UTMD Working Paper
The University of Tokyo
Market Design Center

UTMD-031

Regulating Matching Markets with Constraints: Data-driven Taxation

Akira Matsushita
University of Tokyo

Kei Ikegami
New York University

Kyohei Okumura
Northwestern University

Yoji Tomita
CyberAgent, Inc.

Atsushi Iwasaki
University of Electro-Communications

May 28, 2022

Regulating Matching Markets with Constraints: Data-driven Taxation

Akira Matsushita
University of Tokyo
matsushita@ants-tears.com

Kei Ikegami
New York University
ki2047@nyu.edu

Kyohei Okumura
Northwestern University
kyohei.okumura@gmail.com

Yoji Tomita
CyberAgent, Inc.
tomita_yoji@cyberagent.co.jp

Atsushi Iwasaki
University of Electro-Communications
atsushi.iwasaki@uec.ac.jp

Abstract

This paper develops a framework to conduct a counterfactual analysis to regulate matching markets with regional constraints that impose lower and upper bounds on the number of matches in each region. Our work is motivated by the Japan Residency Matching Program, in which the policymaker wants to guarantee the least number of doctors working in rural regions to achieve the minimum standard of service. Among the multiple possible policies that satisfy such constraints, a policymaker wants to choose the best. To this end, we develop a discrete choice model approach that estimates utility functions of agents from observed data and predicts agents' behavior under different counterfactual policies. Our framework also allows the policymaker to design the welfare-maximizing tax scheme, which outperforms the policy currently used in practice. Furthermore, a numerical experiment illustrates how our method works.

1 Introduction

Matching with constraints, initiated by Kamada and Kojima [16], has been recently paid considerable attention to across economics, computer science, and AI [1, 5, 8, 9, 18, 15, 13, 19]. For example, *diversity constraints* matter in the context of school choice: each school is required to balance its composition of students, typically in terms of socioeconomic status. Recent developments in matching theory enable policymakers to search for the matching outcome under several constraints on the volume of matches.

Preprint. Under review.

This paper develops a framework to design a policy that regulates matching markets with regional constraints [13] that impose lower and upper bounds on the number of matches in each region. The policymaker (hereafter, PM) needs to satisfy such constraints required by society. For example, the government may need to guarantee the minimum number of doctors working in each rural region to achieve the minimum standard of service. In general, there are multiple possible policies to satisfy the constraints but only some of them are used in practice: the Japanese government sets the upper bound on the number of workers working in urban areas so that it guarantees the number of doctors in rural areas. However, it is often unclear which policy is better than the others and there is a need for methods to evaluate and compare different policies.

A seminal work by Choo and Siow [7] initiated the estimation method for the two-sided matching model, i.e., one-to-one marriage markets. Each agent's utility is assumed to be transferable ([22, 4, 17]), where each agent can freely transfer their utility to their partner via monetary transfer. Galichon and Salanié [11] shows that the stable matching coincides with the market clearing allocation in which the agents solve the discrete choice problems, choosing an alternative from the finite set so that he maximizes its utility. They also propose a method to estimate utility functions that are consistent with the observed matching patterns under the *unobserved heterogeneity*, which captures the different behaviors of the agents that have the same covariates and look the same to the PM.

Given the utility functions estimated via the method of Galichon and Salanié [11], our framework allows the PM to predict matching patterns under different counterfactual constraints. Moreover, if the PM can use taxes and subsidies to incentivize agents in the market to satisfy regional constraints, our approach provides a way of computing the tax scheme that maximizes social surplus under the constraints. We define *efficient aggregate equilibrium (EAE)*, which is the welfare maximizing equilibrium under the constraints achievable by taxes and subsidies. We characterize EAE as a solution to a convex optimization problem and show that the solution always exists and is unique. This result enables us to design the welfare-maximizing tax scheme for any arbitrary regional constraints.

Finally, we conduct a numerical experiment to illustrate how our approach works. We imitate the problem that the Japan Residency Matching Program (JRMP) deals with: its policy goal is to mitigate the popularity gap among urban and rural hospitals [16]. To this end, JRMP restricts the maximum number of residents matched within each prefecture and proportionally decreases the capacity of hospitals in each prefecture to meet the bound (*cap-reduced AE* in Section 6). The simulation demonstrates that our method outperforms such a policy used in practice in terms of social welfare. In addition, we show that even when the PM cannot run deficits and needs to balance the budget, we can achieve higher social welfare than the cap-reduced policy.

2 Equilibrium in Matching Market

In this section, we set up the model of transferable utility matching market with regional constraints and define a basic equilibrium concept, called *individual equilibrium*. As we will see later, this equilibrium concept itself is not helpful for the policymaker. However, this will be a building block of a key equilibrium concept we develop in Section 3, which helps the policymaker design the tax scheme using data.

We consider the two-sided decentralized matching market with regional constraints. There are two groups of agents: let I be the set of workers, and let J be the set of job slots. We assume I and J are finite for the moment, but note that we will make a large market approximation in Section 3. Each worker $i \in I$ can be matched with at most one job slot $j \in J$. If i is unmatched, we say i is matched with an outside option j_0. Worker i obtains payoff $u_{ij} \in \mathbb{R}$ when i is matched with $j \in J_0 := J \cup \{j_0\}$. Similarly, job slot $j \in J$ can be assigned to worker $i \in I$ or an outside option i_0. Slot j obtains payoff $v_{ij} \in \mathbb{R}$ when j is matched with $i \in I_0 := I \cup \{i_0\}$. We normalize the unmatch payoffs to be zero: $u_{ij_0} = v_{i_0 j} = 0$ for all $i \in I$ and $j \in J$.

Let Z be a set of finite regions, $Z := \{z_0, z_1, z_2, \dots, z_L\}$ for some $L \in \mathbb{N}$. Each slot $j \in J$ belongs to a *region* $z(j) \in Z$. For convenience, we assume an outside option for workers, say j_0, is in region z_0. With a slight abuse of notation, we write $j \in z$ if $z(j) = z \in Z$.

Each region $z \in Z$ has an upper bound quota $\bar{o}_z \in \mathbb{R}_+$ and a lower bound quota $\underline{o}_z \in \mathbb{R}_+$ ($\bar{o}_z \geq \underline{o}_z \geq 0$ and $\bar{o}_z \neq 0$). The number of workers in region z must be at least $\underline{o}_z$ and at most $\bar{o}_z$. If region z has no upper bound quota, we set $\bar{o}_z = +\infty$. Similarly, if z does not have lower bound, we set $\underline{o}_z = 0$. We assume that there is no restriction on the outside options: $\bar{o}_{z_0} = +\infty$ and $\underline{o}_{z_0} = 0$.

To satisfy the regional constraints, the policymaker taxes the pairs in excessively popular regions while it subsidizes the pairs in unpopular regions. The *tax* on region z is denoted by w_z and all pair in region z pays w_z to the policymaker (NB: a negative tax can be interpreted as a subsidy).[1]

The payoffs are *transferable* between the matched worker and slot in the following sense. The matched pair (i, j) generates the joint surplus $\Phi_{ij} \in \mathbb{R}$ measured by money, say dollars. Let the utility of i when matched with j be u_{ij}; let the utility of j when matched with i be v_{ij}. If $i \in I$ and $j \in J \in z$ match, they divide the net joint surplus, i.e., $\tilde{\Phi}_{ij} \coloneqq \Phi_{ij} - w_z = u_{ij} + v_{ij}$[2]. The values u_{ij} and v_{ij}, or how they divide the net joint surplus, are determined possibly for social reasons (e.g., bargaining power) and treated as given by the agents. Note that each slot is possessed by some entity such as a firm, so the slot also has a preference over workers.

A matching represents who is matched with whom and is expressed as 0-1 vector $d = (d_{ij})_{ij}$ such that $d_{ij} = 1$ iff i and j are matched for $(i, j) \in I_0 \times J_0 \setminus \{(i_0, j_0)\}$. A matching d is *feasible* if

- (population constraint) each worker $i \in I$ satisfies $\sum_{j \in J_0} d_{ij} = 1$; each slot $j \in J$ satisfies $\sum_{i \in I_0} d_{ij} = 1$, and
- (regional constraint) each region $z \in Z$ satisfies $\sum_{i \in I} \sum_{j \in z} d_{ij} \in [\underline{o}_z, \bar{o}_z]$.

In the following (sub)sections, we characterize the equilibrium matching among the feasible matchings. We assume that all agents know the joint surpluses and how they are divided, i.e., $(\Phi_{ij}, u_{ij}, v_{ij})_{i,j}$, and tax scheme $(w_z)_z$ for all regions.

We now define an *individual equilibrium*. Below u_i and v_j are interpreted as i's and j's payoffs in the equilibrium, respectively. For the reason why this equilibrium concept is valid, see Appendix B.

Definition 1 (Individual Equilibrium)**.** A profile $(d, (u, v), w)$ of feasible matching d, equilibrium payoffs of agents $(u, v) \in \mathbb{R}^{|I|} \times \mathbb{R}^{|J|}$, and taxes $w \in \mathbb{R}^L$ is *stable* if

- (Individual rationality) For all $i \in I$, $u_i \geq \tilde{\Phi}_{ij_0} = 0$, with equality if i is unmatched in d; for all $j \in J$, $v_j \geq \tilde{\Phi}_{i_0 j} = 0$, with equality if j is unmatched in d.
- (No blocking pairs) For all i and j, $u_i + v_j \geq \tilde{\Phi}_{ij}$, with equality if i and j are matched in d.

The individual equilibrium is a solution to the social welfare maximization problem exactly as in [22]. However, in practice, the agents outside of the economy, such as the policymaker or the researcher, do not have access to the data on the exact individual preferences. Instead, they rely on coarser information, such as population characteristics, to determine the taxes. In the subsequent sections, we introduce a concept of *aggregate equilibrium*, based on individual equilibrium, to handle such a situation and enable the policymaker to compute the optimal tax scheme.

