The Efficiency of Cap-Based Regulation in Residency Matching: Evidence from the Japan Residency Matching Program

Kei Ikegami Atsushi Iwasaki Akira Matsushita
Kyohei Okumura

September 4, 2024

Abstract

This paper investigates the efficiency of cap-based regulation within the Japan Residency Matching Program (JRMP) in addressing the geographic distributional imbalances of medical residents across Japan. Despite the program's aim to efficiently allocate medical students to residency programs, significant disparities persist, particularly in rural and underserved areas. We develop a theoretical model that extends the classic matching with transferable utility framework to incorporate regional caps and floors as policy constraints. Our model allows for the design of taxation policies that influence equilibrium matching outcomes while respecting these constraints. Using a novel dataset of matching outcomes from 2016 to 2019, we estimate agents' preferences and conduct counterfactual analyses. Our simulation demonstrate that a small monetary intervention and modifying the regional caps could effectively increase the number of residents in rural areas, thereby reducing geographic imbalances.

1 Introduction

Policy interventions are often necessary for matching markets to address distributional imbalances. Examples include race-based affirmative action in the United States to ensure racial diversity (Ellison and Pathak, 2021) and gender quotas in electoral systems worldwide to improve female representation (Besley et al., 2017). Matching markets must respond to the cultural and social demands of the societies in which they operate, and the incorporation of quotas and caps serves as a primary tool to achieve these goals.

In this paper, we focus on the residency matching market in Japan, a notable example of a matching market with distributional imbalances. Established in 2004, the Japan Residency Matching Program (JRMP) was modeled after the National Resident Matching Program in the United States, aiming to allocate medical students to residency programs nationwide with enhanced transparency and efficiency. However, despite these intentions, the JRMP has been criticized since its introduction for potentially causing significant geographic imbalances in the distribution of doctors, resulting in a shortage of healthcare professionals in rural and underserved areas and exacerbating healthcare access issues in these regions.

Various solutions have been proposed by both practitioners and researchers to address these imbalances. The JRMP sought to mitigate the issue by implementing the standard deferred acceptance (DA) algorithm while artificially reducing the number of positions in urban areas, aiming to match more medical students with hospitals in underserved regions. We will describe this approach in detail in Section 1.1. Kamada and Kojima (2015) proposed a more sophisticated version of the DA algorithm, which explicitly accommodates regional caps. However, these methods have two key limitations: First, they do not directly accommodate floors in underserved regions; they instead impose caps on popular regions to encourage inflow into less popular ones, which is inefficient.[1] Second, these methods fail to account for transfers within matched pairs that are determined endogenously in a market equilibrium. In JRMP, factors such as salary, benefits, and the content of medical resident training are crucial differentiators, and hospitals carefully adjust these elements in response to policy changes. Addressing distributional imbalances by merely tweaking the DA algorithm, without considering the equilibrium adjustments

[1]This use of caps partly arises from technical constraints associated with DA-based approaches. A feasible matching that satisfies the floors may not exist as long as submitting unacceptable hospitals or regions is allowed.

in transfers, risks deviating from actual market behavior.

This paper models matching markets with caps and floors, or *regional constraints*, by extending the matching with transferrable utility model introduced by Shapley and Shubik (1971). A regional constraint is an exogenously given policy goal that specifies the lower and upper bounds on the number of matches realized in each region, which might not be satisfied in the market equilibrium without policy intervention. We introduce a policymaker who designs a taxation policy to influence equilibrium matching. Given the taxation policy, the agents form a stable outcome (Shapley and Shubik, 1971). The policymaker aims to design the taxation policy under which the equilibrium matching respects the regional constraints.

In practice, the policymaker often has access only to aggregate-level data and faces *unobserved heterogeneity* with which different agents sharing the same observable characteristics are indistinguishable from the viewpoint of the policymaker. We extend the framework to the setting introduced by Galichon and Salanié (2021*a*). This extension enables counterfactual simulations under different taxation policies (Proposition 1) and the design of a welfare-maximizing taxation policy (Theorem 1) even with aggregate-level data and unobserved heterogeneity.

Aside from the issues of constraints, as a highlight of our contribution to the literature of empirical analysis of matching markets, we introduce a measurement model for the transfer between matched agents. Whereas the transfers defined in our model are not directly observed, we propose a framework for using the observed salaries to quantify all the model components monetary-based. Due to the practical difficulty in measuring the transfers directly and the relative feasibility of observing salaries, our method broadens the scope of the empirical analysis of matching markets.

