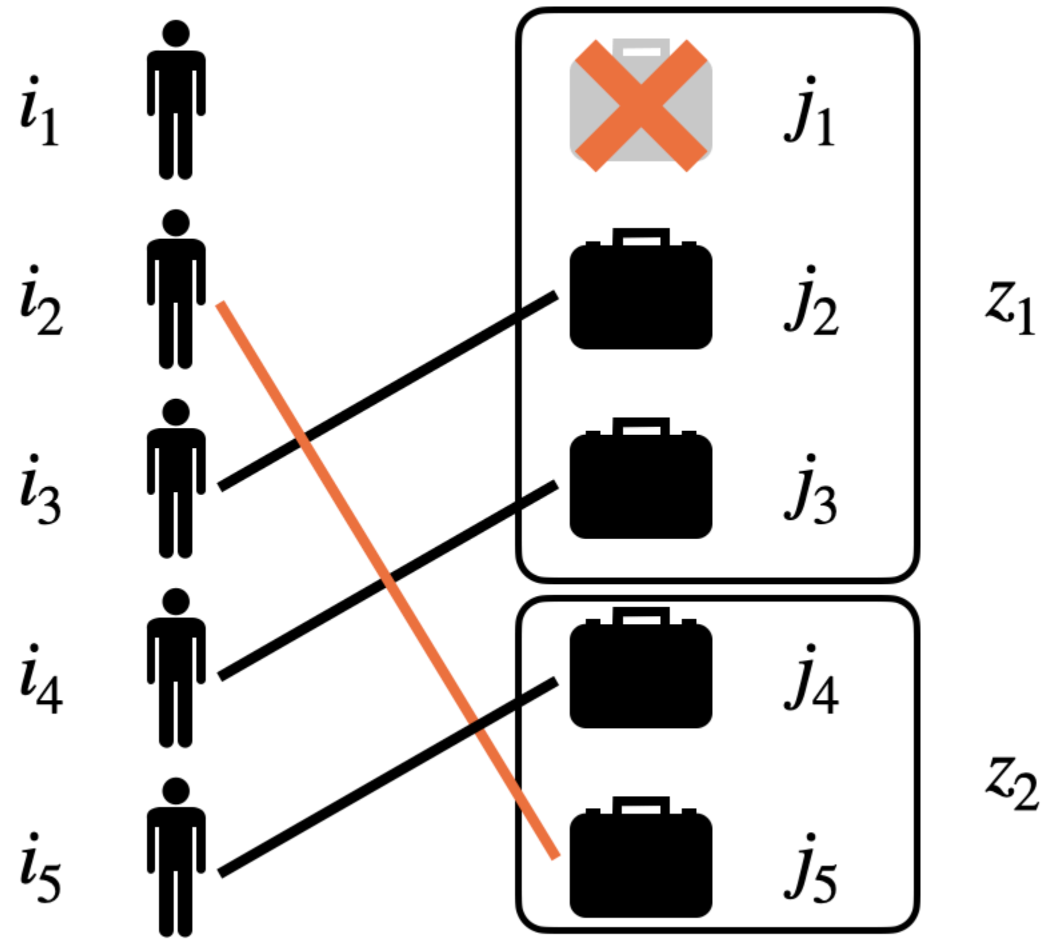
Artificial Caps (AC)
(cap on urban areas)



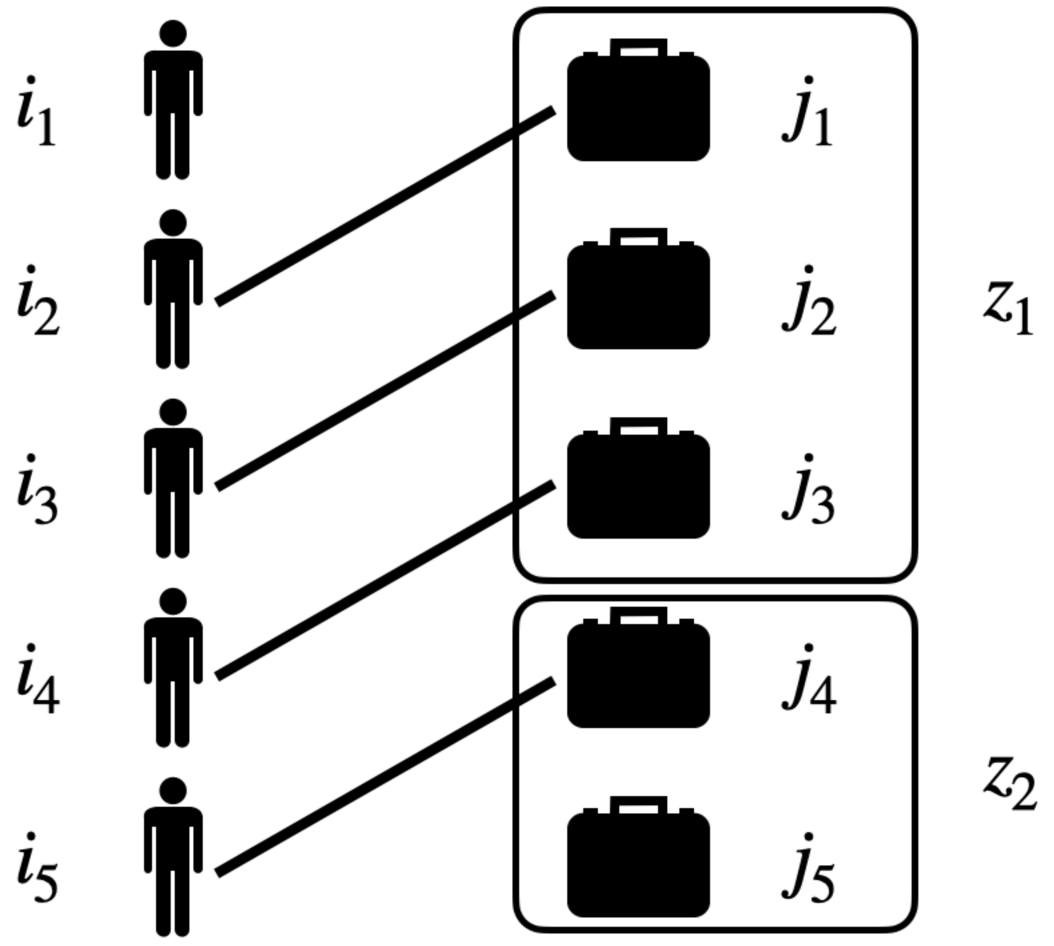
simulation setup

- we simulate the JRMP market in 2019 using the estimated model
- compare three different policies:

Artificial Caps (AC)
(cap on urban areas)



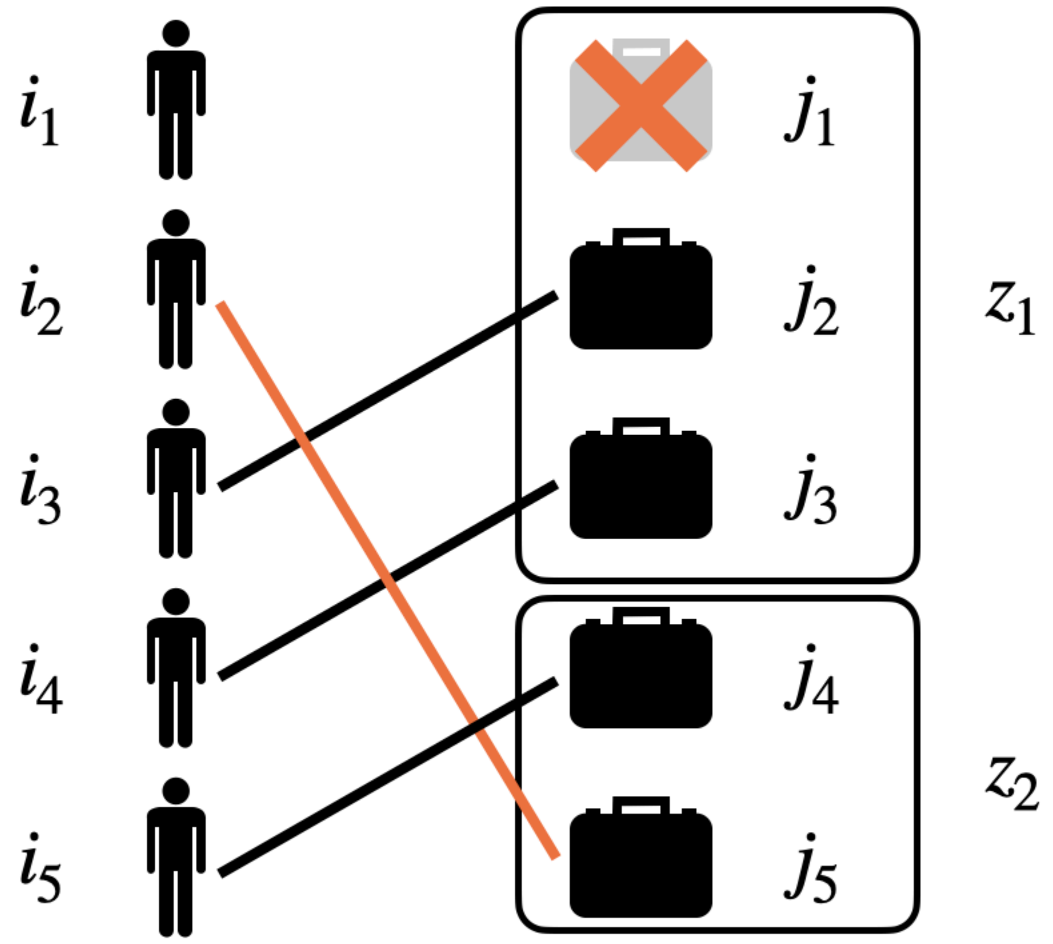
No Cap Tightening (NC)
(all positions available)



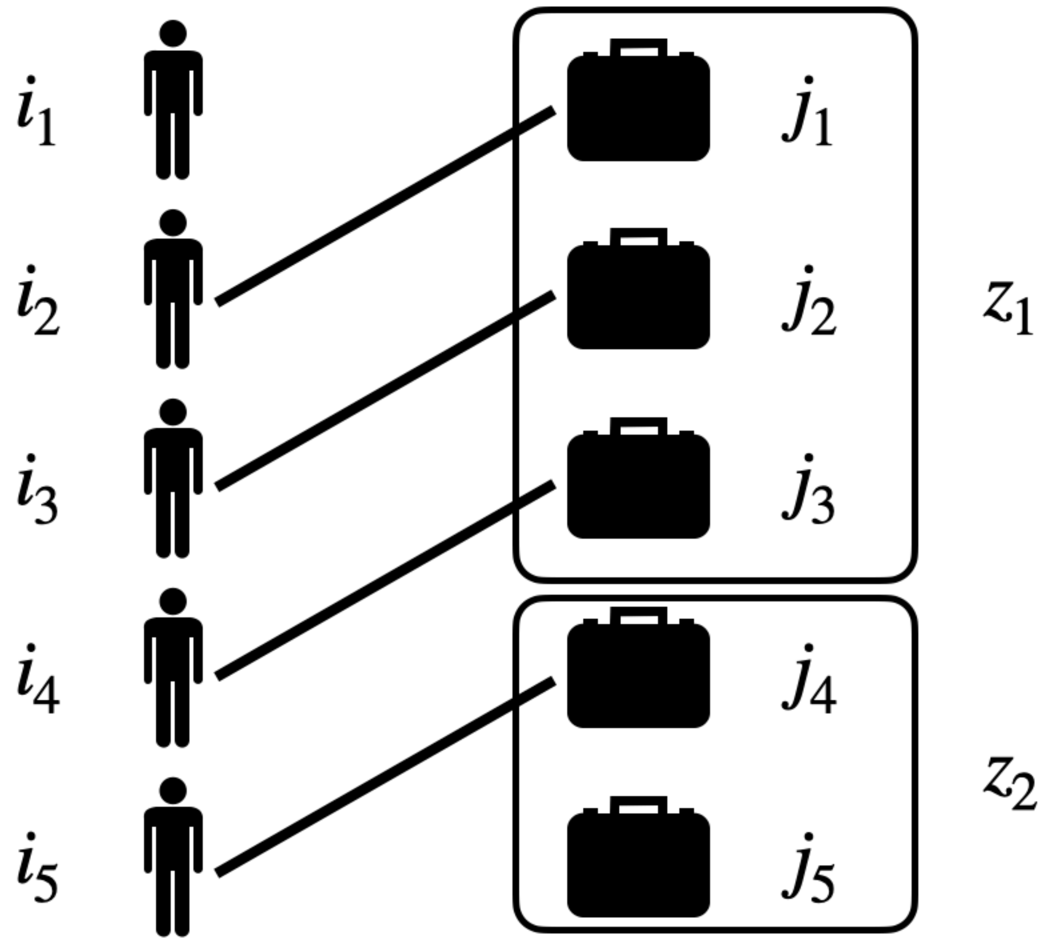
simulation setup

- we simulate the JRMP market in 2019 using the estimated model
- compare three different policies:

Artificial Caps (AC)
(cap on urban areas)



No Cap Tightening (NC)
(all positions available)

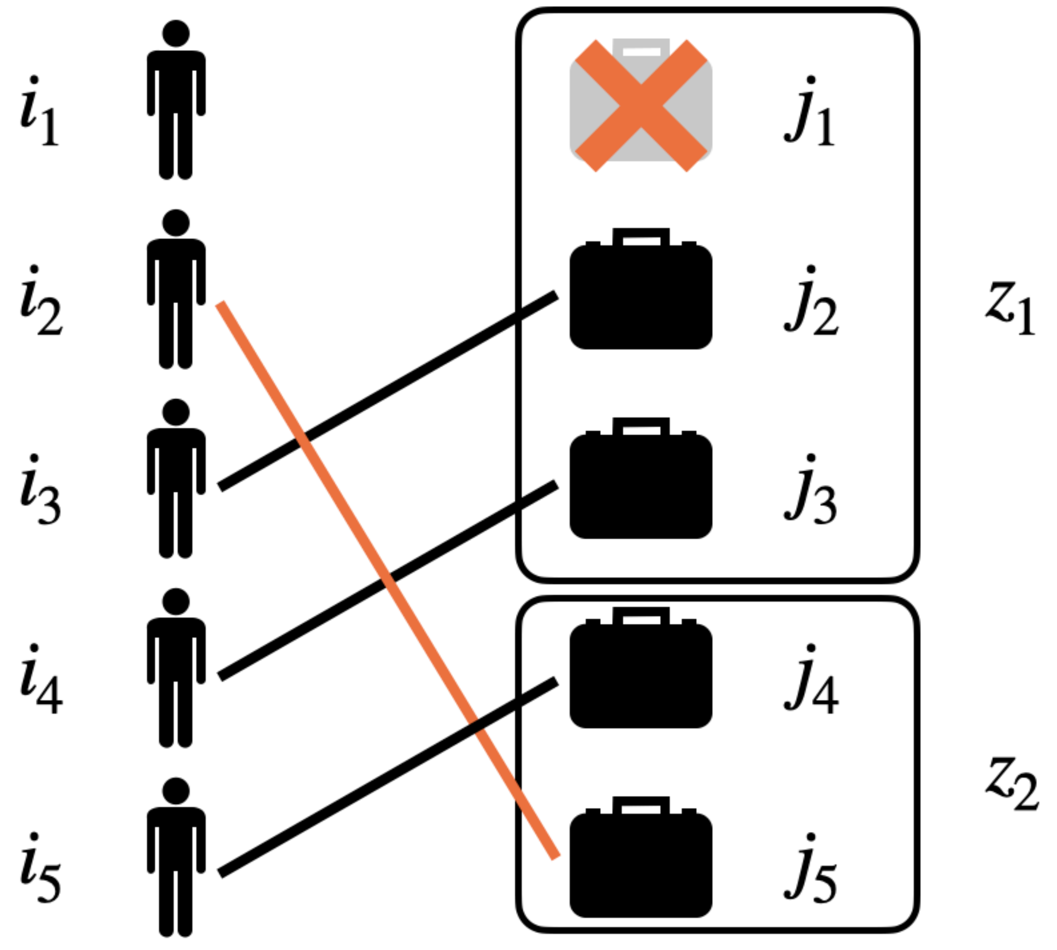


for each hospital, we define its true capacity as
the maximum number of positions it offered up to 2019