3 Aggregate Equilibrium with Unobserved Heterogeneity

3.1 Unobserved Heterogeneity and Separability

Let $X \coloneqq \{x_1, x_2, \ldots, x_N\}$ be the finite set of observable worker types. Each worker $i \in I$ has a type $x(i) \in X$. Similarly, let $Y \coloneqq \{y_1, y_2, \ldots, y_M\}$ be the finite set of observable job slot types. Each slot $j \in J$ has a type $y(j) \in Y$. With a slight abuse of notation, $x(i) = x$ (resp. $y(j) = y$) is denoted by $i \in x$ (resp. $j \in y$). We define x_0 and y_0 as "null types" that are the types of outside options i_0 and j_0. Finally let $X_0 \coloneqq X \cup \{x_0\}$ and $Y_0 \coloneqq Y \cup \{y_0\}$ be the sets of all worker and slot types including the null types, and define $T \coloneqq X_0 \times Y_0 \setminus \{(x_0, y_0)\}$ as the set of all type pairs.

The policymaker can observe these types only; they cannot distinguish same type agents/slots. There is *unobserved heterogeneity* in a sense that even when two agents i and i' have the same type x and look the same to the policymaker, their actual preferences can be different.

[1]The policymaker may design taxes for each pair. Although we restrict the class of tax schemes here, such a restriction is harmless regarding welfare maximization. See Corollary 2 in Section 4.

[2]Technically we define the joint surplus $\tilde{\Phi}_{ij_0} = \tilde{\Phi}_{i_0 j} = 0$ when unmatched.

Type $y \in Y$ and region $z \in Z$ can be interpreted in various manners. One may think of a type as a firm/hospital and region z as a unit of districts. It is also possible that a type is a minor subcategory of occupation (e.g., registered/licensed practical nurse, physician assistant, medical doctor) and a region is a larger category of occupation (e.g., medical jobs). Throughout this paper, we interpret a type as a firm, and a region as a unit of districts for simplicity.

Assumption 1 (Regional Constraint)**.** Each type $y \in Y$ can belong to only one region[3]. Denote the region that type y belongs to by $z(y) \in Z$.

Let us define $n_x \coloneqq \#\{i\colon x(i) = x\}$ the number of type-x agents and $m_y \coloneqq \#\{j\colon y(j) = y\}$ the number of type-y slots. Hereafter we make the large market approximation:

Assumption 2 (Large Market Approximation)**.** Each type x has n_x mass of continuum of agents rather than distinct n_x agents. Similarly we assume there are m_y mass of type y job slots.

In this setting, the matching is defined as the measure of matches for each pair of type (x, y); let $\mu = (\mu_{xy})_{x,y} \in \mathbb{R}_+^{|T|}$. And its feasibility is defined as follows;

Definition 2 (Feasible matching)**.** $\mu = (\mu_{xy})_{xy} \in \mathbb{R}_+^{|T|}$ is a *feasible matching* if it satisfies

- (population constraint) each worker type $x \in X$ satisfies $\sum_{y \in Y_0} \mu_{xy} = n_x$; each slot type $y \in Y$ satisfies $\sum_{x \in X_0} \mu_{xy} = m_y$, and
- (regional constraint) each region $z \in Z$ satisfies $\sum_{x \in X} \sum_{y \in z} \mu_{xy} \in [\underline{o}_z, \bar{o}_z]$.

Regarding the joint surplus $\tilde{\Phi}_{ij}$, we impose two structures on it. First, for some i.i.d. random variable ξ_{ij}, $\tilde{\Phi}_{ij} - \xi_{ij}$ does not depend on the pair of individuals (i, j), but only depend on their types $x(i)$ and $y(j)$. Second, for each i and j such that $x(i) = x$ and $y(j) = y$, ξ_{ij} can be represented as a sum of two i.i.d. random variables ϵ_{iy} and η_{xj}. We call these properties of the joint surplus as *additive separability*. The discussion above can be summarized as follows:

Assumption 3 (Independence)**.** For each x and $i \in x$, error term ϵ_{iy} is drawn from the distribution P_x. Similarly, for each y and $j \in y$, error term η_{xj} is drawn from the distribution Q_y. The error terms are independent across all i's and j's.

Assumption 4 (Additive Separability)**.** For each x and y, $\tilde{\Phi}_{ij} - \epsilon_{iy} - \eta_{xj}$ is constant over the individuals $i \in x$ and $j \in y$, and denoted by Φ_{xy}.

We also impose the following technical assumption on the error terms ϵ_{iy}'s and η_{xj}'s:

Assumption 5 (Smooth Distibution with Full-Support)**.** For each x and y, the cdf's P_x and Q_y are continuously differentiable. Moreover, $\mathrm{supp}(P_x) = \mathbb{R}^{N+1}$ and $\mathrm{supp}(Q_y) = \mathbb{R}^{M+1}$.

Although these assumptions 2-5 are standard in the discrete choice literature, we will discuss the necessity of them later in Appendix B.

Example 1. Consider a job matching market between two types of workers and three types of jobs which are divided into two categories (z_1, z_2). I.e. $|X| = 2, |Y| = 3, L = 2$. Let Φ_{xy} be

$$\Phi = \begin{bmatrix} \Phi_{x_1y_1} & \Phi_{x_1y_2} & \Phi_{x_1y_3} \\ \Phi_{x_2y_1} & \Phi_{x_2y_2} & \Phi_{x_2y_3} \end{bmatrix} = \begin{bmatrix} 2 & 1.5 & 1 \\ 1.5 & 2 & 1 \end{bmatrix}$$

and generate $\Phi_{ij} = \Phi_{xy} + \epsilon_{iy} + \eta_{xj}$ where ϵ_{iy} and η_{xj} follow Gumbel distribution whose location parameter is 0 and the scale parameter is 1. An example for the other variables are; $n = (n_{x_1}, n_{x_2}) = (0.5, 0.5), m = (m_{y_1}, m_{y_2}, m_{y_3}) = (0.3, 0.3, 0.4), z(y_1) = z(y_2) = z_1, z(y_3) = z_2, \bar{o}_{z_1} = 0.5, \bar{o}_{z_2} = 0.4, \underline{o}_{z_1} = 0.1, \underline{o}_{z_2} = 0.05$.

3.2 Discrete Choice Representation and Aggregate Equilibrium

Under Assumption 3, 4, 5, the policymaker can relate the observed matching data to the error distribution.

[3] If type y slots belong to region z_1 and z_2, we redefine the type of slots in z_1 as $y_{(1)}$ and z_2 as $y_{(2)}$.

We first introduce the concept of *systematic utilities* U_{xy} and V_{xy}, defined as

$$U_{xy} \coloneqq \min_{i\,:\,x(i)=x} \{u_i - \epsilon_{iy}\},\ V_{xy} \coloneqq \min_{j\,:\,y(j)=y} \{v_j - \eta_{jx}\},$$

and $U_{xy_0} = V_{x_0 y} = 0$ for each $x \in X$ and $y \in Y$.

For each $i \in x$ and $j \in y$, Assumption 4 assumes that $\tilde{\Phi}_{ij} - \epsilon_{iy} - \eta_{xj} \equiv \Phi_{ij}$ does only depend on x and y. Here we additionally assume that error distributions satisfy the following: for each $i \in x$ and $j \in y$, we have $\tilde{\Phi}_{ij} - \epsilon_{iy} - \eta_{xj} = U_{xy} + V_{xy}$.

The following Lemma 1 tells us that the matching outcome when the agents have unobserved preferences is observationally equivalent to the result of discrete choice with unobserved heterogeneity.

Lemma 1. For any worker $i \in I$ and any slot $j \in J$, we can rewrite u_i and v_j as

$$\begin{cases} u_i = \max_{y \in Y_0}\ U_{x(i),y} + \epsilon_{iy} \\ v_j = \max_{x \in X_0}\ V_{x,y(j)} + \eta_{xj}. \end{cases}$$

Lemma 1 also implies that U_{xy}'s and V_{xy}'s can be interpreted as the part of utilities that depend merely on the type pairs. When we write $U = (U_{xy})_{x \in X, y \in Y}$, $V = (V_{xy})_{x \in X, y \in Y}$, $w = (w_z)_{z \in Z}$, the welfare of side X and side Y are defined as follows:

$$G(U) = \sum_{x \in X} n_x \mathop{\mathbb{E}}_{\epsilon_i \sim P_x} \left[\max_{y \in Y_0} U_{xy} + \epsilon_{iy} \right]$$

$$H(V) = \sum_{y \in Y} m_y \mathop{\mathbb{E}}_{\eta_j \sim Q_y} \left[\max_{x \in X_0} V_{xy} + \eta_{xj} \right]$$

By Daly-Zachary-Williams theorem ([20]), we have

$$\frac{\partial}{\partial U_{xy}} \mathop{\mathbb{E}}_{\epsilon_i \sim P_x} \left[\max_{y \in Y_0} U_{xy} + \epsilon_{iy} \right] = \Pr(i \text{ with type } x \text{ chooses type } y).$$

Under the large market approximation, this value coincides with the fraction of type-x agents choosing type y. Thus, $(\partial G(U))/(\partial U_{xy})$ is the demand of x for y; similarly, $(\partial H(V))/(\partial V_{xy})$ is the demand of y for x. In equilibrium, these two should be equal and coincide with the number of matches between x and y. Since G and H are determined by the error distributions, the fact that the realized matching pattern coincides with the partial derivatives of G and H relates the observed matching data to the errors.