We conduct an empirical analysis of the JRMP using a novel dataset of realized matching outcomes over three years, along with the characteristics of schools and hospitals. We estimate the preference parameters of agents on both sides of the market. For doctors, our estimates align with the existing literature, showing that factors such as the distance between the hospital and their schools mater, as well as the size of the hospital, are significant. Additionally, we find that the number of previous matches is an important determinant of doctors' preferences. On the hospital side, we observe that hospitals exhibit horizontal preferences similar to those of doctors: doctors from distant regions are less preferred, while quality measures of doctors are estimated to increase the surplus of the matching hospital, consistent with existing findings.

Based on the estimates, we simulate matching markets under various settings to identify the source of the inefficiency of the current JRMP market. In our simulation, we select 15 prefectures as regions requiring lower bounds on the number of matches in the market and simulate the matching patterns under optimal subsidy levels. Our results suggest that both doctors and hospitals experience welfare gains under the optimal subsidy policy. Specifically, a subsidy of approximately $40,000$ (JPY) per match per month in the target prefectures is sufficient to achieve the lower bounds. Furthermore, we identify that the welfare gains primarily stem from increasing caps in urban prefectures. This implies that the current regional caps set by the JRMP are far from the optimal ones.

The rest of this paper is organized as follows. Section 1.1 provides the institutional background of the JRMP. Section 2 details the data sources, key variables, and descriptive statistics. Section 3 introduces the theoretical model, extending the classic matching with transferable utility framework to incorporate regional caps and floors as policy constraints. Section 4 presents the main theoretical results, including the design of optimal taxation policies under regional constraints. Section 5 outlines the empirical strategy, describing the estimation of structural parameters and the construction of moment conditions. Section 6 reports the results of our Monte Carlo simulations. Section 7 provides the results of our estimation. Section 8 conducts counterfactual analyses to evaluate the efficiency of the current policy and the potential impact of alternative policies.

1.1 Institutional Background of the JRMP

Established in 2004, the Japan Residency Matching Program (JRMP) is modeled after the National Resident Matching Program in the United States. The JRMP uses a deferred acceptance algorithm to match medical students seeking clinical training with hospitals offering residency programs based on mutual preferences. Typically, sixth-year medical students who plan to take the national exam participate in the program.[2]

The JRMP process is structured as follows: Students and hospitals must register for the system and submit their preferences. Students have access to information such as hospital size, location, training program details, salary, and workload, and they can parti-

[2] In Japan, students can enter medical school immediately after high school, so sixth-year students are typically 24-25 years old.

cipate in job fairs and visit hospitals for more details. Before submitting their preference lists, students must take exams conducted by each hospital they wish to list. After the initial submission of preference lists, the distribution of students' first-choice hospitals is disclosed once, allowing both students and hospitals to adjust their preferences before finalization. The deferred acceptance algorithm then runs to determine the matches based on the finalized preference lists.[3] Unmatched students can reach out to hospitals with vacant slots individually or wait to participate in the matching process the following year. In 2023, the JRMP saw the participation of 10,202 students and 1,209 hospitals offering 10,895 positions. The algorithm successfully matched 87.9% of the students, with 64.3% securing their first-choice hospitals, 16.3% their second choice, and 9.0% their third choice.

Although internships were not mandatory before 2004, most medical students still undertook them, typically at hospitals affiliated with their medical schools. For example, in 2001, 71.2% of students worked in such affiliated hospitals. These hospitals often offered poor financial compensation, forcing many students to take part-time jobs to cover their living expenses. Since 2004, two-year internships have become mandatory for becoming clinicians. The government began subsidizing hospitals offering residency programs, ensuring that residents receive sufficient salaries and eliminating the need for part-time work.[4] Consequently, an increasing number of students are now choosing to work in non-university hospitals, with 69.27% doing so between 2017 and 2019.[5](Kitamura and Takagi, 2006)

Many argue that the introduction of the JRMP has contributed to significant geographic imbalances in the distribution of doctors. A widely accepted opinion is that when most doctors were affiliated with universities, the universities had strong control over doctor placements and could send some of them to underserved areas. However, under the JRMP, more doctors complete their internships in urban and non-university hospitals, loosening the connections between doctors and universities and preventing uni-

[3]At this stage, it is nearly impossible for students or hospitals to add more options to their preference lists. Given the (one-sided) strategy-proofness of the DA algorithm, the usefulness of this information disclosure is questionable.

[4]Another primary goal of the new internship system is to introduce a rotating internship system, allowing residents to gain a broader understanding of primary care. Before 2004, residents typically concentrated on their specific areas of interest without receiving comprehensive training in other fields.