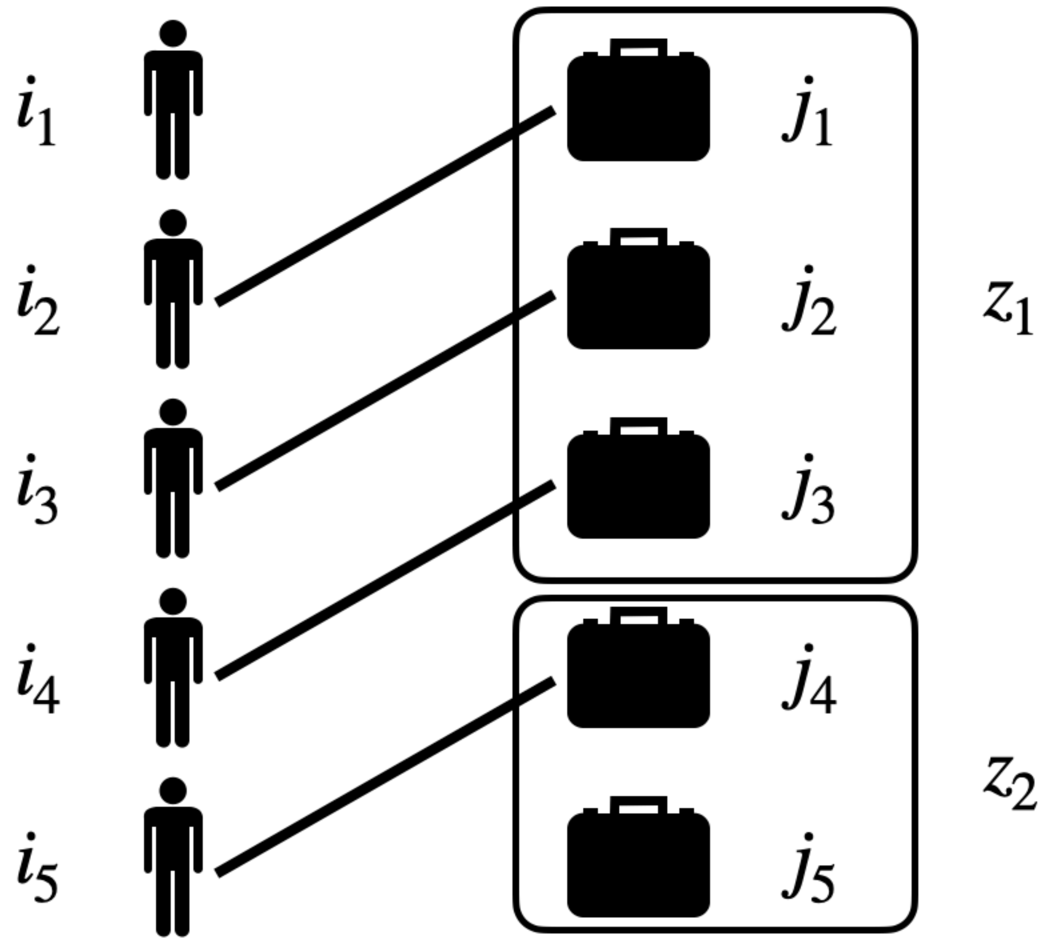
simulation setup

- we simulate the JRMP market in 2019 using the estimated model
- compare three different policies:

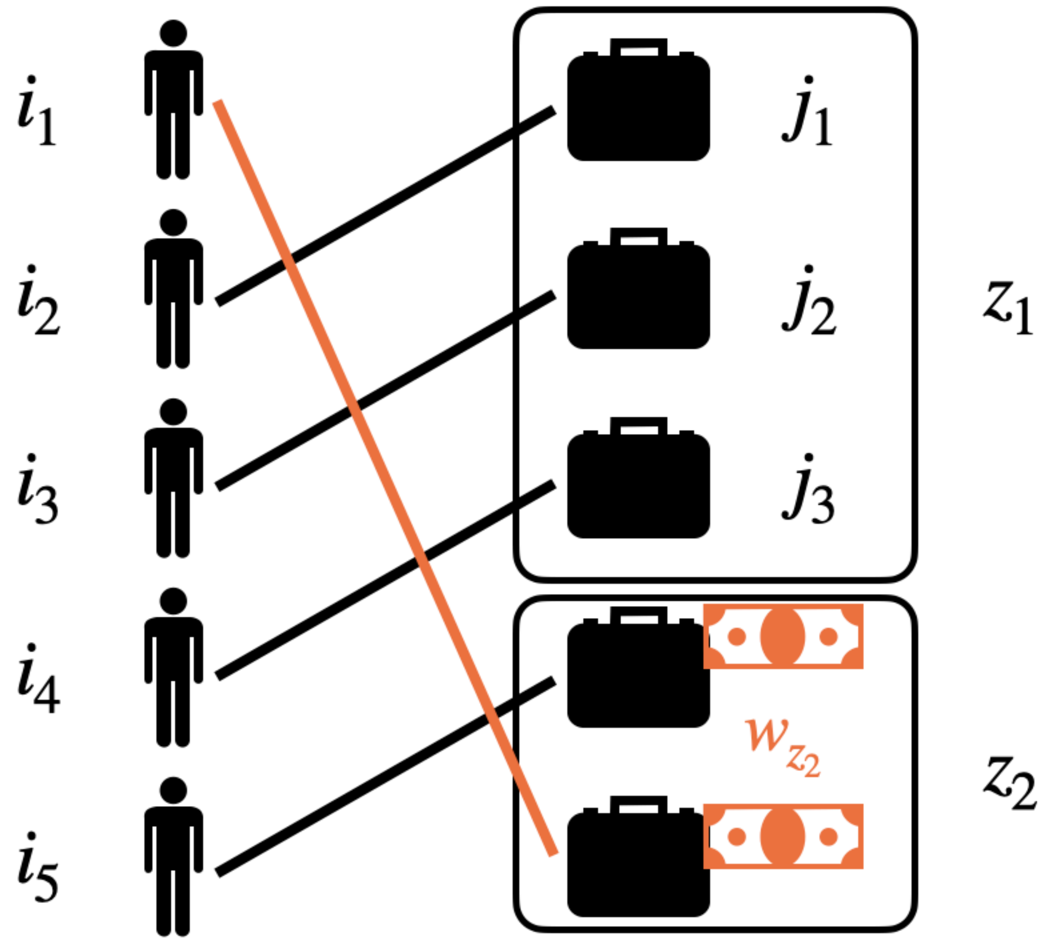
Artificial Caps (AC)
(cap on urban areas)



No Cap Tightening (NC)
(all positions available)



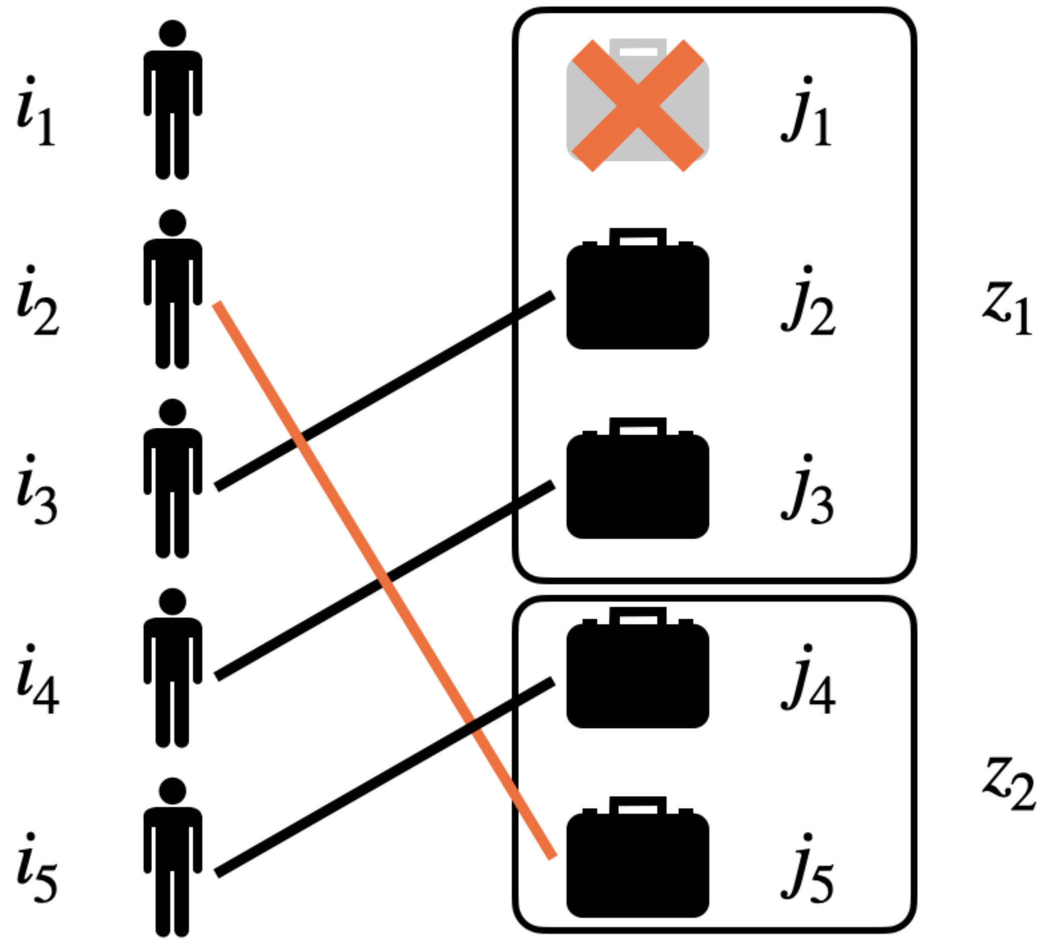
Optimal Subsidy (OS)
(all positions available + subsidy)



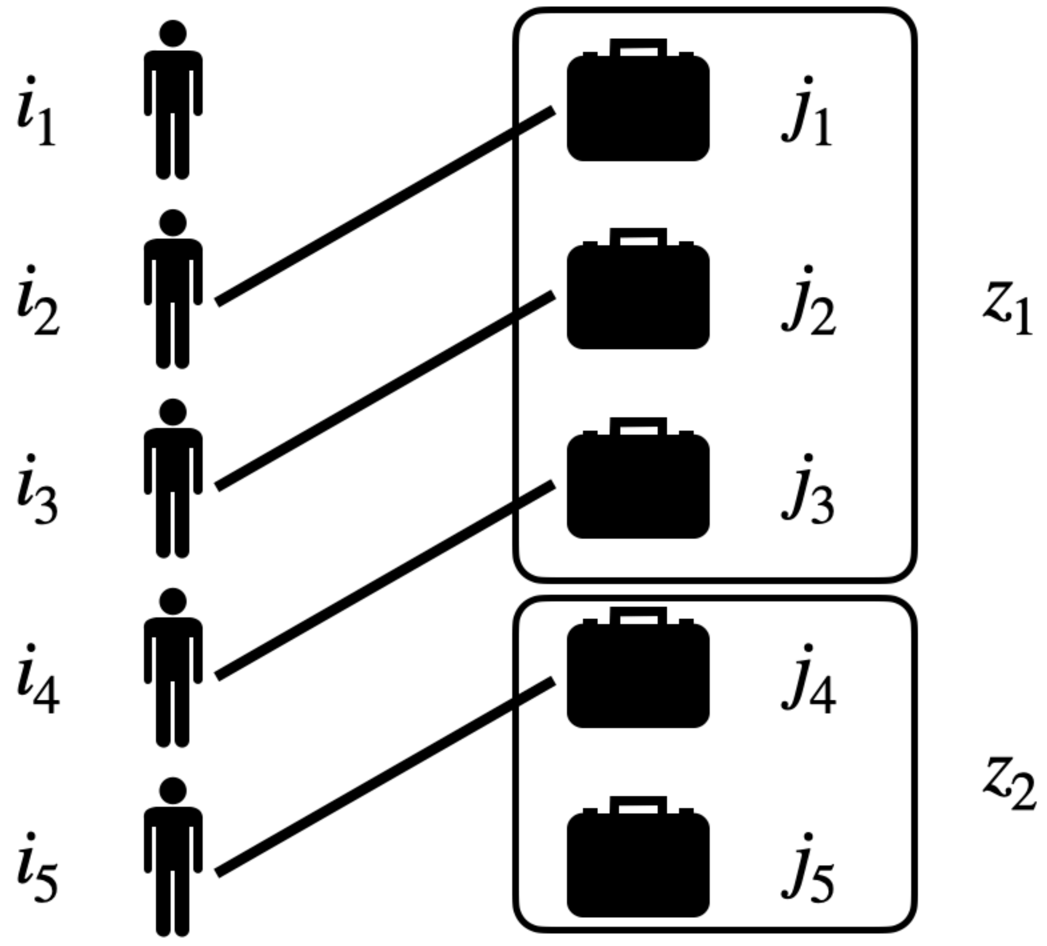
simulation setup

- we simulate the JRMP market in 2019 using the estimated model
- compare three different policies:

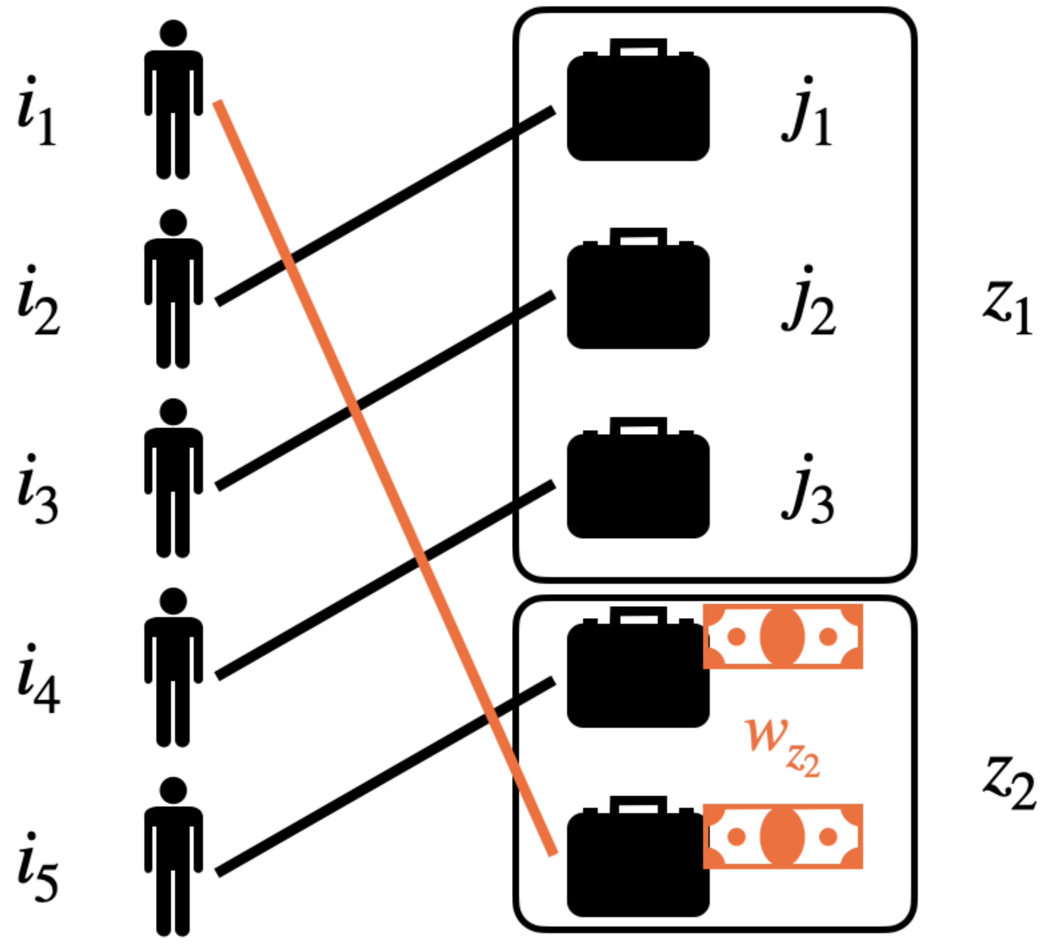
Artificial Caps (AC)
(cap on urban areas)