We define *aggregate equilibrium with regional constraint*. Let Φ_{xy} be the observed part of the joint surplus generated by a match $i \in x$ and $j \in y$. This term is usually parameterized by the researcher and estimated from the data: [11] provides a method to estimate Φ_{xy}'s given the matching data in the market without regional constraint. In the following, we suppose such data is available and consider Φ_{xy}'s as given when designing a tax scheme.

Definition 3. Given $(\Phi_{xy})_{x,y}$, profile $(\mu, (U, V, w))$ is an *aggregate equilibrium with regional constraints*, if it satisfies the following four conditions:

1. Population constraint; for any $x \in X$, $\sum_{y \in Y_0} \mu_{xy} = n_x$; for any $y \in Y$, $\sum_{x \in X_0} \mu_{xy} = m_y$.
2. No-blocking pairs; for any $(x, y) \in T$, $U_{xy} + V_{xy} \geq \Phi_{xy} - w_{z(y)}$.
3. Market clearing; for any $(x, y) \in T$, $\mu_{xy} = \nabla_{xy} G(U) = \nabla_{xy} H(V)$.
4. Regional constraint; for any $z \in Z$, $\sum_{y \in z} \sum_{x \in X} \mu_{xy} \in [\underline{o}_z, \bar{o}_z]$.

Note that the additive separability and Lemma 1 together imply that condition 2 is equivalent to $u_i + v_j \geq \tilde{\Phi}_{ij}$, which corresponds to the no-blocking pairs condition in the individual equilibrium. Thus the aggregate equilibrium with regional constraints does coincide with an individual equilibrium in the market with unobserved heterogeneity.

4 Efficient Aggregate Equilibrium

The aggregate equilibria need not be unique because the multiple combinations of taxes can adjust the number of matches (see Example ?? below). Hence, we define *efficient aggregate equilibrium*, hereafter EAE, as a refinement. EAE is an aggregate equilibrium that does not impose any tax or subsidy on the regions whose constraints are not binding (complementary slackness).

Definition 4 (Efficient Aggregate Equilibrium)**.** $(\mu, (U, V, w))$ is an *efficient aggregate equilibrium*, if it is an aggregate equilibrium and satisfies the following additional condition:

5. Complementary slackness; for any $z \in Z$,

$$\Big[w_z > 0 \implies \sum_{y \in z} \sum_{x \in X} \mu_{xy} = \bar{o}_z\Big] \text{ and} \Big[w_z < 0 \implies \sum_{y \in z} \sum_{x \in X} \mu_{xy} = \underline{o}_z\Big].$$

Our main result is that EAE always uniquely exists and *efficient* in the sense that it maximizes social welfare among aggregate equilibria. This result is derived as the corollary of Theorem 1, which characterizes EAE as a solution to a convex optimization problem. Furthermore, this enables us to compute EAE by solving the optimization problem. Proofs are deferred to the appendix.

Theorem 1. Fix any $\Phi \in \mathbb{R}^{(N+1)\times(M+1)}$. If $(\mu, (U, V, w))$ is an EAE, then $(U, V, \bar{w}, \underline{w})$ is a solution to the optimization problem (D) and μ satisfies the market clearing condition, where $\bar{w}_z := \max\{0, w_z\}, \underline{w}_z := -\min\{0, w_z\}$. Conversely, if $(U, V, \bar{w}, \underline{w})$ is a solution to the optimization problem (D) and μ satisfies the market clearing condition, then $(\mu, (U, V, w))$ is an EAE.

$$\text{(D)} \qquad \left| \begin{array}{l} \min\limits_{U,V,\bar{w}_z,\underline{w}_z} G(U) + H(V) + \sum\limits_{z \in Z} \bar{o}_z \bar{w}_z - \sum\limits_{z \in Z} \underline{o}_z \underline{w}_z \\ \text{s.t.} \quad \forall (x, y) \in T,\ U_{xy} + V_{xy} \geq \Phi_{xy} - \bar{w}_{z(y)} + \underline{w}_{z(y)}, \\ \qquad \forall z \in Z,\ \bar{w}_z \geq 0,\ \underline{w}_z \geq 0 \end{array} \right.$$

Corollary 1. EAE always exists and is unique.

We can show that the dual of (D) corresponds to the social welfare maximization problem under the population constraints and the regional constraints ((P) in Appendix B). The optimal value of (P) coincides with that of (D) by the strong duality. Since (P) does not restrict how to impose taxes, the region-specific taxation $(w_z)_{z \in Z}$ in the EAE maximizes the social welfare among all possible taxation policy [4] (See Appendix B for details.)

Corollary 2. EAE maximizes the social welfare under regional constraints, whose tax scheme is not conditioned on the types of pairs but dependent only on the region.

Here we give one possible scenario in which Theorem 1 and related results are useful to obtain the optimal tax scheme. Suppose that the data on the existing matching in the market without regional constraints is available. As mentioned in Section 3, then we can estimate Φ_{xy}'s, assuming specific error term distributions, such as Gumbel distribution as in Example 1. The distributional assumption also determines the form of G and H. Given Φ, G, and H, we can solve (D) to obtain tax scheme w, which can be backed up from $\bar{w}$ and $\underline{w}$.

Example 1 (Continued)**.** The yellow area in Figure 1 represents the set of aggregate equilibria: for each (w'_1, w'_2) in the yellow area, there exists an AE $(\mu, (U, V, w))$ in which $w = (w'_1, w'_2)$. We draw the contours of social welfare (Panel (a)), and surplus (or deficit if negative and dotted lines) of the policymaker (Panel (b)). As we see in Corollary 2, the red point in Panel (a), $(w_1, w_2) = (0.5825, 0)$, that tangents to the contour is the EAE. In Section 5, we discuss about budget balanced AE located in the orange region in Panel (b).

[4] Note that since the policymaker can observe x and y, the taxes for each pair may take the form of w_{xy}. Corollary 2 states that we can actually restrict our attention to the taxes of form w_z when maximizing social welfare.

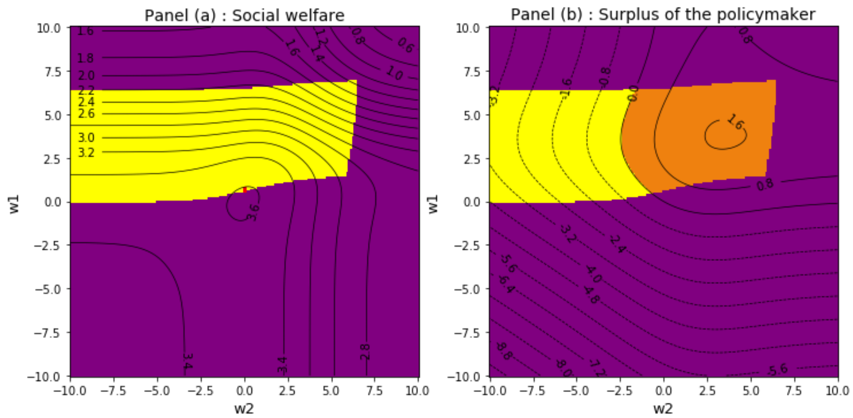

Figure 1: Panel (a) illustrates the social welfare and (b) does the surplus of the PM in the AE given (w_1, w_2). The vertical (horizontal) axis indicates the tax (or subsidy if negative) on region z_1 (z_2). The yellow + orange regions are the set of AE. The red point is the EAE. In Panel (b), the orange region indicates the AE in which the surplus of the PM is nonnegative; BBAE defined in Section 6.

5 Estimation of the Joint Surplus

So far we see how to compute the welfare-maximizing matching μ_{xy} given the known joint surplus Φ_{xy}. In this section, on the contrary, we briefly explain how to estimate the joint surplus Φ_{xy} given the observed matching patterns $\tilde{\mu}_{xy}$.

We take the set of agent types X and Y, their population (n_x) and (m_y). and regions $z(y)$ as given. Now suppose we have the data of

1. observed matching patterns $\tilde{\mu} = (\tilde{\mu}_{xy})_{x,y}$,
2. current tax levels $(w_z)_z$, and
3. type-pair specific covariates $c = (c_{xy})_{x,y}$ (here $c_{xy} \in \mathbb{R}^S$ for some S).

The candidates of c_{xy} are, for example, physical distances between the living area of type-x workers and the office area of type-y job slots, compatibility between workers' skills and job description, or characteristics that depend only on type x or y (such as worker's age or the average wage level around its office). It can simply be the vector of indicator functions of type pairs.

We estimate the Φ_{xy} by the following procedure. First, we choose some parametric function F_λ that maps c_{xy} to Φ_{xy} (e.g. linear regression $F_\lambda(c_{xy}) = \lambda^\top c_{xy}$). Then, we initialize $\lambda = \lambda^0$ and minimize the error computed as follows: in step k,

1. Compute $\Phi^k_{xy} = F_{\lambda^k}(c_{xy})$ for given λ^k
2. Solve (D') and obtain the simulated matching μ^k using $\nabla_{xy} G = \nabla_{xy} H = \mu^k_{xy}$
3. Compute the error (distance) between the simulated matching μ^k and the observed matching $\tilde{\mu}$, $d(\mu^k, \tilde{\mu})$
4. If $d(\mu^k, \tilde{\mu})$ is small enough, finish the estimation. The current $\Phi^k_{xy} = F_{\lambda^k}(c_{xy})$ is the point estimate. Otherwise update $\lambda^k \to \lambda^{k+1}$ so that d becomes smaller[5] and go back to step 1.

$$\text{(D')} \qquad \left| \min_{U,V} \quad G(U) + H(V) \quad \text{s.t.} \quad \forall x \in X, y \in Y,\ U_{xy} + V_{xy} \geq \Phi_{xy} - w_{z(y)} \right.$$

(D') is a convex programming problem and the existence and the uniqueness of the solution are guaranteed in a manner similar to Corollary 1.