[5]The popularity of non-university hospitals may be attributed to better working environment, greater opportunities for practical experience with common diseases, and increased flexibility in location.

versities from functioning as an adjustment mechanism.[6] To compensate for the lack of residents as a labor force, university hospitals often need to request doctors working at other related non-university hospitals to return, further reducing the number of doctors in underserved areas.[7] Additionally, hospitals participating in the JRMP are required to have enough doctors to supervise residents, leading to a concentration of doctors in certain hospitals. (Endo, 2019)

To address distributional imbalances, the JRMP began implementing regional caps in 2010.[8] The regional caps are set through a systematic process: first, the total number of positions in the country is determined by multiplying the total number of medical students by a constant. This constant was approximately 1.22 in 2015, but the government plans to reduce it to 1.05 by 2025. Once the total number of positions is determined, they are allocated to each prefecture based on variables such as population, medical school capacity, current number of doctors, and geographic factors like the number of doctors per unit area and the population of isolated islands. The allocation rule is designed to favor underserved areas, with positions in urban areas being reduced more significantly. The rationale behind these caps is that by tightening the total capacity and limiting positions in urban areas, more students will secure placements in rural regions and remain there after their internships.

2 Data

Our analysis covers the four years of matching results generated by the JRMP from 2016 to 2019. To estimate our model, we need three key elements: the matching patterns

[6]Due to the competitive nature of medical schools, many students from urban areas opt to attend medical schools in rural regions. These students are more likely to choose to work in hospitals near their home areas.

[7]University hospitals want to have many medical staff members for both clinical and research purposes, but financial constraints make this difficult. Therefore, they send staff to non-university hospitals on temporary assignments, with the non-university hospitals paying their salaries. Non-university hospitals accept this arrangement because it provides a stable supply of doctors, or they request such arrangements due to doctor shortages. Although interns lack sufficient knowledge and skills to be effective in underserved areas, they can handle some tasks at university hospitals, reducing the workload of mid-level doctors and enabling these doctors to be dispatched elsewhere. Consequently, a decrease in the number of residents restricts the dispatch of mid-level doctors.

[8]Another possible approach could be to simply increase the number of doctors. However, the Japanese government opted not to pursue this strategy and instead set caps on the total number of medical students nationwide. This decision was based on the expectation that Japan's population will decrease in the coming years, potentially leading to an oversupply of doctors. An oversupply could result in physician-induced demand and increased public expenditure on health insurance.

between medical schools and hospitals, the characteristics of these institutions relevant to the preferences of medical students and hospitals, and the salaries paid to residents during their internships. We begin by explaining the data sources and then present the descriptive statistics in Section 2.1.

The matching patterns between medical schools and hospitals (i.e., the number of matches between any given pair) are derived from the "Physician Registration Report," which includes information such as doctors' registration numbers, workplace postal codes, and the universities from which they graduated. This data allows us to determine the annual number of matches between specific medical schools and hospitals.

We obtained the characteristics of hospitals from the JRMP website, which provides details such as hospital names, program offerings, and capacity. For the characteristics of medical schools, we used the national exam pass rate and whether the university is public, based on publicly available information from hospital websites. Additionally, to measure the expected ability of graduates from a medical school, we used the T score of the entrance exam.[9] T scores are widely recognized as an indicator of university entrance exam difficulty in Japan, with higher scores indicating more challenging universities. We used the most recent T scores available for our estimation.[10]

Finally, we gathered salary data by crawling hospital websites. Due to limited data availability, we used the most recent salary information rather than data from 2016 to 2019, assuming that salary levels remained constant during this period. We also collected additional hospital-related information, such as location, number of beds, and emergency transport cases.

2.1 Descriptive Statistics

Table 1 summarizes the environment and the outcomes of JRMP for the years 2017, 2018, and 2019. Since our estimation uses the matching patterns from the final year as one of the covariates, we focus on these three years. Panel A and Panel B in Table 1 show the results of JRMP. While the environment, including the total number of medical students and the available seats, and their ratios, exhibits minor yearly variations, the matching

[9] The T scores of universities are published by cram schools. These scores are calculated using data from practice exams administered by the cram schools, which gather information on students' actual university entrance exam results. The T scores reflect the relationship between students' practice exam performance and their success in university entrance exams.

[10] The data source is `https://www.keinet.ne.jp/university/ranking/`.