No Cap Tightening (NC)
(all positions available)



Optimal Subsidy (OS)
(all positions available + subsidy)



regional constraints: 15 rural prefectures receive at least as many residents as they did under the AC

welfare ranking

Artificial Caps (AC)
(cap on urban areas)

No Cap Tightening (NC)
(all slots available)

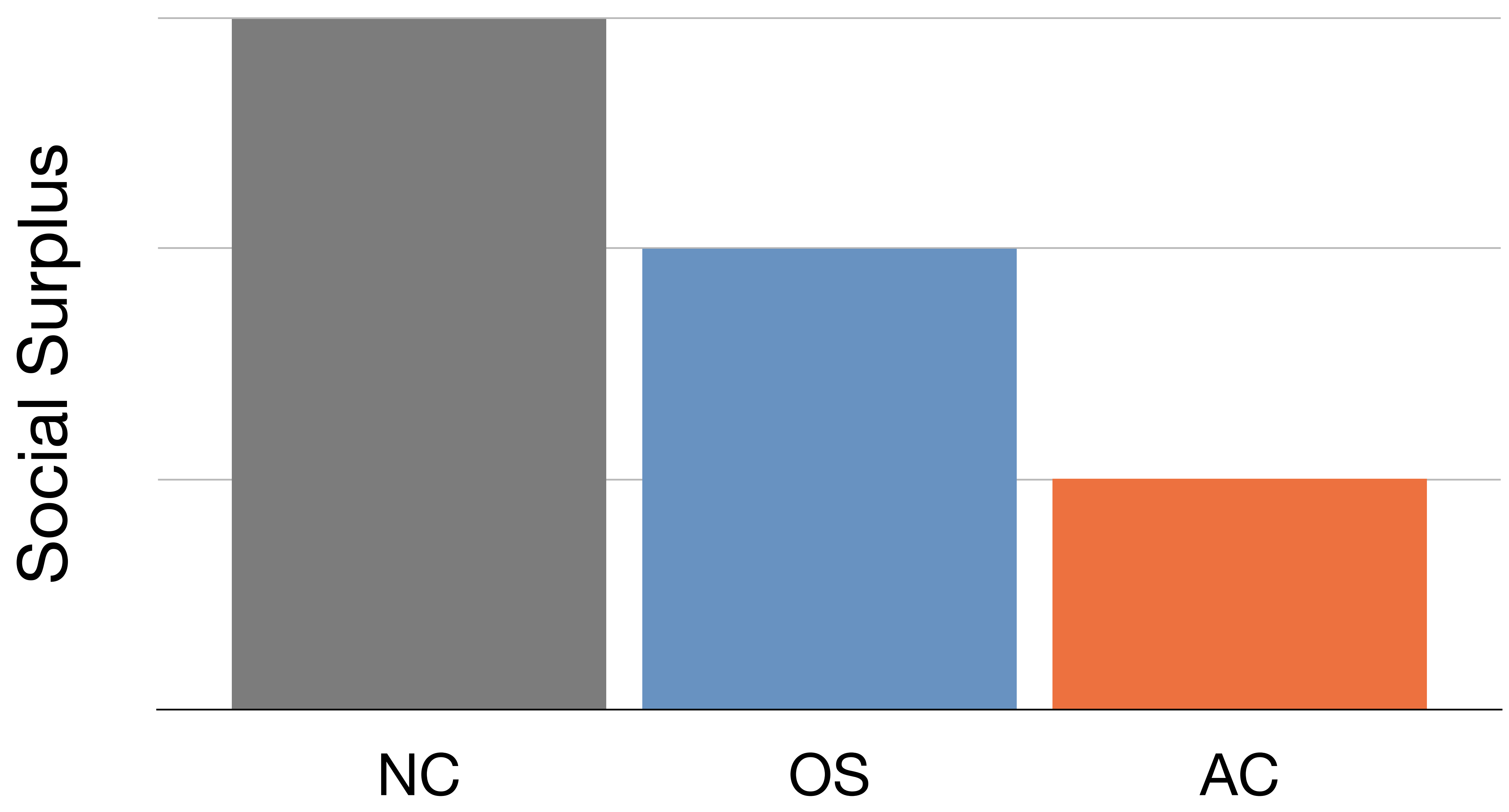
Optimal Subsidy (OS)
(all slots available + subsidy)

welfare ranking

Artificial Caps (AC)
(cap on urban areas)

No Cap Tightening (NC)
(all slots available)

Optimal Subsidy (OS)
(all slots available + subsidy)



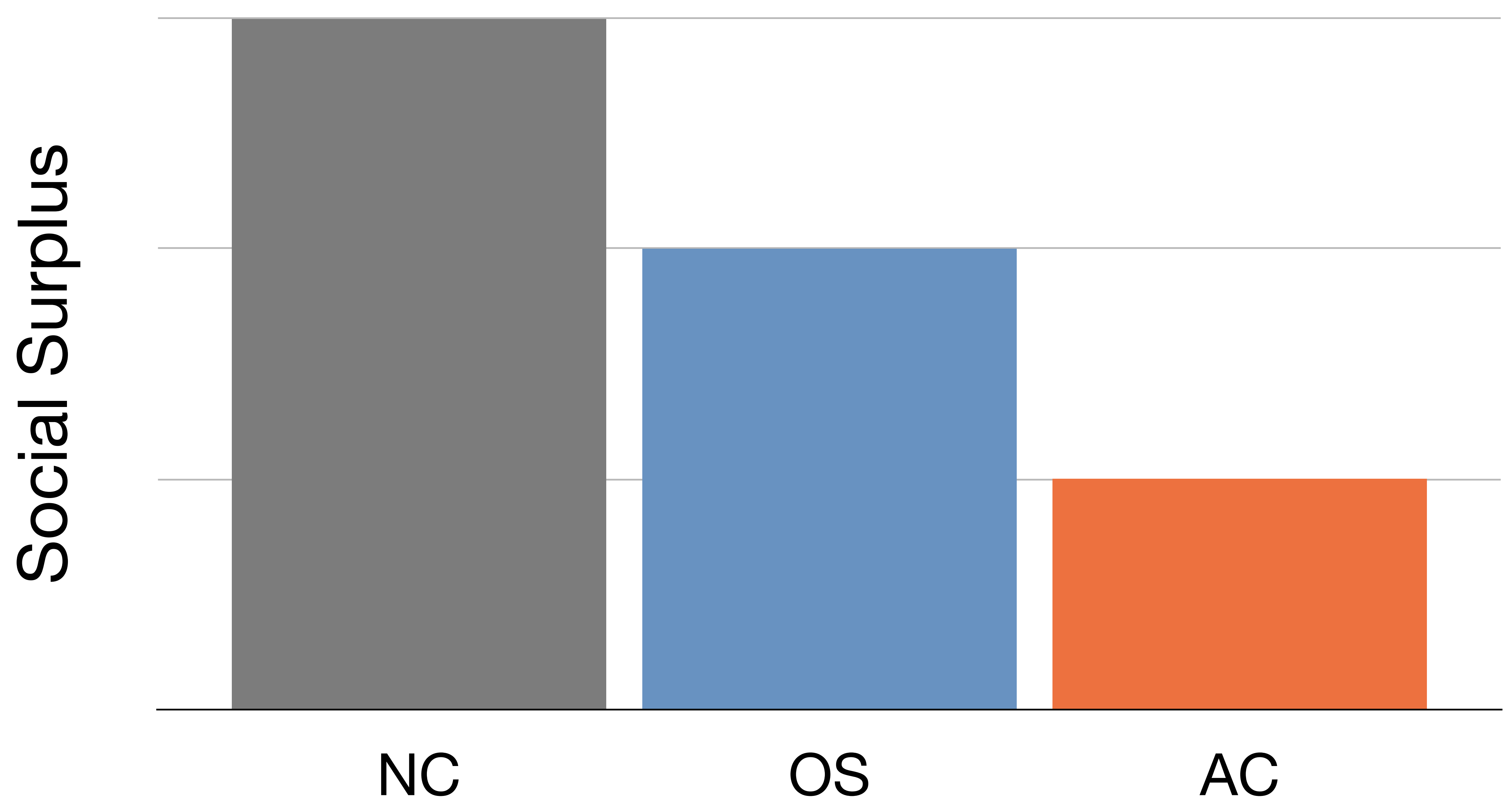
welfare ranking

Artificial Caps (AC)
(cap on urban areas)

No Cap Tightening (NC)
(all slots available)

Optimal Subsidy (OS)
(all slots available + subsidy)

Fact: **NC** maximizes social surplus (w/o regional constraints)



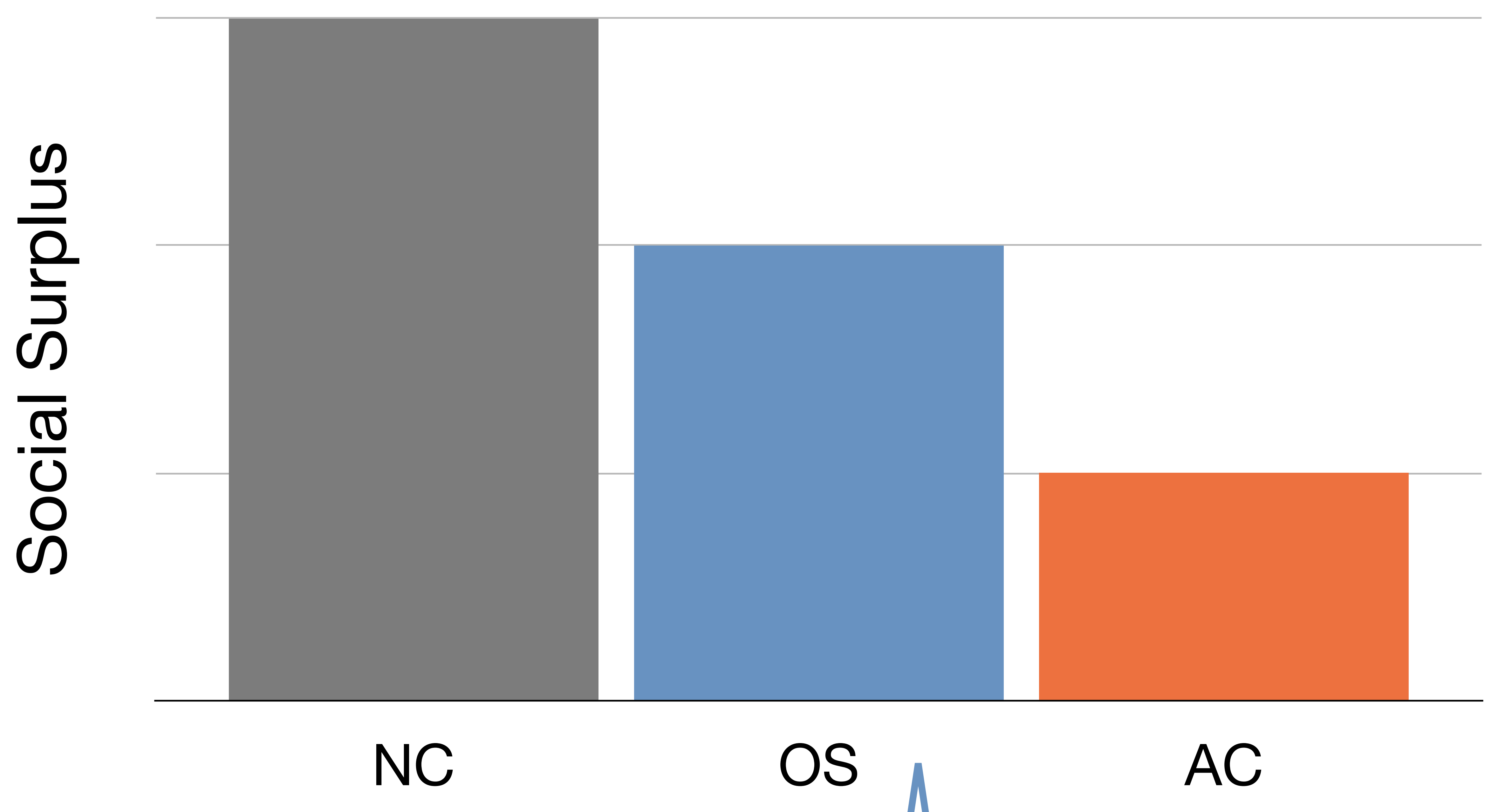
welfare ranking

Artificial Caps (AC)
(cap on urban areas)

No Cap Tightening (NC)
(all slots available)

Optimal Subsidy (OS)
(all slots available + subsidy)

Fact: **NC** maximizes social surplus (w/o regional constraints)



Thm: **OS** maximizes social surplus s.t. regional constraints

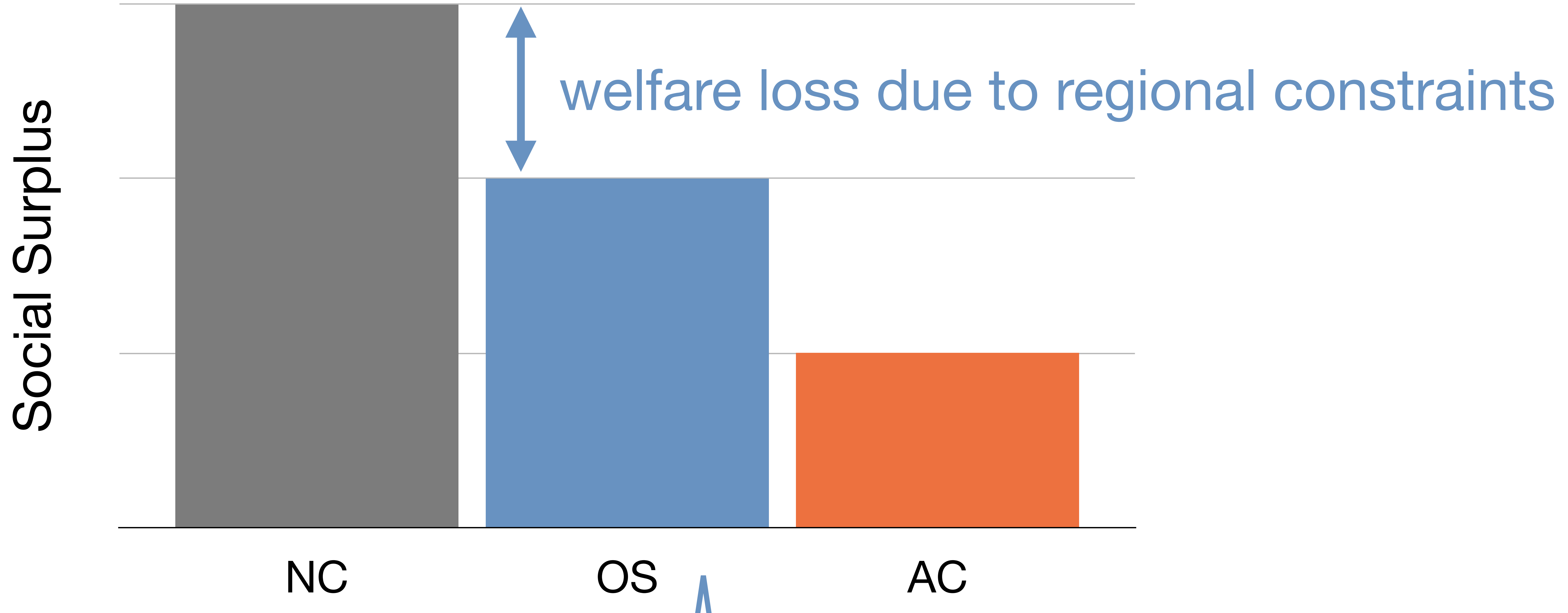
welfare ranking

Artificial Caps (AC)
(cap on urban areas)