[5] How to update the parameter λ depends on the optimization algorithm (e.g. Newton method).

The choice of the parametric function F and the distance function d are arbitrary. See Galichon and Salanié [11] for more details. Here we introduce the maximum likelihood estimation (MLE). Let us adopt the Kullback-Leibler divergence of the multinomial distribution over the type pairs as d,

$$d(\tilde{\mu}, \mu) \equiv \sum_{(x,y)\in T} (\tilde{\mu}_{xy}/|\tilde{\mu}|) \log \frac{(\tilde{\mu}_{xy}/|\tilde{\mu}|)}{(\mu_{xy}/|\mu|)},$$

where $|\mu| \equiv \sum_{(x,y)\in T} \mu_{xy}$. The functional form of F_λ is arbitrary. Given the matching data $\tilde{\mu}$, minimizing the KL-divergence is equivalent to maximizing the log likelihood function

$$\log L(\lambda) \propto \sum_{(x,y)\in T} \tilde{\mu}_{xy} \log \frac{\mu_{xy}}{|\mu|}.$$

Note that λ affects μ_{xy} through $\Phi_{xy} = F_\lambda(c_{xy})$ and (D'). By minimizing d with respect to the parameter λ, we get the estimate of Φ_{xy} by the MLE.

6 Illustrative Experiment

In this section, we compare the EAE with other equilibrium concepts in simulation. Hereafter we take Φ_{xy} as given (already estimated). To get intuition, we simulate a small tractable matching market of residencies with one urban region (z_1) and two rural regions (z_2, z_3) in which all doctors prefer urban hospitals to rural hospitals on average. The policy challenge is to satisfy the minimum standards in the rural regions. We simulate EAE with a larger number of types and regions in Appendix B.

We assume there are 10 types of doctors and 6 types of hospital, $|X| = 10$ and $|Y| = 6$. The population of each type is identical, $n_x = 0.1$ for all $x \in X$ and $m_y = 0.25$ for all $y \in Y$. There are three regions $z_1 = z(y_1) = z(y_2)$, $z_2 = z(y_3) = z(y_4)$, and $z_3 = z(y_5) = z(y_6)$. Region z_1 is attractive for doctors (an urban area) while z_2 and z_3 are not (rural areas). All doctors are identical for each hospital on average. Specifically we set

$$\Phi_{xy} := \begin{cases} 2.0 + \xi_{xy} & (x \in X, y \in z_1) \\ 0.5 + \xi_{xy} & (x \in X, y \in \{z_2, z_3\}) \end{cases},$$

where ξ_{xy} are independent noise drawn from $N(0, 1)$.

We assume there are lower bounds on rural areas, and no other constraints are imposed. We take the same lower bound $\underline{o}_z$ for the rural areas z_2 and z_3 ($\{\underline{o}_z \in \mathbb{R} \mid 0.1 \le \underline{o}_z \le 0.4\}$) and set no lower bound on the urban area $\underline{o}_{z_1} = 0$. We take an average of 30 simulations in Figure 2. (See Appendix B for details of the experiment.)

Here, we compare EAE with three other different aggregate equilibria (plus AE without any regional constraints as baseline).

First one is **EAE upper-bound**. Instead of directly putting the lower bounds on the rural areas, it imposes the "loosest" upper bound on the urban region so that sufficient proportion of doctors move to the unpopular regions under the EAE (See Appendix B for details). For example, if the PM would like to satisfy the lower bounds $\underline{o}_z = 0.4$, the EAE upper-bound instead sets an upper-bound with the smallest $\bar{o}_{z_1}$ satisfying $\sum_x \sum_{y \in z_2} \mu_{xy}^{\text{EAE-UB}} \ge 0.4$ and $\sum_x \sum_{y \in z_3} \mu_{xy}^{\text{EAE-UB}} \ge 0.4$. Replacing floor constraints with ceil constraints is a common technique in the theoretical literature since it is often impossible for algorithms to satisfy desirable properties like strategy-proofness under floor constraints.

Second one is **cap-reduced AE**. Similar to EAE upper bound, it also limits the maximum number of doctors matched in the urban area, instead of directly imposing the lower bounds on the rural areas. The difference is, here the PM artificially reduces the capacities of **each hospital type** in the urban area, instead of imposing an upper bound on the urban region. For example, instead of imposing the regional upper-bound $\bar{o}_{z_1} = 0.4$ in the EAE upper-bound, the PM sets the artificial capacities $m_{y_1} = m_{y_2} = 0.2$ of the urban hospital slots and compute the AE without tax and subsidy. This is obviously an inefficient policy but is frequently used in practice, including the JRMP.

Third one is **budget balanced AE (BBAE)**, in which the PM should attain the budget balance, i.e., $\sum_{x,y} \mu_{xy} w_{z(y)} \ge 0$. Note that BBAE may not be unique. (See Panel (b) of Figure 1.) Here we focus on BBAE which maximizes social welfare.

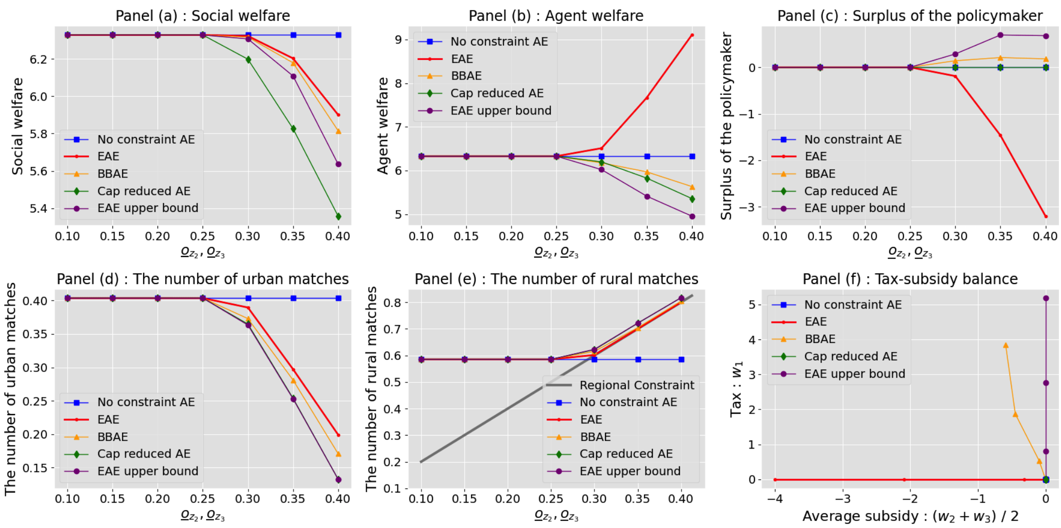

Figure 2: Panel (a)-(e): the horizontal axis is the lower bound constraints of rural regions ($\underline{o}_{z_2} = \underline{o}_{z_3}$), and the vertical axes indicate (a) the levels of the social welfare, (b) the agent welfare (the worker welfare + the job slot welfare), (c) the policymaker surplus, (d) the number of matched workers in the urban region z_1, and (e) in the rural regions z_2, z_3 of each algorithm. The gray line in Panel (e) indicates the sum of the lower bounds of z_2, z_3. Panel (f): the locus of the equilibrium tax and average subsidy of each method when $\underline{o}_{z_2}, \underline{o}_{z_3}$ changes. Note that the no constraint AE and the cap-reduced AE use neither tax nor subsidy, the blue and the green lines stick to the origin.

It is worth mentioning that the order of the equilibrium concepts in terms of social welfare remains unchanged for any problem instance.

Fact 1. For any instances, the levels of social welfare are in the following order:

$$\text{EAE} \geq \text{BBAE} \geq \text{EAE upper-bound} \geq \text{Cap-reduced AE}.$$

Here the first inequality comes from the fact that the EAE is welfare-maximizing as Corollary 2 states, the second one holds because the EAE upper-bound imposes only taxes on urban areas so the PM always makes a positive surplus, and the third one again follows from the fact that the EAE maximizes the social welfare (under the alternative upper-bound constraints).

We summarize the performance comparison results in Figure 2. Panel (a) illustrates how much social welfare each equilibrium achieves at each lower bound level and confirms Fact 1. Panel (b) and (c) represent the agents (the workers and the job slots) welfare and the surplus of the PM, which reveals that in the EAE, the social efficiency is achieved by relocating the surplus from the PM to the agents. Panel (d) and (e) show the number of doctors matched in the urban and the rural areas. We can see that the EAE successfully keeps the number of doctors in z_1 as much as possible while satisfying the rural lower bounds tight. Finally Panel (f) depicts the locus of the tax and the average subsidy levels when $\underline{o}_z$ changes. It show the EAE and the EAE upper-bound are the complete opposites; the EAE uses subsidies on the rural regions only while the EAE upper-bound uses a tax on the urban are only. Here the BBAE balances a tax and subsidies so that it maximizes the social welfare in the range that the surplus of the PM remains nonnegative. When the PM cannot make a deficit, the BBAE is the second best choice that is more desireble than the EAE upper-bound or the cap-reduced AE.

7 Summary and Discussion

This paper develops a framework to conduct a counterfactual analysis to regulate matching markets with constraints. Extending the framework of Galichon and Salanié [11], we propose a method to (1) estimate the utility functions of agents from currently available data about matching patterns, (2) predict outcomes under counterfactual policies, and (3) design the welfare-maximizing tax scheme. Our results suggest that there is a better way to implement regional constraints than the policy currently used in practice.

Our framework should be useful for PMs who want to figure out how costly it is to implement the regional constraints. Although constraints are exogenously given in most papers in the literature on matching with constraints, the PM often needs to set them balancing the tradeoff between their tightness and the welfare loss. Our method allows her to estimate how much she should pay to meet the different levels of regional quotas, and helps her decision making.

In the following, we discuss two possible future directions that must be of interest. First, our framework relies on the large market approximation, assuming that each type has sufficiently many agents. However, when the PM has access to detailed individual data, she may use it to design a better policy. It is worth considering how we can utilize such fine-grained data in this setting. Second, although our framework assumes the transferable utility, there is another thick strand of research on matching that assumes non-transferable utilities. Galichon et al. [12] describes the general framework to handle imperfectly transferable utility matching, which includes both transferable and non-transferable utility matching as special cases. An extension of our framework to such a general setting is also one promising future direction.

References

[1] Atila Abdulkadiroğlu and Tayfun Sönmez. School choice: A mechanism design approach. *American Economic Review*, 93(3):729–747, 2003.

[2] Victor Aguirregabiria and Pedro Mira. Dynamic discrete choice structural models: A survey. *Journal of Econometrics*, 156(1):38–67, 2010. ISSN 0304-4076. doi: https://doi.org/10.1016/j.jeconom.2009.09.007. URL `https://www.sciencedirect.com/science/article/pii/S0304407609001985`. Structural Models of Optimization Behavior in Labor, Aging, and Health.

[3] Eduardo M. Azevedo and Jacob D. Leshno. A supply and demand framework for two-sided matching markets. *Journal of Political Economy*, 124(5):1235–1268, 2016. doi: 10.1086/687476. URL `https://doi.org/10.1086/687476`.

[4] Gary S. Becker. A Theory of Marriage: Part I. *Journal of Political Economy*, 81(4):813–846, 1973.

[5] Peter. Biró, Tamas. Fleiner, Robert.W. Irving, and David.F. Manlove. The college admissions problem with lower and common quotas. *Theoretical Computer Science*, 411(34-36):3136–3153, 2010.

[6] Stéphane Bonhomme. *Econometric analysis of bipartite networks*, pages 83–121. 01 2020. ISBN 9780128117712. doi: 10.1016/B978-0-12-811771-2.00011-0.

[7] Eugene Choo and Aloysius Siow. Estimating a marriage matching model with spillover effects. *Demography*, 43(3):463–490, 2006. ISSN 0070-3370. doi: 10.1353/dem.2006.0023.

[8] Lars Ehlers, Isa E. Hafalir, M. Bumin Yenmez, and Muhammed A. Yildirim. School choice with controlled choice constraints: Hard bounds versus soft bounds. *Journal of Economic Theory*, 153:648–683, 2014.

[9] Daniel Fragiadakis, Atsushi Iwasaki, Peter Troyan, Suguru Ueda, and Makoto Yokoo. Strategyproof matching with minimum quotas. *ACM Transactions on Economics and Computation*, 4, 2016. (an extended abstract appeared in AAMAS, pages 1327–1328, 2012).

[10] David Gale and Lloyd Stowell Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

[11] Alfred Galichon and Bernard Salanié. Cupid's Invisible Hand: Social Surplus and Identification in Matching Models. *Review of Economic Studies, forthcoming*, 2021.

[12] Alfred Galichon, Scott Duke Kominers, and Simon Weber. Costly concessions: An empirical framework for matching with imperfectly transferable utility. *Journal of Political Economy*, 127(6):2875–2925, 2019. doi: 10.1086/702020.

[13] Masahiro Goto, Atsushi Iwasaki, Yujiro Kawasaki, Ryoji Kurata, Yosuke Yasuda, and Makoto Yokoo. Strategyproof matching with regional minimum and maximum quotas. *Artificial Intelligence*, 235:40–57, 2016.

[14] Michael Greinecker and Christopher Kah. Pairwise stable matching in large economies. *Econometrica*, 89(6):2929–2974, 2021. doi: https://doi.org/10.3982/ECTA16228. URL `https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA16228`.

[15] Isa E. Hafalir, M. Bumin Yenmez, and Muhammed A. Yildirim. Effective affirmative action in school choice. *Theoretical Economics*, 8(2):325–363, 2013.

[16] Yuichiro Kamada and Fuhito Kojima. Efficient matching under distributional constraints: Theory and applications. *American Economic Review*, 105(1):67–99, 2015.

[17] Alexander Kelso and Vincent Crawford. Job matching, coalition formation, and gross substitutes. *Econometrica*, 50(6):1483–1504, 1982.

[18] Fuhito Kojima. School choice: Impossibilities for affirmative action. *Games and Economic Behavior*, 75(2):685–693, 2012.

[19] Ryoji Kurata, Naoto Hamada, Atsushi Iwasaki, and Makoto Yokoo. Controlled school choice with soft bounds and overlapping types. *Journal of Artificial Intelligence Research*, 58:153–184, 2017.

[20] Daniel McFadden. Econometric models for probabilistic choice among products. *The Journal of Business*, 53(3):S13–S29, 1980.

[21] Georg Nöldeke and Larry Samuelson. The implementation duality. *Econometrica*, 86(4): 1283–1324, 2018. doi: https://doi.org/10.3982/ECTA13307. URL `https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA13307`.

[22] Lloyd S Shapley and Martin Shubik. The assignment game I: The core. *International Journal of game theory*, 1(1):111–130, 1971.

A Omitted proofs

A.1 Proof of Lemma 1

Proof. For any type x worker $i \in I$, by definition of U, we have

$$\begin{aligned} & U_{xy} \le u_i - \epsilon_{iy},\ \forall y \in Y_0 \\ \iff\ & u_i \ge U_{xy} + \epsilon_{iy},\ \forall y \in Y_0 \\ \iff\ & u_i \ge \max_{y \in Y_0} U_{xy} + \epsilon_{iy}. \end{aligned}$$

Similarly, for any type y worker $j \in J$, we have $v_j \ge \max_{x \in X_0}\ V_{xy} + \eta_{xj}$.

Suppose to the contrary that there exists type x worker $i \in I$ that satisfies

$$u_i > \max_{y \in Y_0} U_{xy} + \epsilon_{iy}.$$

If i is matched with slot $j \in J$, then

$$\begin{aligned} u_i + v_j &= \Phi_{ij} - w_{z(y)} \\ &> \left(\max_{y' \in Y_0}\ U_{xy'} + \epsilon_{iy'}\right) + \left(\max_{x' \in X_0}\ V_{x'y} + \eta_{x'j}\right) \\ &\ge U_{xy(j)} + \epsilon_{iy(j)} + V_{xy(j)} + \eta_{xj} \\ &\ge \Phi_{xy} - w_{z(y)} + \epsilon_{iy(j)} + \eta_{jx} \\ &= \Phi_{ij} - w_{z(y)}, \end{aligned}$$

which includes a contradiction. Similarly we see a contradiction when we assume $v_j > \max_{x \in X_0}\ V_{xy} + \eta_{xj}$. $\square$

A.2 Theorem 1

First, we show two lemmas used in the proof.

Lemma 2. G and H are strictly increasing and strictly convex.

Proof. $\underline{G \text{ is strictly increasing.}}$ Take any $U^1, U^2 \in \mathbb{R}^{N \times M}$ such that $U^1 \geq U^2$ and $U^1 \neq U^2$. Then $G(U^1) \geq G(U^2)$ by definition. In addition, note that $U^1_{xy} > U^2_{xy}$ holds for some $x \in X$ and $y \in Y$. Since P_x has full support, we have

$$\Pr_{\epsilon_i}(u_i = U^1_{xy} + \epsilon_{iy}) \geq \Pr_{\epsilon_i}(u_i = U^2_{xy} + \epsilon_{iy}) > 0.$$

Because $\mathbb{E}_{\epsilon_i}[u_i \mid u_i = U_{xy} + \epsilon_{iy}]$ is strictly increasing in U_{xy}, we have

$$\begin{aligned}\mathbb{E}_{\epsilon_i}\left[u_i \mid u_i = U^1_{xy} + \epsilon_{iy}\right] &\cdot \Pr_{\epsilon_i}(u_i = U^1_{xy} + \epsilon_{iy}) \\ &> \mathbb{E}_{\epsilon_i}\left[u_i \mid u_i = U^2_{xy} + \epsilon_{iy}\right] \cdot \Pr_{\eta_j}(u_i = U^2_{xy} + \epsilon_{iy}),\end{aligned}$$

and thus $G(U^1) > G(U^2)$ holds.

$\underline{G \text{ is strictly convex.}}$ Take any $U^1, U^2 \in \mathbb{R}^{N \times M}$ and $s \in [0,1]$. Since

$$\begin{aligned} sG(U^1) + (1-s)G(U^2) &= \sum_x n_x \mathbb{E}\left[\Big(\max_y \ s(U^1_{xy} + \epsilon_{iy})\Big) + \Big(\max_y \ (1-s)(U^2_{xy} + \epsilon_{iy})\Big)\right] \\ &\geq \sum_x n_x \mathbb{E}\left[\max_y \ sU^1_{xy} + (1-s)U^2_{xy} + \epsilon_{iy}\right] \\ &= G\Big(sU^1 + (1-s)U^2\Big) \end{aligned}$$

holds, G is a convex function.

Now suppose $U^1 \neq U^2$. Then $U^1_{xy} \neq U^2_{xy}$ holds for some $x \in X$, $y \in Y$. Without loss of generality, assume $U^1_{xy} > U^2_{xy}$. Since P_x is full support,

$$\Pr\left(\left\{\epsilon_i \mid U^1_{xy} + \epsilon_{iy} > \max_{y' \neq y} U^1_{xy'} + \epsilon_{iy'} \quad \wedge \quad \max_{y' \neq y} U^2_{xy'} + \epsilon_{iy'} > U^2_{xy} + \epsilon_{iy}\right\}\right) > 0$$

holds. Therefore for any $s \in (0,1)$, we have

$$sG(U^1) + (1-s)G(U^2) > G\Big(sU^1 + (1-s)U^2\Big),$$

which implies G is strictly convex. Similarly, we can show H is also strictly increasing and strictly convex. $\square$

Lemma 3. For each x and y, if $\mu_{xy} > 0$, then $U_{xy} + V_{xy} = \Phi_{xy} - w_{z(y)}$.