No Cap Tightening (NC)
(all slots available)

Optimal Subsidy (OS)
(all slots available + subsidy)

Fact: **NC** maximizes social surplus (w/o regional constraints)



Thm: **OS** maximizes social surplus s.t. regional constraints

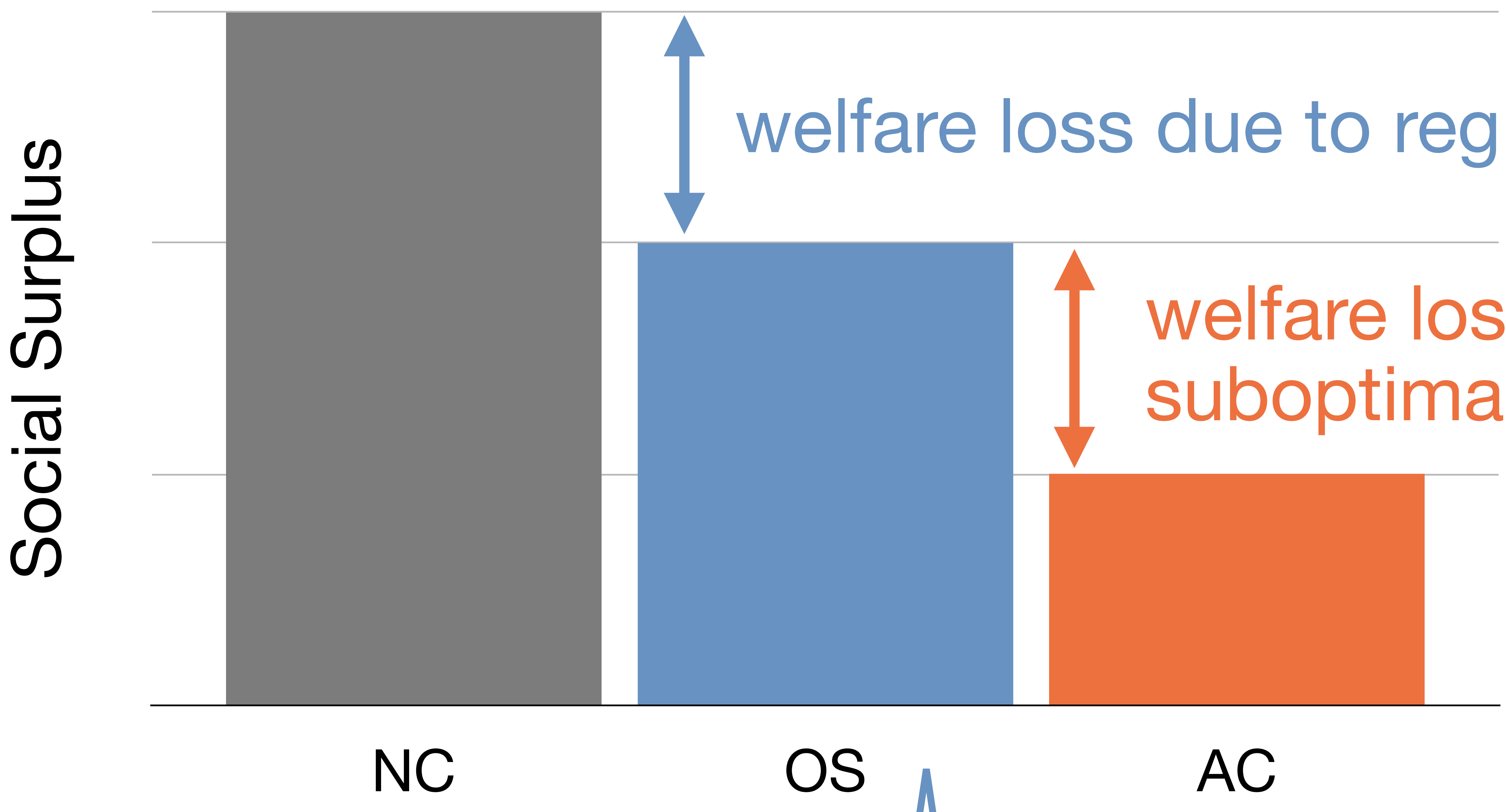
welfare ranking

Artificial Caps (AC)
(cap on urban areas)

No Cap Tightening (NC)
(all slots available)

Optimal Subsidy (OS)
(all slots available + subsidy)

Fact: **NC** maximizes social surplus (w/o regional constraints)



welfare loss due to regional constraints

welfare loss from a suboptimal policy instrument

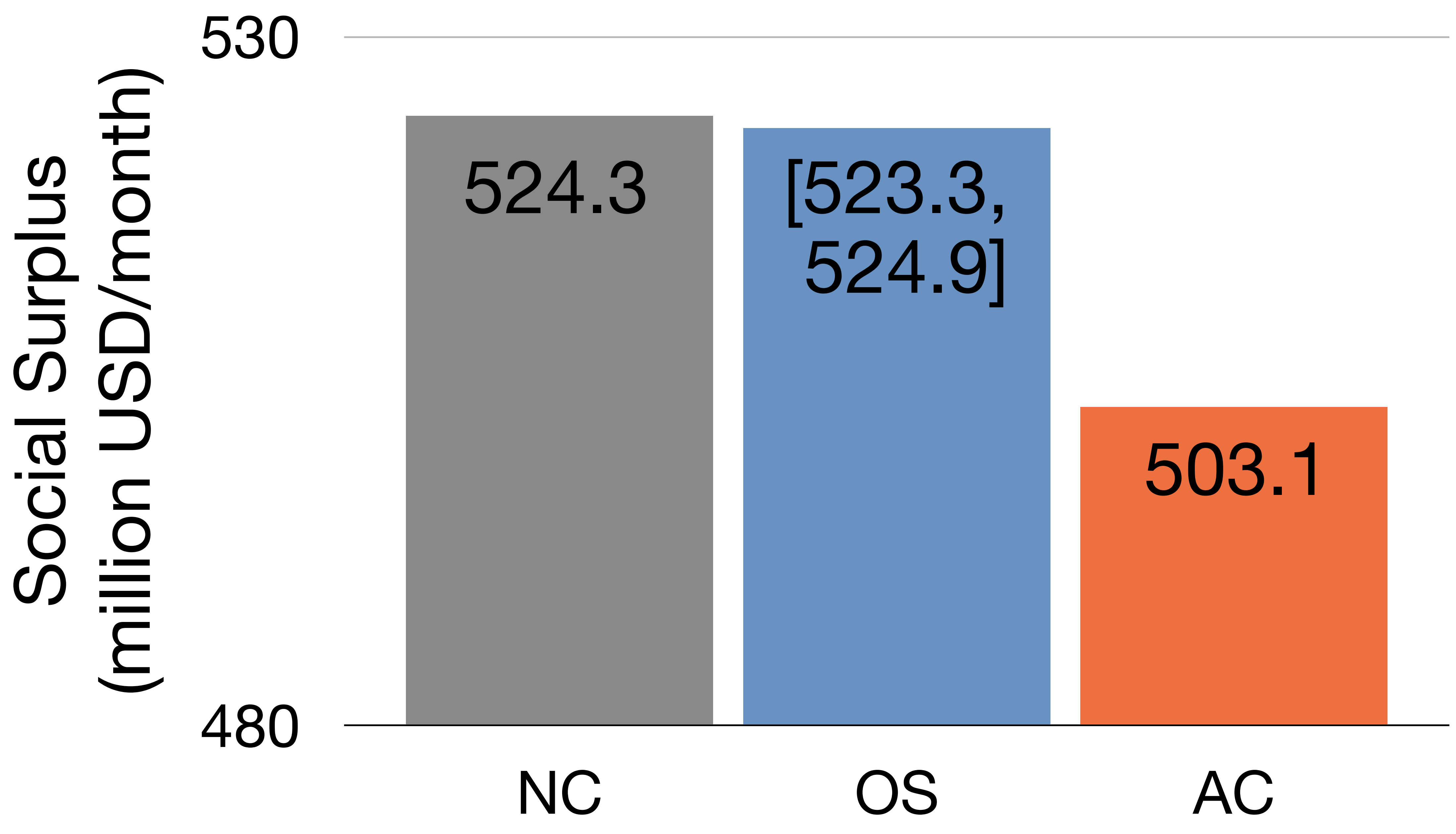
Thm: **OS** maximizes social surplus s.t. regional constraints

simulation results

Artificial Caps (AC)
(cap on urban areas)

No Cap Tightening (NC)
(all slots available)

Optimal Subsidy (OS)
(all slots available + subsidy)

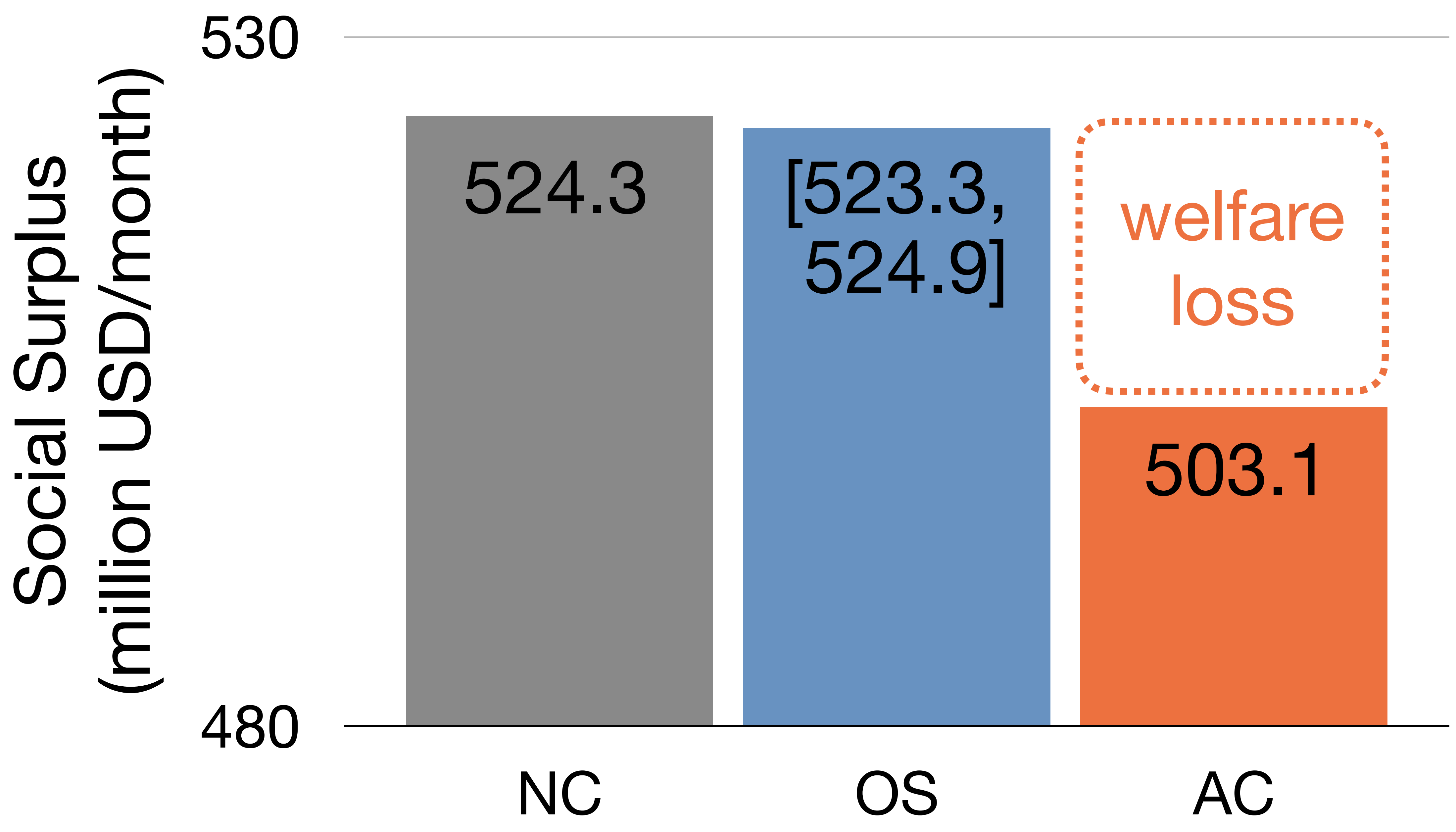


Artificial Caps (AC)
(cap on urban areas)

No Cap Tightening (NC)
(all slots available)