Proof. Fix any x and y. Suppose that $\mu_{xy} > 0$. Then there exists i and j such that $x(i) = x$, $y(j) = y$, and $d_{ij} = 1$. Suppose toward contradiction that $U_{xy} + V_{xy} > \Phi_{xy} - w_{z(y)}$. By the definition of U_{xy} and V_{xy}, we have

$$u_i - \epsilon_{iy} + v_j - \eta_{xj} > \Phi_{xy} - w_{z(y)},$$

which implies that $u_i + v_j > \tilde{\Phi}_{ij} - w_{ij}$. A contradiction. $\square$

Proof of Theorem 1.

Proof. Let's consider the necessary and sufficient conditions of the solution to (D). Let $(\lambda_{xy})_{x \in X, y \in Y}, (\bar{\lambda}_z, \underline{\lambda}_z)_{z \in Z}$ be lagrange multipliers, then the Lagrangean denoted by $\mathcal{L}$ is computed as follows;

$$\begin{aligned}\mathcal{L} = G(U) + H(V) &+ \sum_{z \in Z} \bar{o}_z \bar{w}_z - \sum_{z \in Z} \underline{o}_z \underline{w}_z \\ &+ \sum_{x \in X, y \in Y} \lambda_{xy}\left(U_{xy} + V_{xy} - \Phi_{xy} + \bar{w}_{z(y)} - \underline{w}_{z(y)}\right) + \sum_{z \in Z} \bar{\lambda}_z \bar{w}_z + \sum_{z \in Z} \underline{\lambda}_z \underline{w}_z.\end{aligned}$$

The KKT conditions are,

$$\frac{\partial \mathcal{L}}{\partial U_{xy}} = \nabla_{xy} G(U) + \lambda_{xy} = 0 \quad \forall x \in X, y \in Y, \tag{1}$$

$$\frac{\partial \mathcal{L}}{\partial V_{xy}} = \nabla_{xy} H(V) + \lambda_{xy} = 0 \quad \forall x \in X, y \in Y, \tag{2}$$

$$\frac{\partial \mathcal{L}}{\partial \bar{w}_z} = \bar{o}_z + \sum_{y \in z} \sum_{x \in X} \lambda_{xy} + \bar{\lambda}_z = 0 \quad \forall z \in Z, \tag{3}$$

$$\frac{\partial \mathcal{L}}{\partial \underline{w}_z} = -\underline{o}_z - \sum_{y \in z} \sum_{x \in X} \lambda_{xy} + \underline{\lambda}_z = 0 \quad \forall z \in Z, \tag{4}$$

$$U_{xy} + V_{xy} - \Phi_{xy} + \bar{w}_{z(y)} - \underline{w}_{z(y)} \geq 0 \quad \forall x \in X, y \in Y \tag{5}$$

$$\lambda_{xy} \leq 0, \quad \lambda_{xy} \left(U_{xy} + V_{xy} - \Phi_{xy} + \bar{w}_{z(y)} - \underline{w}_{z(y)} \right) = 0 \quad \forall x \in X, y \in Y \tag{6}$$

$$\bar{w}_z \geq 0, \ \bar{\lambda}_z \leq 0, \ \bar{\lambda}_z \bar{w}_z = 0 \quad \forall z \in Z \tag{7}$$

$$\underline{w}_z \geq 0, \ \underline{\lambda}_z \leq 0, \ \underline{\lambda}_z \underline{w}_z = 0 \quad \forall z \in Z. \tag{8}$$

This satisfies the linearly independent constraint qualification, which implies these are the necessary conditions for the optimality.

From Lemma 2, the objective function of (D) is convex with respect to $(U, V, \bar{w}, \underline{w})$. Because the constraints are linear in the parameters, KKT conditions are also sufficient conditions for the optimality.

EAE $\implies$ solution of (D): Take any EAE $(\mu, (U, V, w))$ and define $\bar{w}_z = \max\{0, w_z\}$ and $\underline{w}_z = -\min\{0, w_z\}$ for all $z \in Z$.

Define $\lambda_{xy} = -\mu_{xy}$ for all $x \in X, y \in Y$. Then from condition 3 of EAE, the following holds;

$$-\lambda_{xy} = \mu_{xy} = \nabla_{xy} G(U) = \nabla_{xy} H(V), \quad \forall x \in X, y \in Y.$$

This implies that (A.2) and (A.2)are satisfied. Now $\mu_{xy} \geq 0$ implies $\lambda_{xy} \leq 0$. From condition 2 of EAE,

$$U_{xy} + V_{xy} - \Phi_{xy} + \bar{w}_{z(y)} - \underline{w}_{z(y)} = 0.$$

This implies that (A.2)) and (A.2) are satisfied.

Next, when we define

$$\bar{\lambda}_z = \sum_{y \in z} \sum_{x \in X} \mu_{xy} - \bar{o}_z$$

$$\underline{\lambda}_z = \sum_{y \in z} \sum_{x \in X} -\mu_{xy} + \underline{o}_z,$$

(A.2) and (A.2) are directly implied. Furthermore, by definition, $\bar{w}_z, \underline{w}_z \geq 0$ for every $z \in Z$. And condition 4 of EAE implies that $\bar{\lambda}_z \leq 0$, $\underline{\lambda}_z \leq 0$. From condition 5 of EAE gives;

$$\bar{w}_z > 0 \implies w_z > 0 \implies \sum_{y \in z} \sum_{x \in X} \mu_{xy} = \bar{o}_z$$

$$\underline{w}_z > 0 \implies w_z < 0 \implies \sum_{y \in z} \sum_{x \in X} \mu_{xy} = \underline{o}_z,$$

which implies that (A.2) and (A.2). So we are done with this part.

A solution of (D) $\implies$ EAE: Take any $(U, V, \bar{w}, \underline{w}, \lambda)$ satisfying KKT conditions and define $\mu_{xy} = -\lambda_{xy}$ then (A.2) and (A.2) implies that $\mu_{xy} = \nabla_{xy} G(U) = \nabla_{xy} H(V)$ and so condition 3 is satisfied.

(A.2) and (A.2) implies;

$$\bar{\lambda}_z = -\sum_{y \in z} \sum_{x \in X} \lambda_{xy} - \bar{o}_z \leq 0 \iff \sum_{y \in z} \sum_{x \in X} \mu_{xy} \leq \bar{o}_z.$$

Similarly, (A.2) and (A.2) implies;

$$\underline{\lambda}_z = \sum_{y\in z}\sum_{x\in X}\lambda_{xy} + \underline{o}_z \le 0 \iff \sum_{y\in z}\sum_{x\in X}\mu_{xy} \ge \underline{o}_z.$$

These says that condition 4 is satisfied.

Define $w_z = \bar{w}_z - \underline{w}_z$, then from (A.2) we get

$$w_z > 0 \implies \bar{w}_z > 0 \implies \bar{\lambda}_z = \sum_{y\in z}\sum_{x\in X}\mu_{xy} - \bar{o}_z = 0.$$

Now from (A.2),

$$w_z < 0 \implies \underline{w}_z > 0 \implies \underline{\lambda}_z = -\sum_{y\in z}\sum_{x\in X}\mu_{xy} + \underline{o}_z = 0.$$

Thus, we obtain condition 5 of EAE.

Next, from $\mu_{xy} = \nabla_{xy}G(U) = \nabla_{xy}H(V)$, we get

$$\sum_{y\in Y_0}\mu_{xy} = \sum_{y\in Y_0}\nabla_{xy}G(U) = n_x\sum_{y\in Y_0}\Pr\Big(\{\epsilon_{iy} \mid u_i = U_{xy} + \epsilon_{iy}\}\Big) = n_x, \quad \forall x \in X$$

$$\sum_{x\in X_0}\mu_{xy} = \sum_{x\in X_0}\nabla_{xy}H(V) = m_y\sum_{x\in X_0}\Pr\Big(\{\eta_{xj} \mid v_j = V_{xy} + \eta_{xj}\}\Big) = m_y, \quad \forall y \in Y.$$

This is equivalent to condition 1 of EAE.

Lastly, Assumption 5 assures us the following; for any $x \in X$, $y \in Y$,

$$-\lambda_{xy} = \nabla_{xy}G(U) = \Pr\Big(\{\epsilon_{iy} \mid u_i = U_{xy} + \epsilon_{iy}\}\Big) > 0$$
$$-\lambda_{xy} = \nabla_{xy}H(V) = \Pr\Big(\{\eta_{xj} \mid v_j = V_{xy} + \eta_{xj}\}\Big) > 0.$$

(A.2) implies

$$U_{xy} + V_{xy} = \Phi_{xy} - w_z(y), \quad \forall x \in X, y \in Y.$$

Hence, condition 2 of EAE is satisfied.

Therefore we are done with this part.

□

A.3 Proof of Corollary 1

Proof.

Existence: First, observe that the feasible set of (D) is nonempty and convex. Then by the theorem of convex duality, (D) has a solution.

Uniqueness: Fix any EAEs $(\mu, (U, V, w))$ and $(\mu', (U', V', w'))$. We want to show that $(\mu, (U, V, w)) = (\mu', (U', V', w'))$.

Let $I(w) \coloneqq \sum_{z\in Z}\bar{o}_z\max\{0, w_z\} - \sum_{z\in Z}\underline{o}_z(-\min\{0, w_z\})$. Then, the objective function of (D) can be rewritten as

$$g(U, V, w) \coloneqq G(U) + H(V) + I(w).$$

Note that I is convex, and hence g is also convex.