Optimal Subsidy (OS)
(all slots available + subsidy)

simulation results



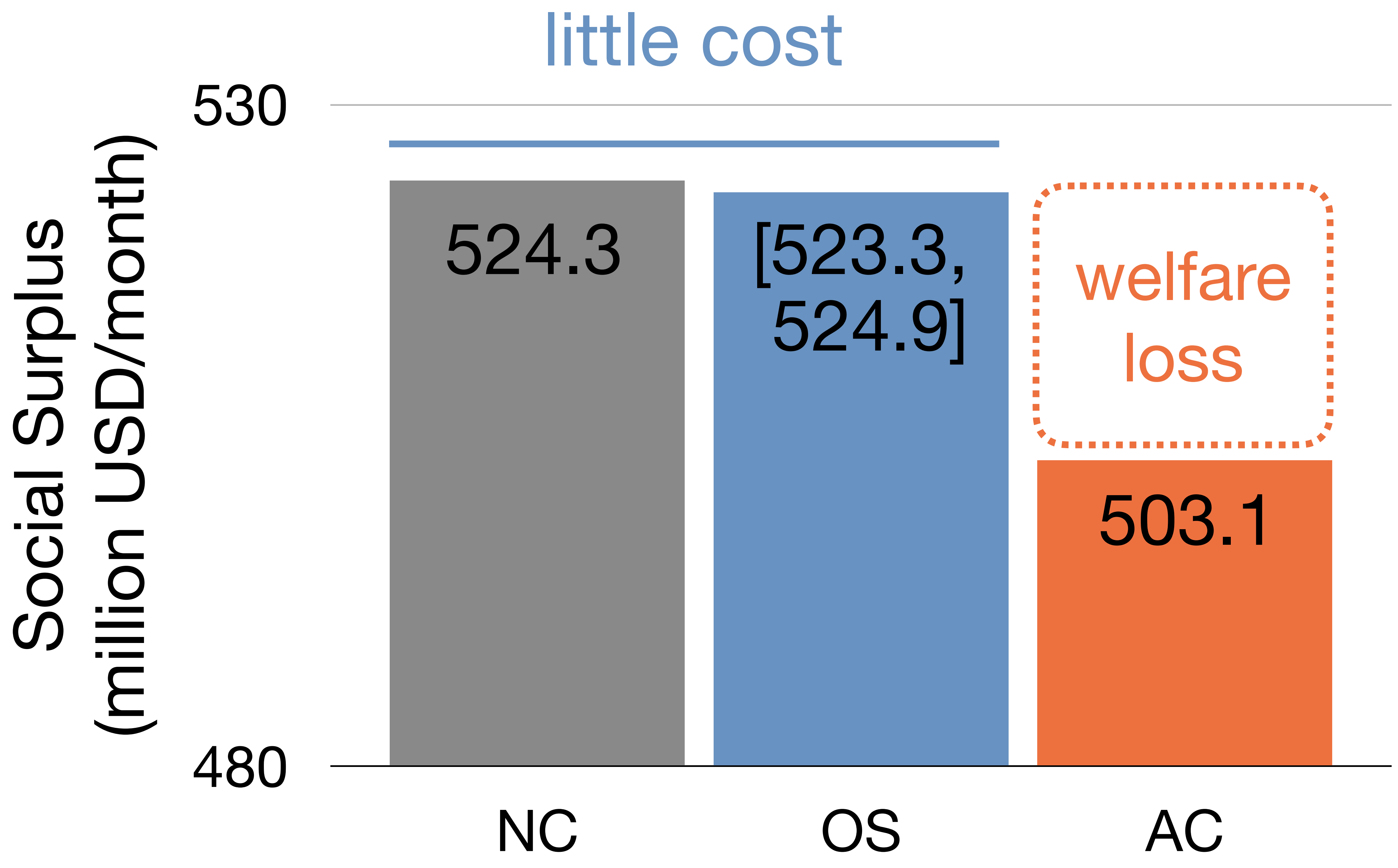
- status quo policy (AC) generates a significant welfare loss ($21.3M per month)

Artificial Caps (AC)
(cap on urban areas)

No Cap Tightening (NC)
(all slots available)

Optimal Subsidy (OS)
(all slots available + subsidy)

simulation results



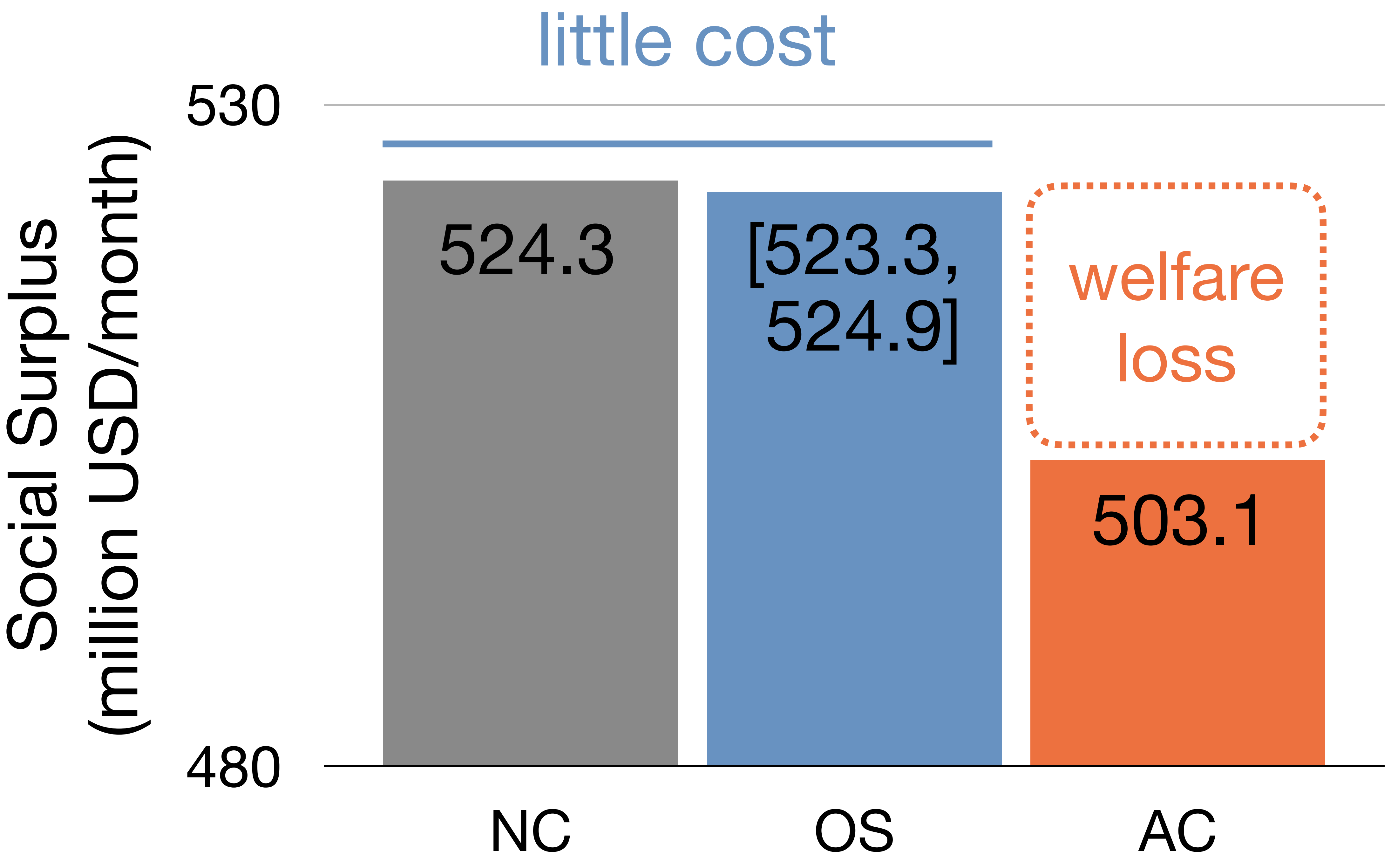
- status quo policy (AC) generates a significant welfare loss ($21.3M per month)
- same distributional outcome can be achieved by the optimal subsidy policy with little welfare loss from regional constraints

Artificial Caps (AC)
(cap on urban areas)

No Cap Tightening (NC)
(all slots available)

Optimal Subsidy (OS)
(all slots available + subsidy)

simulation results



- status quo policy (AC) generates a significant welfare loss ($21.3M per month)
- same distributional outcome can be achieved by the optimal subsidy policy with little welfare loss from regional constraints
- subsidy cost: $[0.60M, 2.20M] per month
    - $[48.6, 178.3] per student
    - $[497.1, 1824.9] per recipient
    - substantial, but comparable to existing policies

mechanism

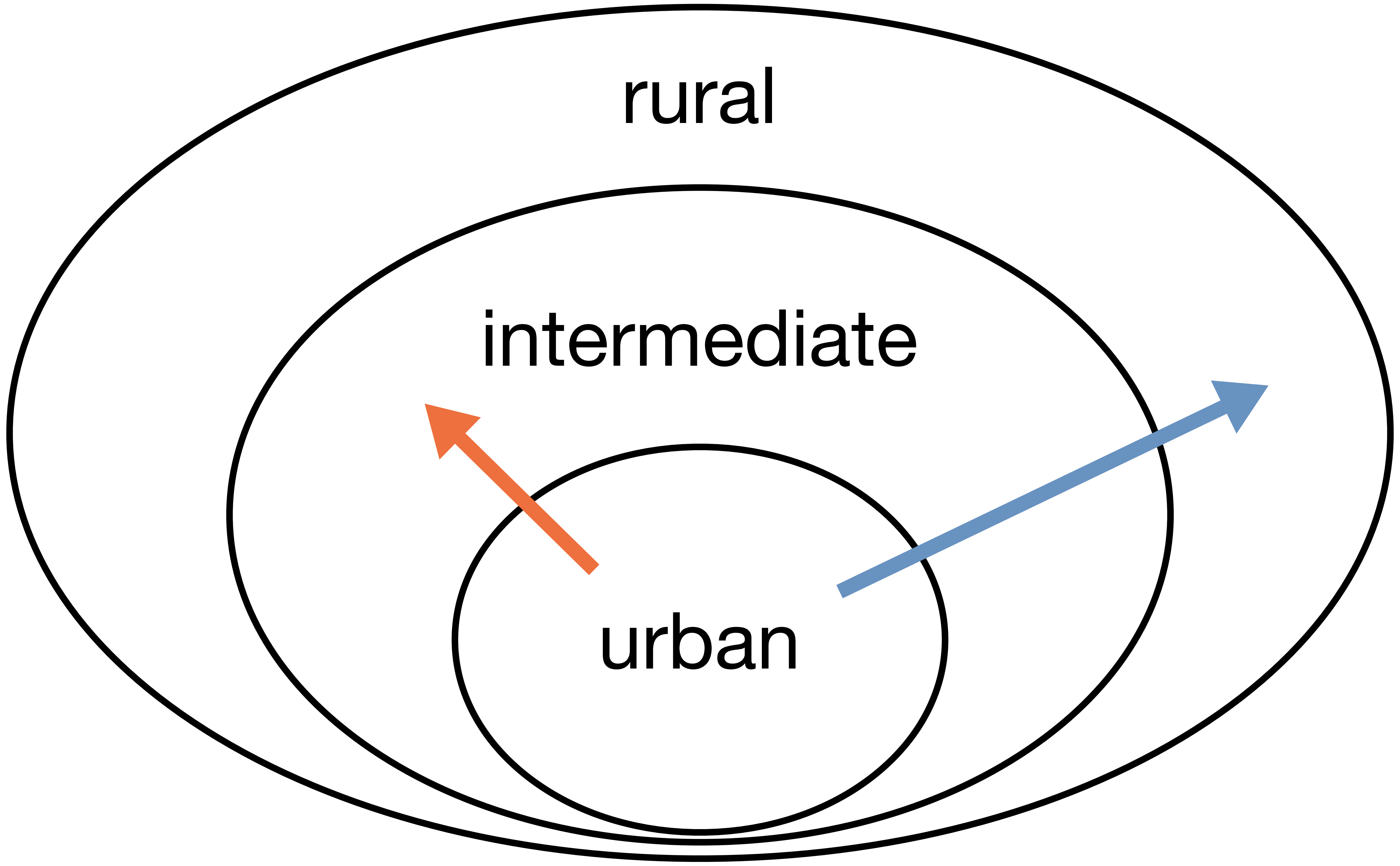
- why is the subsidy effective in this application?

mechanism

- why is the subsidy effective in this application?
- it can (i) target rural areas more precisely, avoiding spillovers to intermediate regions;

mechanism

- why is the subsidy effective in this application?
- it can (i) target rural areas more precisely, avoiding spillovers to intermediate regions;



mechanism

- why is the subsidy effective in this application?
- it can (i) target rural areas more precisely, avoiding spillovers to intermediate regions; (ii) account for preference intensity

mechanism

- why is the subsidy effective in this application?
- it can (i) target rural areas more precisely, avoiding spillovers to intermediate regions; (ii) account for preference intensity

$$\begin{array}{cc} & \begin{array}{cc} \text{urban} & \text{rural} \end{array} \\ \begin{array}{c} i_1 \\ i_2 \end{array} & \begin{bmatrix} 100 & 10 \\ 20 & -10 \end{bmatrix} \end{array}$$

mechanism

- why is the subsidy effective in this application?

- it can (i) target rural areas more precisely, avoiding spillovers to intermediate regions;
  (ii) account for preference intensity

$$\begin{array}{cc} & \begin{array}{cc} \text{urban} & \text{rural} \end{array} \\ \begin{array}{c} i_1 \\ i_2 \end{array} & \begin{bmatrix} 100 & 10 \\ 20 & -10 \end{bmatrix} \end{array}$$

social surplus:

- NC: 100

mechanism

- why is the subsidy effective in this application?

- it can (i) target rural areas more precisely, avoiding spillovers to intermediate regions; (ii) account for preference intensity

$$\begin{array}{r|cc} & \text{urban} & \overset{\geq 1}{\text{rural}} \\ \hline i_1 & 100 & 10 \\ i_2 & 20 & -10 \end{array}$$

social surplus:

- NC: 100

mechanism

- why is the subsidy effective in this application?

- it can (i) target rural areas more precisely, avoiding spillovers to intermediate regions;
  (ii) account for preference intensity

$$\begin{array}{ccc} & & \geq 1 \\ & \text{urban} & \text{rural} \\ i_1 & \cancel{100} & 10 \\ i_2 & \cancel{20} & -10 \end{array}$$

social surplus:

- NC: 100
- AC: 10

mechanism

- why is the subsidy effective in this application?

- it can (i) target rural areas more precisely, avoiding spillovers to intermediate regions; (ii) account for preference intensity

$$\begin{array}{c|cc|} & & \geq 1 \\ & \text{urban} & \text{rural} \\ i_1 & 100 & 20 \\ i_2 & 20 & 0 \\ & & +10 \end{array}$$

social surplus:

- NC: 100
- AC: 10
- OS: 90 (subsidy $w_{\text{rural}} = 10$)

mechanism

- why is the subsidy effective in this application?

- it can (i) target rural areas more precisely, avoiding spillovers to intermediate regions;
  (ii) account for preference intensity

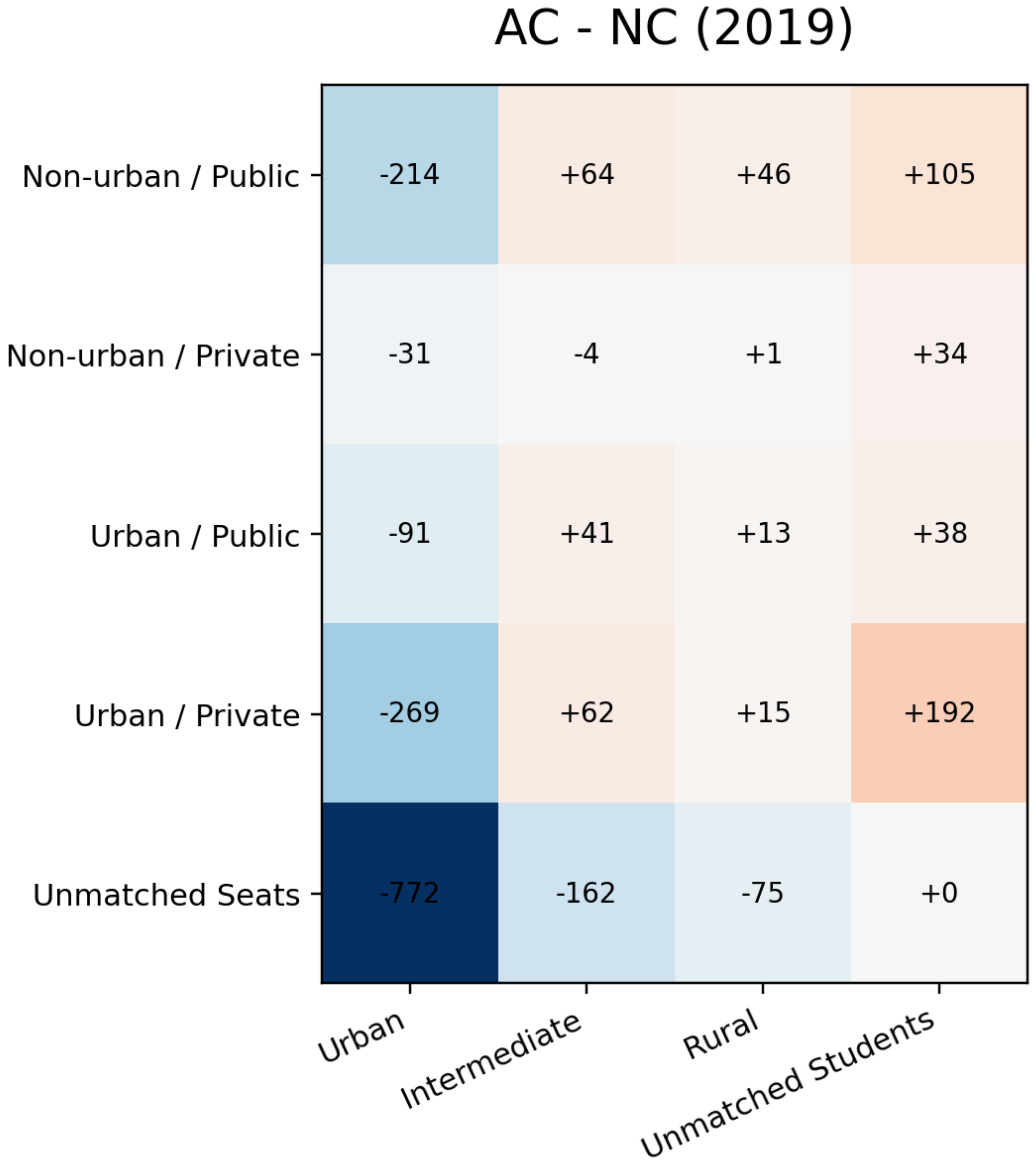
$$\begin{array}{cc} & \text{urban} \quad \text{rural} \\ \begin{array}{c} i_1 \\ i_2 \end{array} & \begin{bmatrix} 100 & 10 \\ 20 & -10 \end{bmatrix} \end{array}$$

social surplus:

- NC: 100
- AC: 10
- OS: 90 (subsidy $w_{\text{rural}} = 10$)

- a targeted subsidy can affect marginal participants with little welfare losses

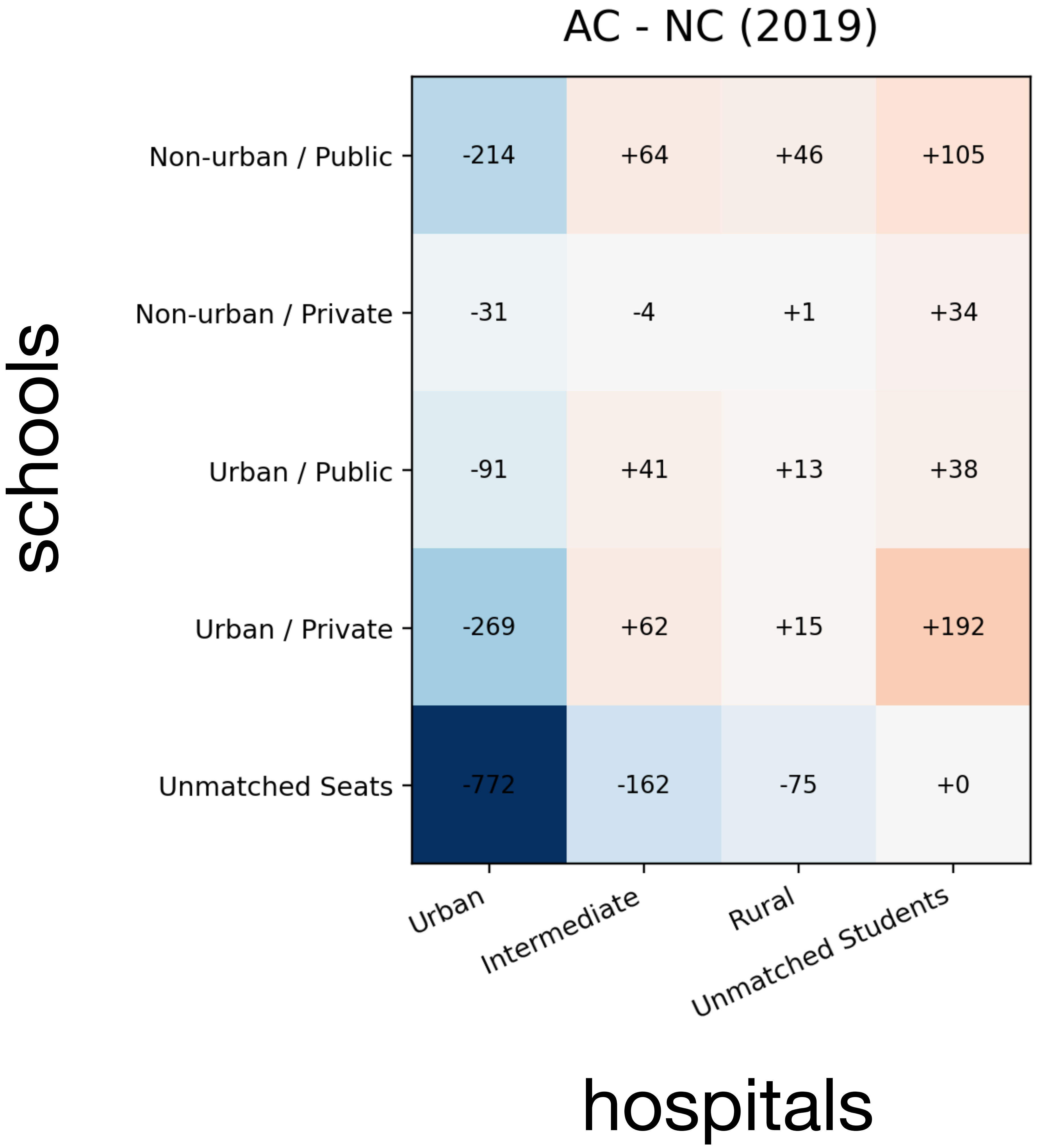
change in matching patterns



AC - NC (2019)

	Urban	Intermediate	Rural	Unmatched Students
Non-urban / Public	-214	+64	+46	+105
Non-urban / Private	-31	-4	+1	+34
Urban / Public	-91	+41	+13	+38
Urban / Private	-269	+62	+15	+192
Unmatched Seats	-772	-162	-75	+0

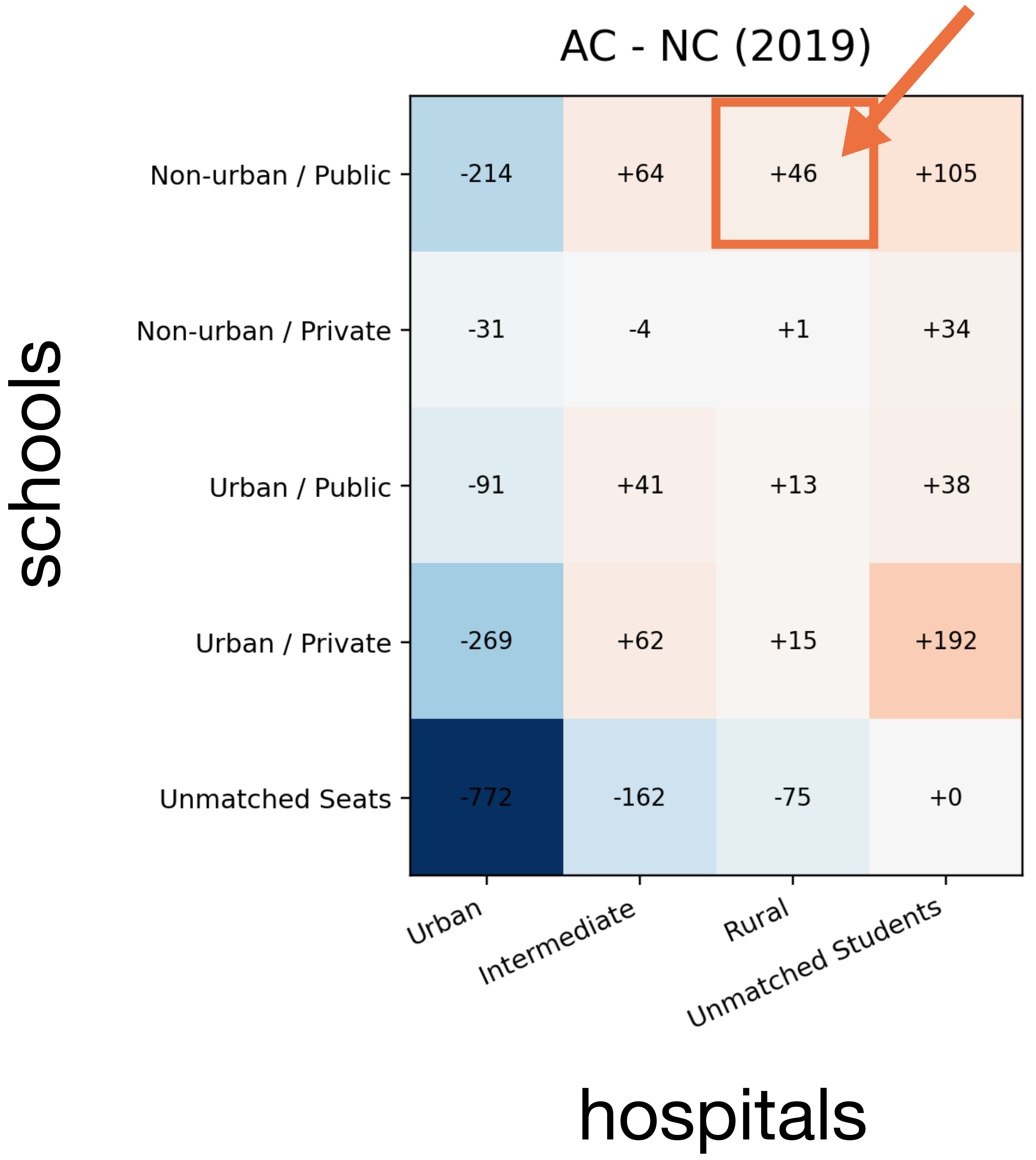
change in matching patterns



AC - NC (2019)

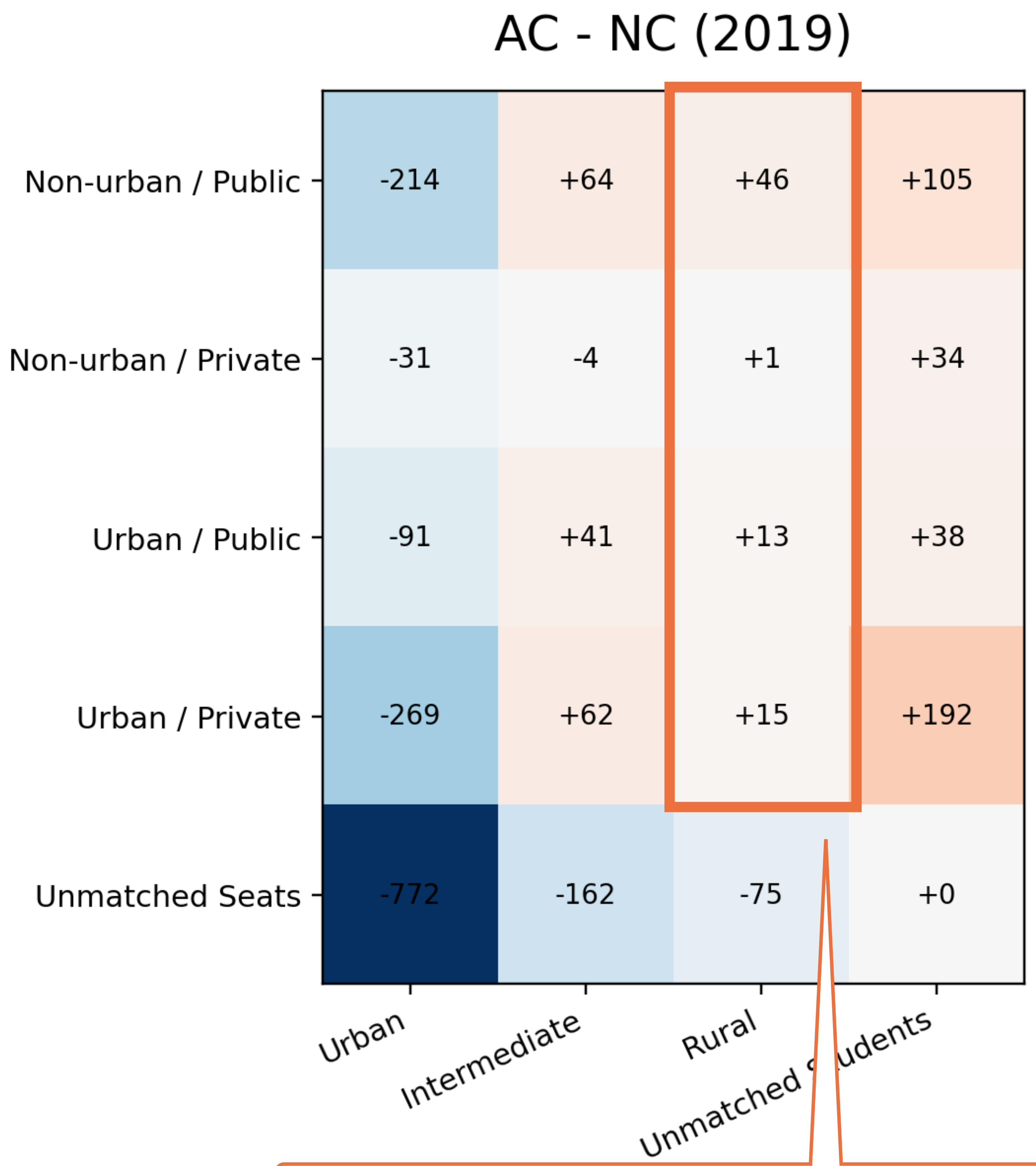
	Urban	Intermediate	Rural	Unmatched Students
Non-urban / Public	-214	+64	+46	+105
Non-urban / Private	-31	-4	+1	+34
Urban / Public	-91	+41	+13	+38
Urban / Private	-269	+62	+15	+192
Unmatched Seats	-772	-162	-75	+0

change in matching patterns



	Urban	Intermediate	Rural	Unmatched Students
Non-urban / Public	-214	+64	+46	+105
Non-urban / Private	-31	-4	+1	+34
Urban / Public	-91	+41	+13	+38
Urban / Private	-269	+62	+15	+192
Unmatched Seats	-772	-162	-75	+0