Consider

$$(\mu'', (U'', V'', w'')) \coloneqq \frac{1}{2}(\mu, (U, V, w)) + \frac{1}{2}(\mu', (U', V', w')).$$

Note that $(\mu'', (U'', V'', w''))$ is feasible in (D). Since G and H are strictly convex and any EAE should be a solution to (D), we have $U = U' = U''$ and $V = V' = V''$; otherwise $g(U'', V'', w'') < g(U, V, w)$ and this contradicts the optimality of (U, V, w).

Suppose toward contradiction that $w \neq w'$, or there exists z_0 such that $w_{z_0} \neq w'_{z_0}$. First, note that since $U = U'$ and $V = V'$, we have $\mu = \mu'$ by the market clearing condition.

Case (i): $\sum_{y\in z_0}\sum_x \mu_{xy} \in (\bar{o}_{z_0}, \underline{o}_{z_0})$**.** By the complementary slackness condition, we have $w_{z_0} = w'_{z_0}$. A contradiction.

Case (ii): $\sum_{y\in z_0}\sum_x \mu_{xy} = \bar{o}_{z_0}$ **(**>0**).** By the complementary slackness condition, we have $w_{z_0} \geq 0$ and $w'_{z_0} \geq 0$. Since $w_{z_0} \neq w'_{z_0}$, assume without loss of generality that $w_{z_0} > w'_{z_0} \geq 0$. Let $\tilde{w} \coloneqq (w''_{z_0}, w_{-z_0})$. Observe that $(U, V, \tilde{w})$ is feasible in (D).

Then, we have

$$\begin{aligned} g(U,V,\tilde{w}) - g(U,V,w) &= \bar{o}_{z_0}(w''_{z_0} - w_{z_0}) \\ &< 0 \quad (\because\ w''_{z_0} < w_{z_0}), \end{aligned}$$

which contradicts the optimality of (U, V, w).

Case (iii): $\sum_{y\in z}\sum_x \mu_{xy} = \underline{o}_{z_0}$**.** If $\underline{o}_{z_0} > 0$, we can show $w_{z_0} = w'_{z_0}$ in a similar manner to case (ii).

Suppose that $\underline{o}_{z_0} = 0$. Assume without loss of generality that $w_{z_0} > w'_{z_0} \geq 0$. Let $\tilde{w} \coloneqq (0, w_{-z_0})$, and $\tilde{U} \coloneqq (U_{xy} - w_{z_0}, U_{-(xy)})$. Observe that $(\tilde{U}, V, \tilde{w})$ is feasible in (D). Since function G is strictly increasing in U_{xy} by Lemma 2, we have

$$g(\tilde{U}, V, \tilde{w}) < g(U, V, w),$$

which contradicts the optimality of (U, V, w). $\square$

A.4 Proof of Corollary 2

The dual expression of (D) is

$$\text{(P)} \quad \left| \begin{array}{ll} \text{maximize}_{\mu_{xy}\geq 0} & \displaystyle\sum_{(x,y)\in T} \mu_{xy}\Phi_{xy} + \mathcal{E}(\mu) \\ \text{subject to} & \displaystyle\sum_{y\in Y_0} \mu_{xy} = n_x,\ \forall x \in X \\ & \displaystyle\sum_{x\in X_0} \mu_{xy} = m_y,\ \forall y \in Y \\ & \displaystyle\underline{o}_z \leq \sum_{y\in z}\sum_{x\in X} \mu_{xy} \leq \bar{o}_z,\ \forall z \in Z \end{array} \right.$$

where $\mathcal{E}(\mu) \coloneqq -G^*(\mu) - H^*(\mu)$, and G^*, H^* are the Legendre-Fenchel transform of G, H:

$$\begin{aligned} G^*(\mu) &\coloneqq \sup_U \Big\{ \sum_{x\in X}\sum_{y\in Y_0} \mu_{xy}U_{xy} - G(U) \Big\} \\ H^*(\mu) &\coloneqq \sup_V \Big\{ \sum_{y\in Y}\sum_{x\in X_0} \mu_{xy}V_{xy} - H(V) \Big\}. \end{aligned}$$

First, we will show the following lemma:

Lemma 4. For each i and j, let

$$Y_i^* \in \arg\max_{y\in Y_0} \{U_{xy} + \epsilon_{iy}\}, \quad X_j^* \in \arg\max_{x\in X_0} \{V_{xy} + \eta_{xj}\}.$$

Then, we have

$$\mathcal{E}(\mu) = \sum_{x\in X} n_x \mathop{\mathbb{E}}_{\epsilon_i\sim P_x}\left[\epsilon_{i,Y_i^*}\right] + \sum_{y\in Y} m_y \mathop{\mathbb{E}}_{\eta_j\sim Q_y}\left[\eta_{X_j^*,j}\right]$$

Proof of Lemma 4. The proof can be found in Section 3 and 4 in Galichon and Salanié [11]. For this paper to be self-contained, we will give the proof here.

For each $x \in X$, let $U_x \coloneqq (U_{xy})_{y \in Y_0}$ and

$$G_x(U_x) \coloneqq \mathop{\mathbb{E}}_{\epsilon_i \sim P_x}\left[\max_{y \in Y_0}\{U_{xy} + \epsilon_{iy}\}\right].$$

Then, by definition, we have $G(U) = \sum_{x \in X} n_x G_x(U_x)$.

Let $\bar{\mu}_{xy} \coloneqq \mu_{xy}/n_x$, which is the proportion of type-x agents matched with type-y agents. The Legendre-Fenchel transform of G_x is

$$G_x^*(\bar{\mu}_x) \coloneqq \begin{cases} \sup_{U_x}\left\{\sum_{y \in Y_0} \bar{\mu}_{xy} U_{xy} - G_x(U_x)\right\} & \left(\sum_{y \in Y} \bar{\mu}_y \leq 1\right) \\ +\infty & \text{(o.w.)} \end{cases}, \tag{9}$$

and it follows that $G^*(\mu) = \sum_{x \in X} n_x G_x^*(\bar{\mu}_x)$. By the theory of convex duality, G_x is also a Legendre-Fenchel transform of G_x^*:

$$G_x(U_x) = \sup_{\bar{\mu}_x}\left\{\sum_y \bar{\mu}_{xy} U_{xy} - G_x^*(\bar{\mu}_x)\right\}. \tag{10}$$

Suppose that $\bar{\mu}_x$ attains the supremum in (A.4). By (A.4) and (A.4), we have

$$G_x(U_x) + G_x^*(\bar{\mu}_x) = \sum_y \bar{\mu}_{xy} U_{xy}. \tag{11}$$

Note that, under the large market approximation, we have

$$G_x(U_x) = \sum_{i \in x} \bar{\mu}_{xy} U_y + \mathop{\mathbb{E}}_{\epsilon_i \sim P_x}\left[\epsilon_{i, Y_i^*}\right]. \tag{12}$$

Then, by (A.4) and (A.4), we have

$$G_x^*(\bar{\mu}_x) = -\mathop{\mathbb{E}}_{\epsilon_i \sim P_x}\left[\epsilon_{i, Y_i^*}\right].$$

Therefore, we have

$$-G^*(\mu) = \sum_{x \in X} n_x \mathop{\mathbb{E}}_{\epsilon_i \sim P_x}\left[\epsilon_{i, Y_i^*}\right].$$

By a similar argument, we can show that

$$-H^*(\mu) = \sum_{y \in Y} m_y \mathop{\mathbb{E}}_{\eta_j \sim Q_y}\left[\eta_{X_j^*, j}\right].$$

□

Lemma 4 states that $\mathcal{E}(\mu)$ captures the social surplus unobserved by the policy maker, which is the summation of the error terms that contribute to the social surplus. Hence, the objective function of (P) indeed represents the social welfare in this economy. Because the objective function is concave and the constraints are linear, the optimal value of (P) coincides with that of (D). Therefore the EAE maximizes the social welfare under the regional constraints. Furthermore, the optimal tax scheme is obtained as the lagrange multipliers for the regional constraints, they only depend on the binding region z. □

B Omitted Explanations

B.1 Validity of Individual Equilibrium

First, we recall the definition of the individual equilibrium. Below u_i and v_j are interpreted as i's and j's payoffs in the equilibrium, respectively.

Definition 5 (Individual Equilibrium)**.** A profile $(d, (u, v), w)$ of feasible matching d, equilibrium payoffs of agents $(u, v) \in \mathbb{R}^{|I|} \times \mathbb{R}^{|J|}$, and taxes $w \in \mathbb{R}^L$ is *stable* if

- (Individual rationality) For all i, $u_i \geq \tilde{\Phi}_{ij_0} := 0$, with equality if i is unmatched in d; for all j, $v_j \geq \tilde{\Phi}_{i_0 j} := 0$, with equality if j is unmatched in d.
- (No blocking pairs) For all i and j, $u_i + v_j \geq \tilde{\Phi}_{ij}$, with equality if i and j are matched in d.

We can expect that $(d, (u, v), w)$ satisfies the stability condition in the economy by the following reason. First, if i and j match, their payoff should satisfy $u_i + v_j = \tilde{\Phi}_{ij}$. Here we only assume that they divide the joint surplus without any waste. Suppose, to the contrary, that $u_i + v_j < \tilde{\Phi}_{ij}$ holds for some i and j in a matching. These i and j are not matched with each other under the equilibrium matching d. If they deviate from the current match and form a pair (i, j), they will produce the net joint surplus $\tilde{\Phi}_{ij}$. Then they can divide the net joint surplus so that both of them can be strictly better off. (We implicitly assume that they can reach such an agreement.) Therefore, $u_i + v_j < \tilde{\Phi}_{ij}$ cannot occur in the equilibrium.

B.2 Additional Explanation for Assumptions

To derive Theorem 1, we made five assumptions in total. We restate them below.