change in matching patterns

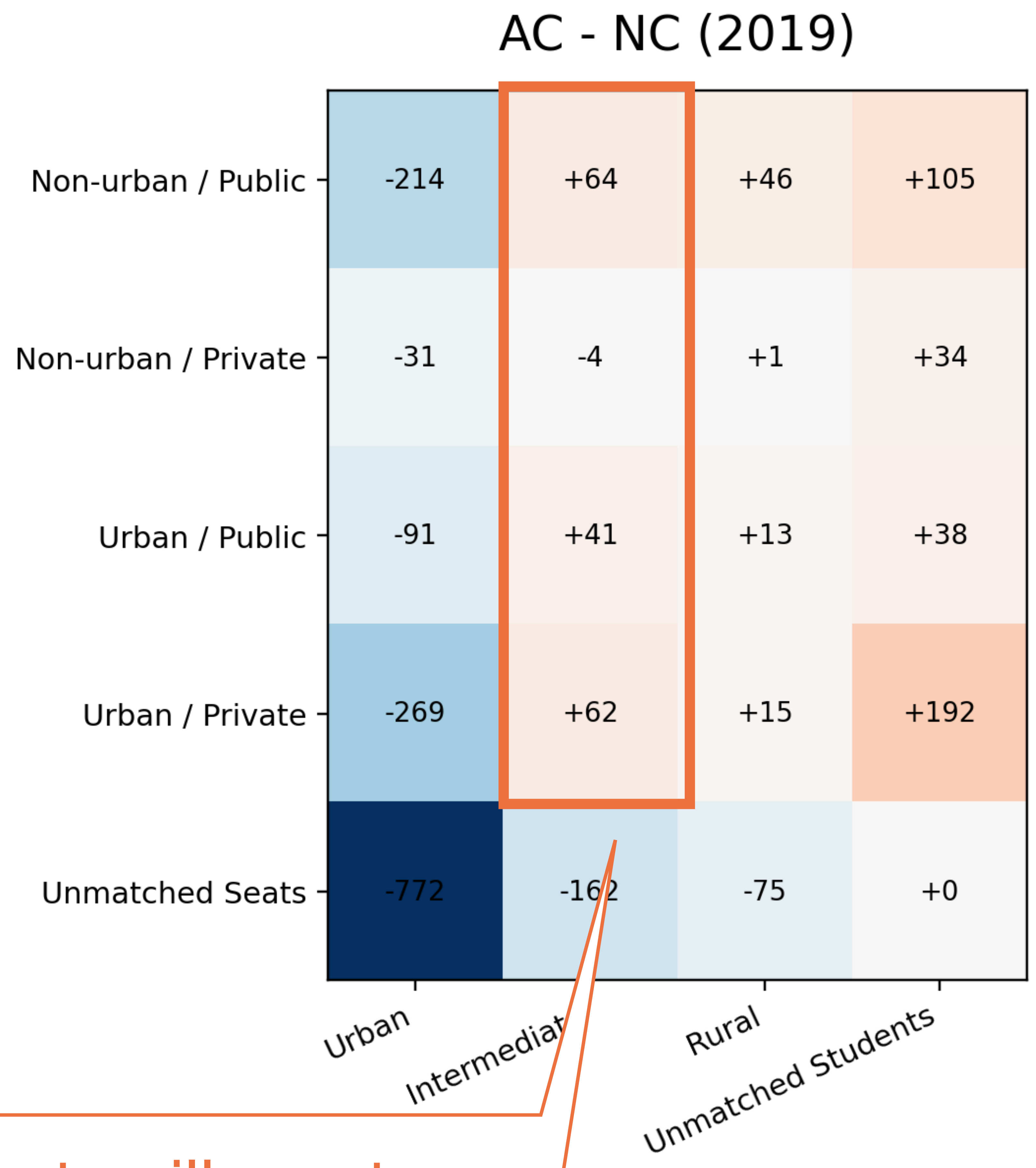


AC - NC (2019)

	Urban	Intermediate	Rural	Unmatched students
Non-urban / Public	-214	+64	+46	+105
Non-urban / Private	-31	-4	+1	+34
Urban / Public	-91	+41	+13	+38
Urban / Private	-269	+62	+15	+192
Unmatched Seats	-772	-162	-75	+0

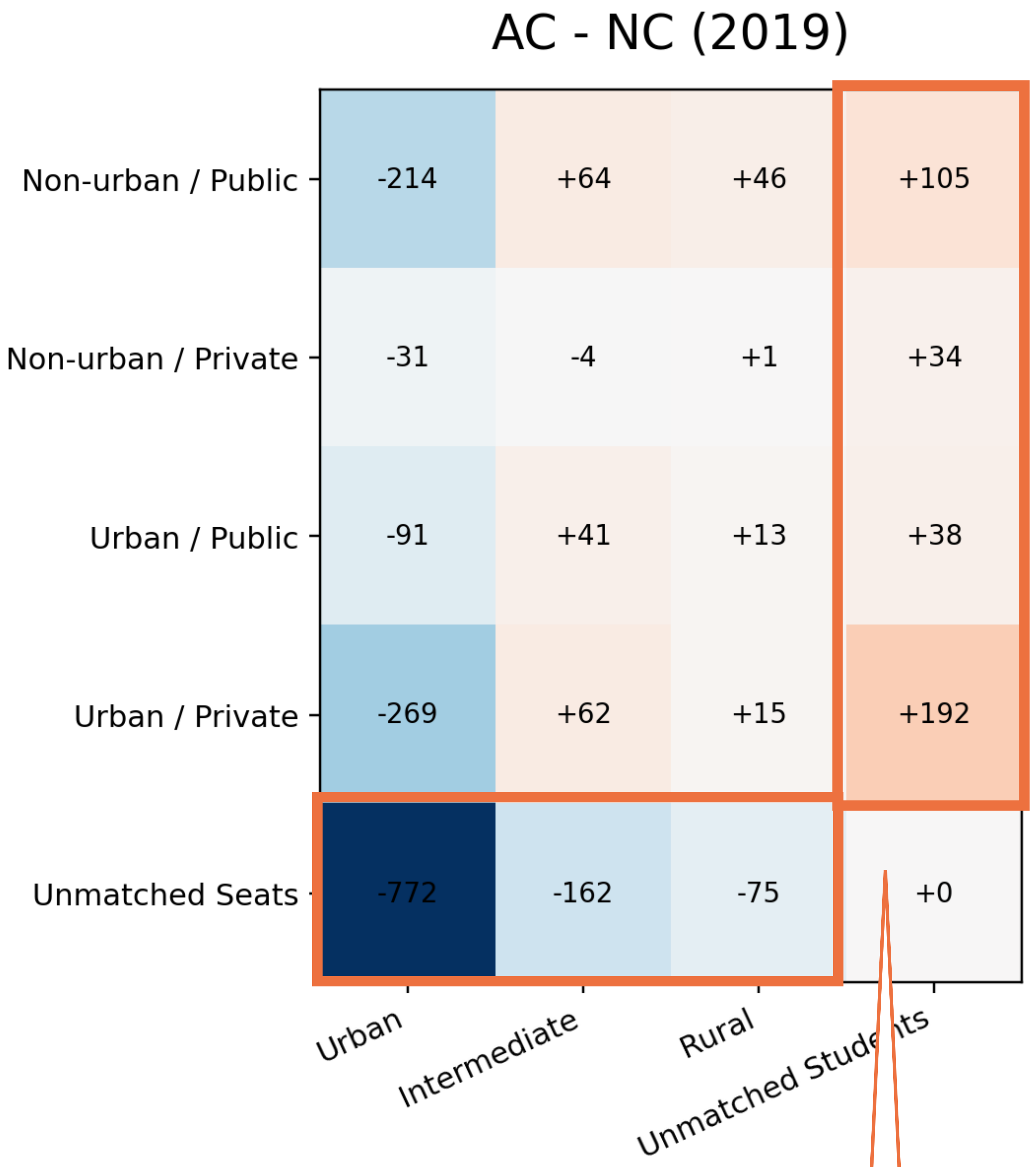
matches in rural areas certainly increases

change in matching patterns



significant spillover to intermediate regions

change in matching patterns

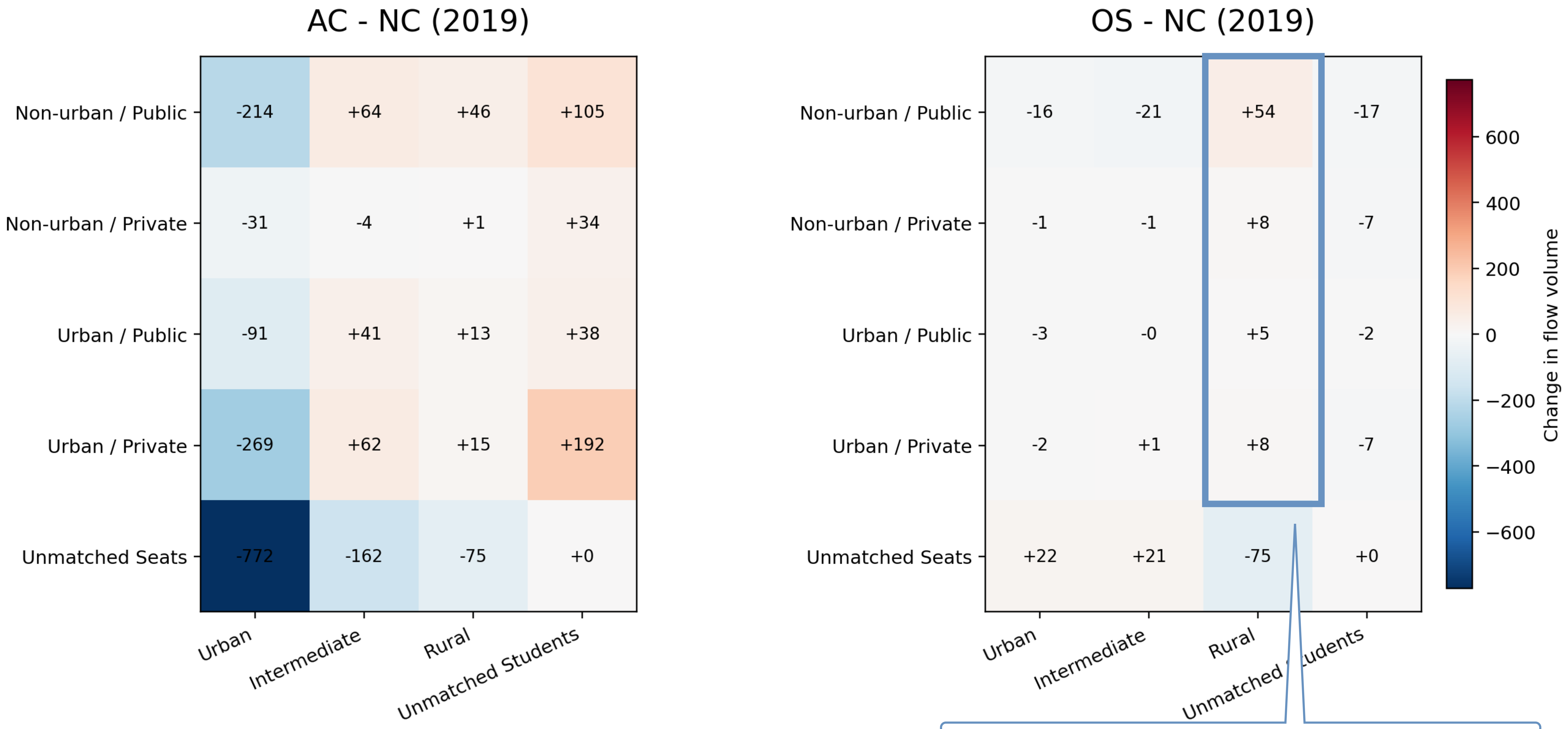


AC - NC (2019)

	Urban	Intermediate	Rural	Unmatched Students
Non-urban / Public	-214	+64	+46	+105
Non-urban / Private	-31	-4	+1	+34
Urban / Public	-91	+41	+13	+38
Urban / Private	-269	+62	+15	+192
Unmatched Seats	-772	-162	-75	+0

# of unmatches increases

change in matching patterns



AC - NC (2019)

	Urban	Intermediate	Rural	Unmatched Students
Non-urban / Public	-214	+64	+46	+105
Non-urban / Private	-31	-4	+1	+34
Urban / Public	-91	+41	+13	+38
Urban / Private	-269	+62	+15	+192
Unmatched Seats	-772	-162	-75	+0

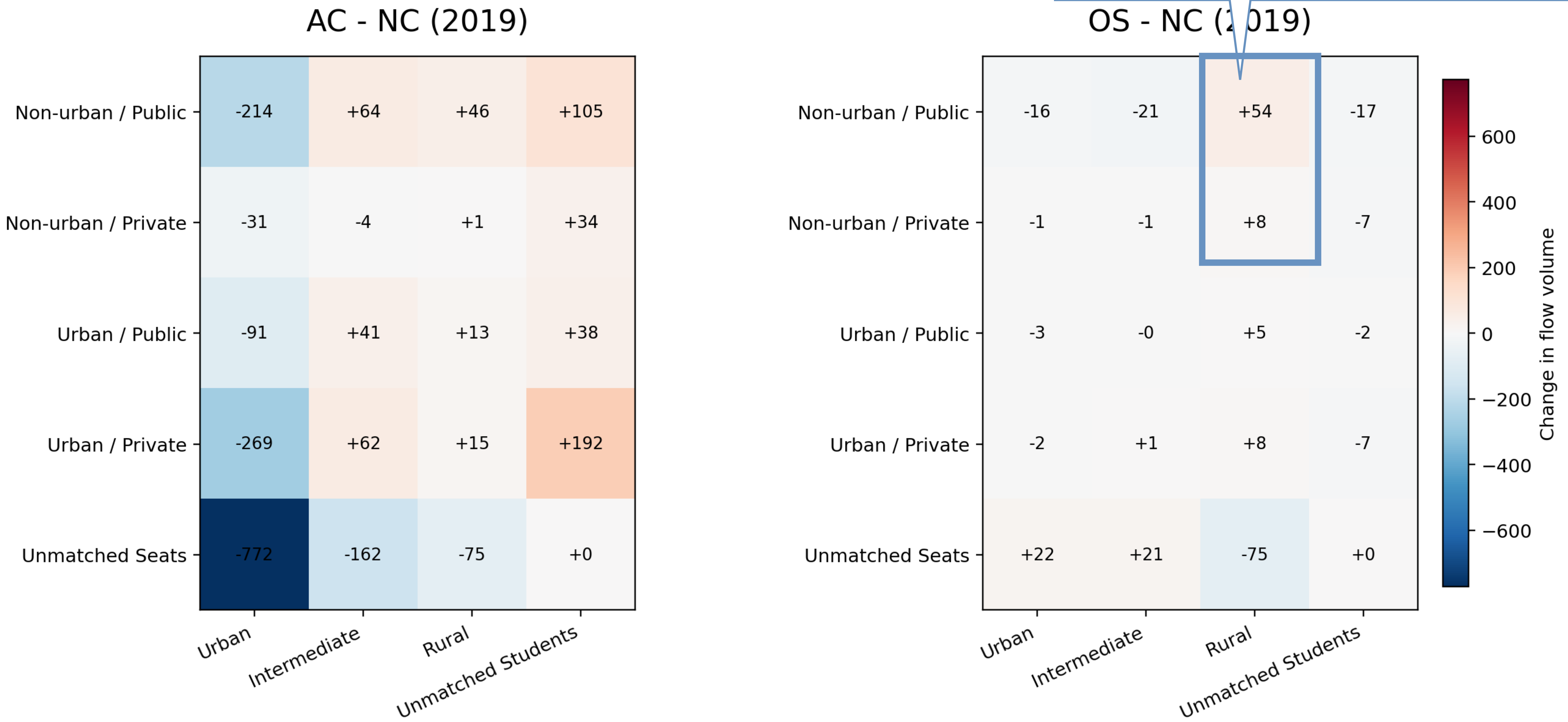
OS - NC (2019)

	Urban	Intermediate	Rural	Unmatched Students
Non-urban / Public	-16	-21	+54	-17
Non-urban / Private	-1	-1	+8	-7
Urban / Public	-3	-0	+5	-2
Urban / Private	-2	+1	+8	-7
Unmatched Seats	+22	+21	-75	+0

precisely target rural areas

change in matching patterns

target students from nearby areas who are less costly to incentivize



AC - NC (2019)

	Urban	Intermediate	Rural	Unmatched Students
Non-urban / Public	-214	+64	+46	+105
Non-urban / Private	-31	-4	+1	+34
Urban / Public	-91	+41	+13	+38
Urban / Private	-269	+62	+15	+192
Unmatched Seats	-772	-162	-75	+0

OS - NC (2019)

	Urban	Intermediate	Rural	Unmatched Students
Non-urban / Public	-16	-21	+54	-17
Non-urban / Private	-1	-1	+8	-7
Urban / Public	-3	-0	+5	-2
Urban / Private	-2	+1	+8	-7
Unmatched Seats	+22	+21	-75	+0

conclusion

- we develop a framework to evaluate the efficiency of policies in matching markets with distributional constraints
  - optimal taxation policy can be approximately implemented from observed data

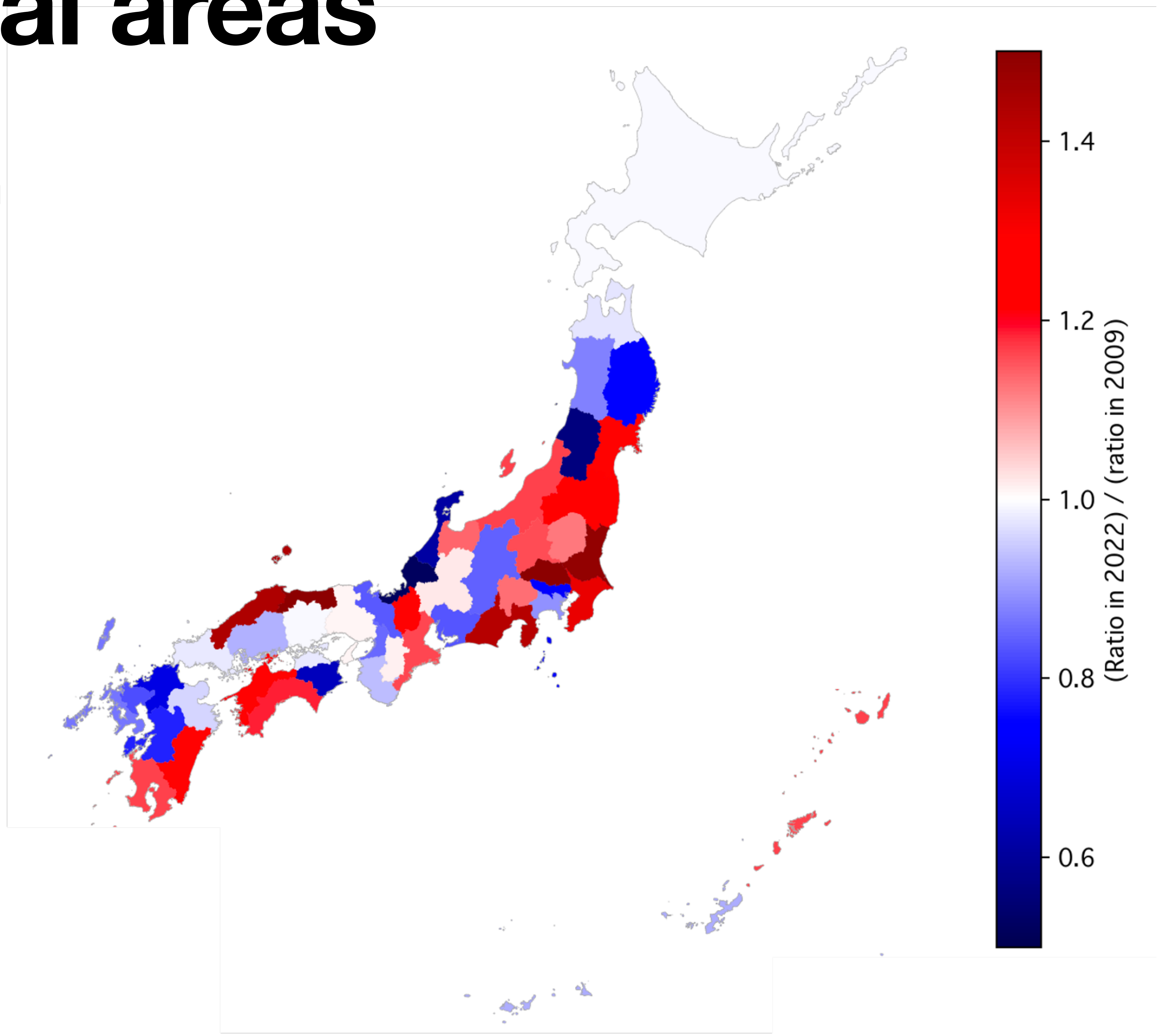
- we apply the framework to **JRMP** data:
  - current cap-based policy generates a significant welfare loss
    - caps are blunt instruments that fail to account for preferences intensities
  - moderate subsidy can address distributional imbalances, improving social welfare

thank you ☺

questions or comments?

unequal impact on rural areas

- cap-based policy has an unequal impact on rural prefectures
- change in share of national matches by prefecture from 2009 to 2022
- the share increases in
  - prefectures around Tokyo
  - some rural areas
- the share decreases in
  - the 6 urban prefectures
  - **some rural areas**
    (north Tohoku, Hokuriku, north Kyushu)



IVs

- number of university hospitals within a 10km radius
- number of government hospitals within a 10km radius
- ave. number of beds in hospitals within a 10km radius
- ave. dist. to the medical schools of students matched under AC, for hospitals within a 10km radius
- ave. number of past matches with those schools under AC, for hospitals within a 10km radius
- ave. number of affiliations with those schools under AC, for hospitals within a 10km radius

- When setting salaries, hospitals can take into account observable local conditions like those listed above.

- In contrast, the non-salary portion of the transfer (the unobserved component, program-specific features) is fixed and cannot be adjusted.

comments on TU stability assumption

- in our paper, we provide a justification for its use for the JRMP market
    - consider a game that mimics the JRMP process
        - positions first propose wages
        - doctors and positions submit rank-ordered lists
        - DA algorithm determines the final match
    - under certain conditions,
        - any pure NE in this game induces a stable outcome in the TU model
        - any stable outcome in the TU model can be supported as a pure NE

institutional background

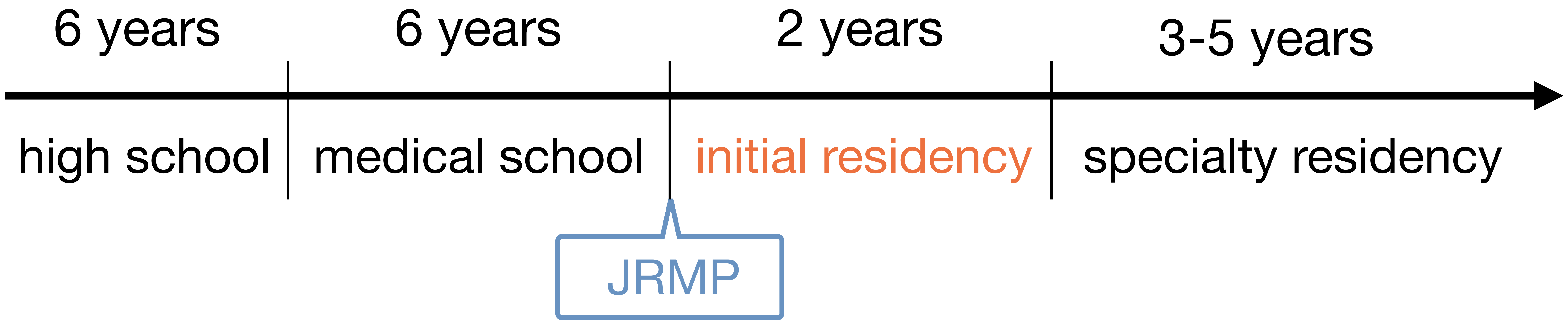
model

theoretical results

estimation

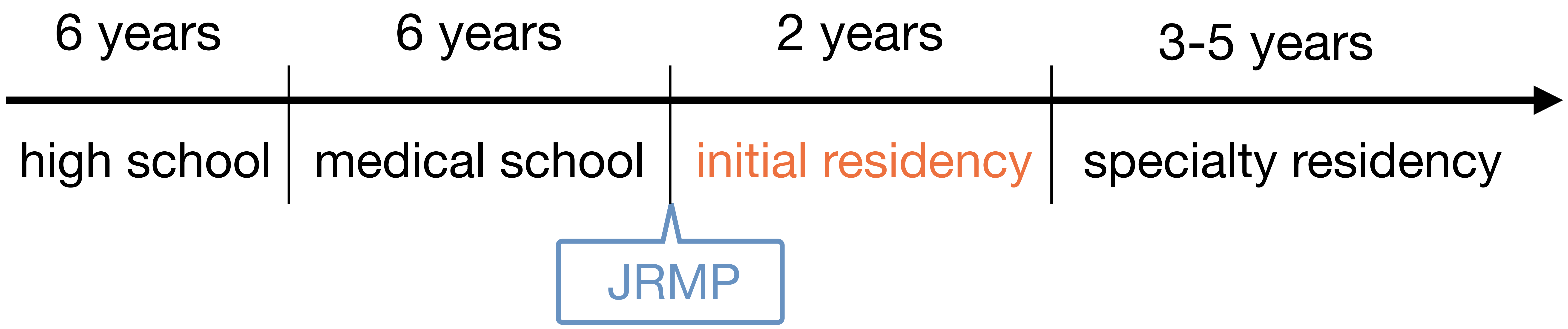
simulation

medical training in Japan



- in Japan, students enter a six-year medical program directly from high school
  - admission is highly competitive, based on entrance examinations

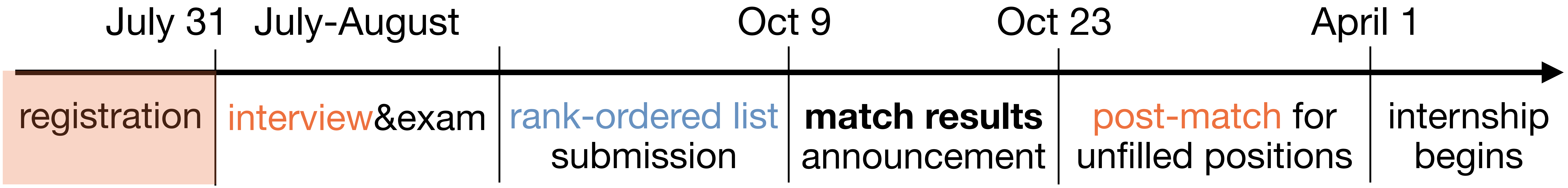
medical training in Japan



- in Japan, students enter a six-year medical program directly from high school
  - admission is highly competitive, based on entrance examinations

- two-year initial residency training starts immediately after graduation
  - it became mandatory in **2004**, when medical training in Japan was reformed
  - as part of the reform, JRMP was introduced to centralize the matching process

JRMP timeline



July 31

July-August

Oct 9

Oct 23

April 1

registration

interview&exam

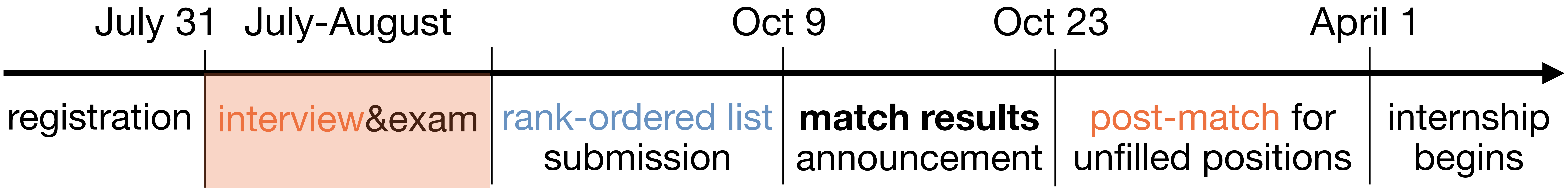
rank-ordered list submission

match results announcement

post-match for unfilled positions

internship begins

JRMP timeline

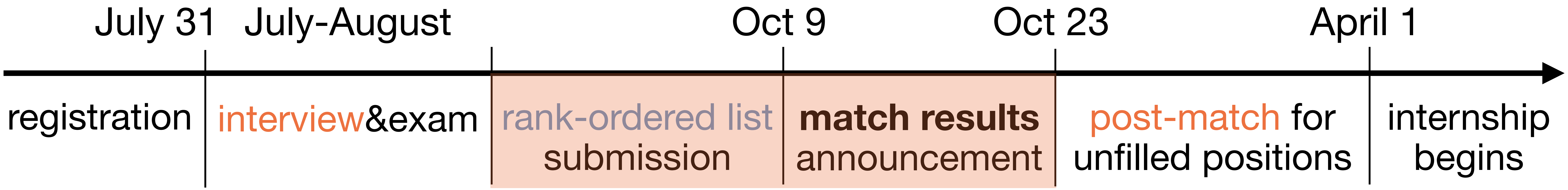


July 31 July-August Oct 9 Oct 23 April 1

registration | interview&exam | rank-ordered list submission | **match results** announcement | post-match for unfilled positions | internship begins

- students must complete interviews with the programs they wish to include in their preference list
    - median length of the list is 4 in 2024

JRMP timeline



July 31

July-August

Oct 9

Oct 23

April 1

registration

interview&exam

rank-ordered list submission

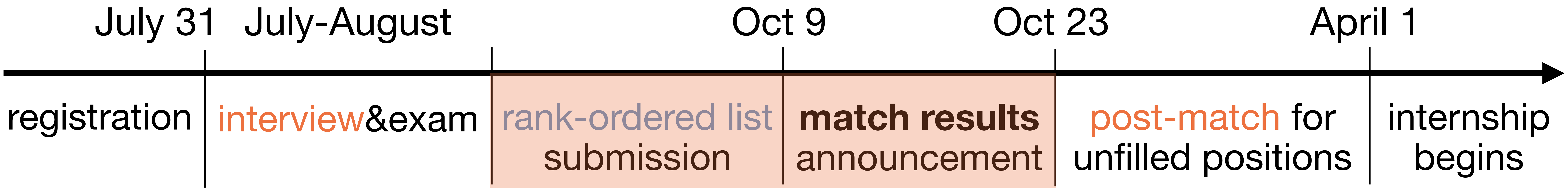
match results announcement

post-match for unfilled positions

internship begins

- students must complete interviews with the programs they wish to include in their preference list
  - median length of the list is 4 in 2024
- students and programs submit their rank-ordered lists to JRMP
- **DA algorithm** determines the matching using the submitted lists

JRMP timeline



July 31 | July-August | Oct 9 | Oct 23 | April 1

registration | interview&exam | rank-ordered list submission | **match results** announcement | post-match for unfilled positions | internship begins

- students must complete interviews with the programs they wish to include in their preference list
  - median length of the list is 4 in 2024

- students and programs submit their rank-ordered lists to JRMP
- **DA algorithm** determines the matching using the submitted lists

- medical students evaluate programs based on a range of factors
  - e.g., location, salary, university affiliation, quality of training, work-life balance, etc.

debates on JRMP

- following the introduction of JRMP, researchers and Japanese media have reported that the distributional disparity of physicians has worsened
  - **before**: a majority of students did their residency at the **univ. hospital** affiliated with their schools
  - **after**: with the centralized matching system, medical students can freely express their preferences, leading to a concentration of applicants in non-university and/or urban programs
- JRMP introduced regional caps starting in **2010**
  - [total # of residency positions] := [# students] × [constant]
  - positions are allocated across prefectures, reducing urban positions disproportionately
  - **rationale**: encourage students to pursue training in rural areas, expecting them to remain there

decreasing over time
(**1.22** in 2015 → **1.05** by 2025)

evidence of endogenous transfer

- there is a large variation in salary among different programs

- rural hospitals offer higher salaries to attract students

- these suggest that hospitals are **strategically choose salaries in response to the market conditions**

- difficult to capture within the NTU framework

	min (USD/month)	**max (USD/month)**	**std/mean**
Japan (2017-2019)	1241	5896	0.26
U.S. (2010)	2583	5147	0.06

	mean (USD/month)	**std (USD/month)**
overall	2665	686
rural hospitals	2813 ∨	679
urban hospitals	2381	603