Assumption 1 (Regional Constraint)**.** Each type $y \in Y$ can belong to only one region. Denote the region that type y belongs to by $z(y) \in Z$.

Assumption 2 (Large Market Approximation)**.** Each type x has n_x mass of continuum of agents rather than distinct n_x agents. Similarly we assume there are m_y mass of type y job slots.

Assumption 3 (Independence)**.** For each x and $i \in x$, error term ϵ_{iy} is drawn from the distribution P_x. Similarly, for each y and $j \in y$, error term η_{xj} is drawn from the distribution Q_y. The error terms are independent across all i's and j's.

Assumption 4 (Additive Separability)**.** For each x and y, $\tilde{\Phi}_{ij} - \epsilon_{iy} - \eta_{xj}$ is constant over the individuals $i \in x$ and $j \in y$, and denoted by Φ_{xy}.

Assumption 5 (Smooth Distibution with Full-Support)**.** For each x and y, the cdf's P_x and Q_y are continuously differentiable. Moreover, $\mathrm{supp}(P_x) = \mathbb{R}^{N+1}$ and $\mathrm{supp}(Q_y) = \mathbb{R}^{M+1}$.

Assumption 1 is without loss of generality: if type y slots belong to region z_1 and z_2, we redefine the type of slots in z_1 as $y_{(1)}$ and z_2 as $y_{(2)}$.

Assumption 2 implies that this is not a *finite agent model* as in [10, 22] but a *large economy model* as in [3, 12, 21]. As mentioned in [14], large economy model usually disposes of the existence of individuals and considers the matching outcomes defined as the pair of observable types as in the current paper. This is originally to diminish the influence of the individual to compute the stable matching under some externality. [11, 12] recently takes advantage of this model to identify (i.e., express the parameter of interest as a function of the distribution of the observed data) the matching model. We follow this strategy to get the identification result from the econometric viewpoint.

Assumption 3 can be relaxed in some sense, but as mentioned in Aguirregabiria and Mira (2010), it is rare to drop the independence assumption in the discrete choice literature because the model becomes highly non-tractable.

Assumption 4 allows us to characterize the stable matching as the market equilibrium of discrete choice problems from both the worker side and the job slot side over the observable types of the other side. Additivity and separability are common assumptions in the econometrics literature [2, 6].

Assumption 5 ensures that we observe at least one match between any two observable types. This is an essential assumption for our identification result. Imagine we do not observe any match between a pair of (x, y). There are two possible cases for this to happen: (1) $\Phi_{xy} = -\infty$, or (2) just a realization error. Full support assumption eliminates the second case, and thus we can assume that a type pair suffers from infinite loss when we do not observe any matching between the pair. In this paper, we focus finite surplus setting so we exclude the above situation by setting Assumption 5.

B.3 Simulate EAE in a Larger Market

Although the EAE is relatively easy to compute by the convex programming (D), we measured how long it takes to compute. We use CVXPY solver[6] on our M1 Max Macbook Pro. We simulate the markets with $|X| = 10, 20$ doctor types and $|Z| = 5, 10, \ldots, 100$ regions. We assume each region has 10 types of hospitals.

The actual JRMP problem has 47 regions (all prefectures in Japan) and approximately 10000 residents in total each year. Here the market with 10 doctor types and 50 regions (500 hospital types) has 5000 doctor-hospital type pairs (two residents for each pair on average), which is enough large to imitate the actual market.

For each $|X|, |Y|$, and $|Z|$, we define the populations as $n_x = 1.0/|X|$ for $x \in X$ and $m_y = 1.5/|Y|$ for $y \in Y$. We set the lower bounds for all regions; $\underline{o}_z = 0.3/|Z|$. We do not impose upper bounds on the regions. We set

$$\Phi_{xy} := 2.0 + \xi_{xy}$$

for all $x \in X, y \in Y$, where ξ_{xy} are independent noise drawn from $N(0, 1)$. We measure the time to compute the EAE 10 times for each $|X|$ and $|Z|$, and take the average of them. The result is illustrated in Figure 3. It is clear that EAE is fast enough to be used for estimation and counterfactual simulation.

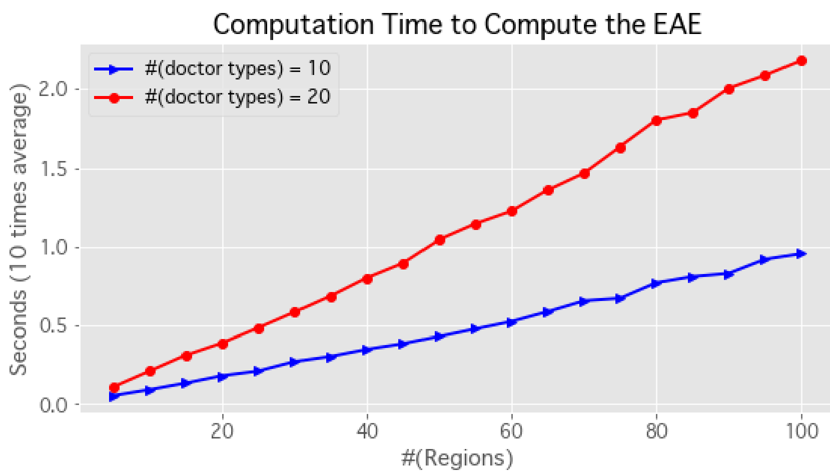

Figure 3: Average computation time (seconds) to compute the EAE (an average of 10 trials). The number of doctor types are $|X| = 10, 20$ and the number of regions are $|Z| = 5, 10, 15, \ldots, 100$. Each region has 10 hospital types.

B.4 How to Compute Alternative Equilibria

B.4.1 EAE upper-bound

Given a lower bound for $\underline{o}_z = \underline{o}_{z_2} = \underline{o}_{z_3}$, we compute EAE upper-bound as follows; we set the 41 candidates of the upper bound for z_1 as $B = \{0.10, 0.11, 0.12, \ldots, 0.50\}$, and

1. We take an upper bound for z_1, $\bar{o}_{z_1}$, from B from the smallest.
2. Compute the EAE under the regional constraint $\bar{o}_{z_1}$ (we do not impose other constraints; we set $\bar{o}_{z_2} = \bar{o}_{z_3} = \infty$ and $\underline{o}_z = 0$ for all z). Let the EAE matching be μ.
3. Check if the lower bounds for z_2 and z_3 are satisfied in the EAE; check whether both $\sum_{x \in X} \sum_{y \in z_2} \mu_{xy} \geq \underline{o}_z$ and $\sum_{x \in X} \sum_{y \in z_3} \mu_{xy} \geq \underline{o}_z$ are satisfied.
   - If they are satisfied, then the current EAE is the *EAE upper-bound*.
   - Otherwise, we take another candidate of the upper bound for z_1 from B which is one step larger.

[6] https://www.cvxpy.org/

4. Repeat the above process until we find an EAE upper-bound.

In the current setting, we successfully find an EAE upper-bound for every lower bound.

B.4.2 Cap-reduced AE

Given a lower bound for $\underline{o}_z = \underline{o}_{z_2} = \underline{o}_{z_3}$, we compute the AE without regional constraints as follows; we set the 41 candidates of the artificial capacities of the urban hospitals y_1, y_2 as $\hat{m} = \hat{m}_{y_1} = \hat{m}_{y_2} \in B = \{0.050, 0.055, 0.060, \ldots, 0.25\}$, and

1. We take an artificial capacity $\hat{m}$ from B from the smallest.
2. Compute the AE without constraints under the artificial capacities $\hat{m}_{y_1} = \hat{m}_{y_2} = \hat{m}$ by (D') setting $w_z = 0$ for all $z \in Z$ (note that we do not impose any regional constraints). Let the AE matching be μ.
3. Check if the lower bounds for z_2 and z_3 are satisfied in the AE; check whether both $\sum_{x\in X}\sum_{y\in z_2}\mu_{xy} \geq \underline{o}_z$ and $\sum_{x\in X}\sum_{y\in z_3}\mu_{xy} \geq \underline{o}_z$ are satisfied.
   - If they are satisfied, then the current AE is the *cap-reduced AE*.
   - Otherwise, we take another candidate of $\hat{m}$ from B which is one step larger.
4. Repeat the above process until we find a cap-reduced AE.

In the current setting, we successfully find a cap-reduced AE for every lower bound.

B.4.3 Budget-balanced AE

To compute the BBAE, we run the grid search; we set the candidates (grids) of taxes as

$$W = \Big\{(w_1, w_2, w_3) \in \mathbb{R}^3 \mid w_1 \in \{0, 0.5, 1.0, \ldots, 10.0\} \text{ and } w_2, w_3 \in \{-0.2, -0.19, -0.18, \ldots, 0.00\}\Big\}.$$

For each $(w_{z_1}, w_{z_2}, w_{z_3}) \in W$, we compute the AE without regional constraints by (D') setting $(w_{z_1}, w_{z_2}, w_{z_3})$. Then for each lower bound for $\underline{o}_z = \underline{o}_{z_2} = \underline{o}_{z_3}$, we select the equilibrium that maximizes the social welfare among all budget-balanced equilibria that meet the given lower bounds as BBAE; the BBAE is the solution to

$$\max_{\mu\in\hat{M}} \sum_{x\in X}\sum_{y\in Y}\mu_{xy}\left(\Phi_{xy} - w_{z(y)}\right)$$

where $\hat{M}$ is the set of all μ (the AE without constraints for some $w \in W$) that satisfies

$$\sum_{x\in X}\sum_{y\in Y}\mu_{xy}w_{z(y)} \geq 0 \quad \text{and} \quad \sum_{x\in X}\sum_{y\in z}\mu_{xy} \geq \underline{o}_z \text{ for each } z = z_2, z_3.$$