Table 1. Environments and Outcomes of JRMP

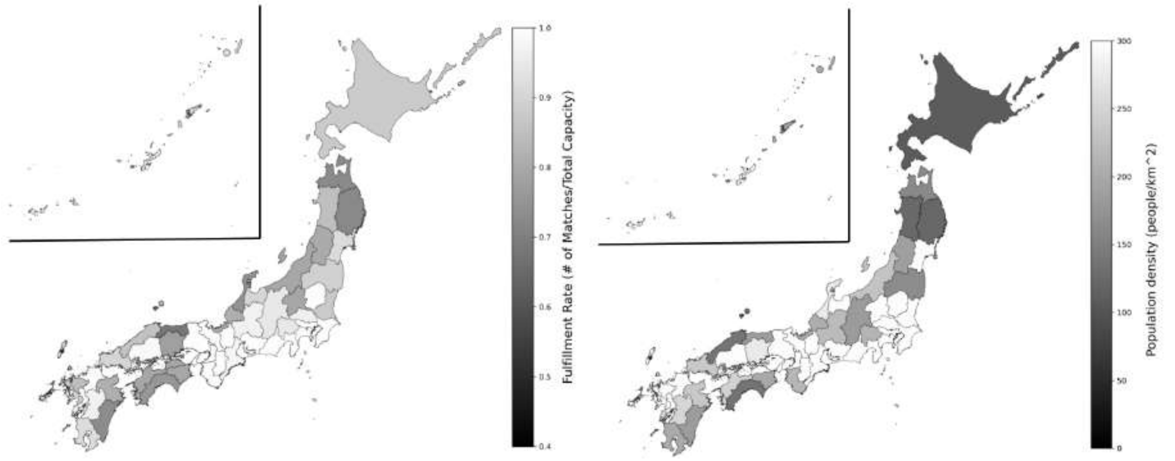
	2017	**2018**	**2019**
Panel A: Doctor side			
Number of universities	78	78	78
Number of students	9830	9916	9932
Number of matched students	8530	8369	8634
Number of unmatched students	1300	1547	1298
Unmatch rate (%)	13.22	18.48	15.03
Panel B: Hospital side			
Number of hospitals	1025	1025	1022
Number of total seats	11716	11468	11730
Number of matched seats	8530	8369	8634
Number of unmatched seats	3186	3099	3096
Unmatch rate (%)	27.19	27.02	26.39
Number of excess seats	1886	1551	1798
Excess rate (%)	16.01	13.52	15.33
Panel C: Salaries (thousand yen per month)			
Average salary	386.4	386.4	386.1
Standard deviation of salaries	99.5	99.6	99.4
Min salary	180.0	180.0	180.0
Max salary	855.0	855.0	855.0

outcomes—such as the number of matches, unmatches, and the unmatch rate—fluctuate over this period. As noted in Section 1.1, the unmatch rate had been increasing before this period due to the tightening caps in urban areas. Although the unmatch rate remains high, our data cover a relatively stationary market.

Despite the presence of unoccupied seats overall, as shown in Table 1, the fulfillment rate by prefecture, defined as the ratio of the number of matches to the total number of positions in each prefecture, exhibits substantial regional variation. Figure 1a displays a choropleth map of fulfillment rates, while Figure 1b illustrates population densities by prefecture in 2019. As noted in Section 1.1, these figures suggest that rural regions, characterized by lower population density, are less popular and suffer lower fulfillment rates than urban areas.

Panel C of Table 1 reports significant variation in the salaries of medical interns across Japan. For example, in 2017, the average annual salary was approximately \$31,714, with a standard deviation of \$8,166.[11] For comparison, Table 1 of Agarwal (2015) reports

[11] The conversion from yen to dollars was based on the exchange rate as of August 30, 2024.



(a) Fulfillment Rates (b) Population Density

Figure 1. Fulfillment Rate and Population Density

that the mean salary for similar medical interns in the United States was \$47,331, with a standard deviation of \$2,953. While average salaries are higher in the U.S., the standard deviation in Japan is 2.77 times larger, indicating greater salary dispersion.

Table 2 presents descriptive statistics for the covariates used to parameterize social surplus and the utility of medical students and hospitals. The T score and graduation exam pass rate of medical schools serve as proxies for student ability, with higher values being more desirable to hospitals. Public medical schools exhibit higher average T scores, indicating that graduation from a public school signals greater ability. This difference is statistically significant. On the hospital side, we define an indicator for urban hospitals located in the six prefectures with officially set caps on the number of matches in the JRMP: Tokyo, Kanagawa, Aichi, Kyoto, Osaka, and Fukuoka. Hospital size is proxied by the number of beds, with Table 2 showing that university hospitals and those in urban areas tend to be larger.

Public medical schools in Japan play a pivotal role in maintaining standardized medical services nationwide, with a significant proportion of their graduates continuing to nearby hospitals for internships. Figure 2 illustrates the matching pattern between hospitals and medical schools in 2017, where each cell represents a pairing, with black cells indicating at least one match. Hospitals are indexed according to the official prefectural index set by the Japanese government, ordered by location, and within each prefecture, by latitude. Medical schools are first divided into public and private groups, with the same ordering method applied within each group.

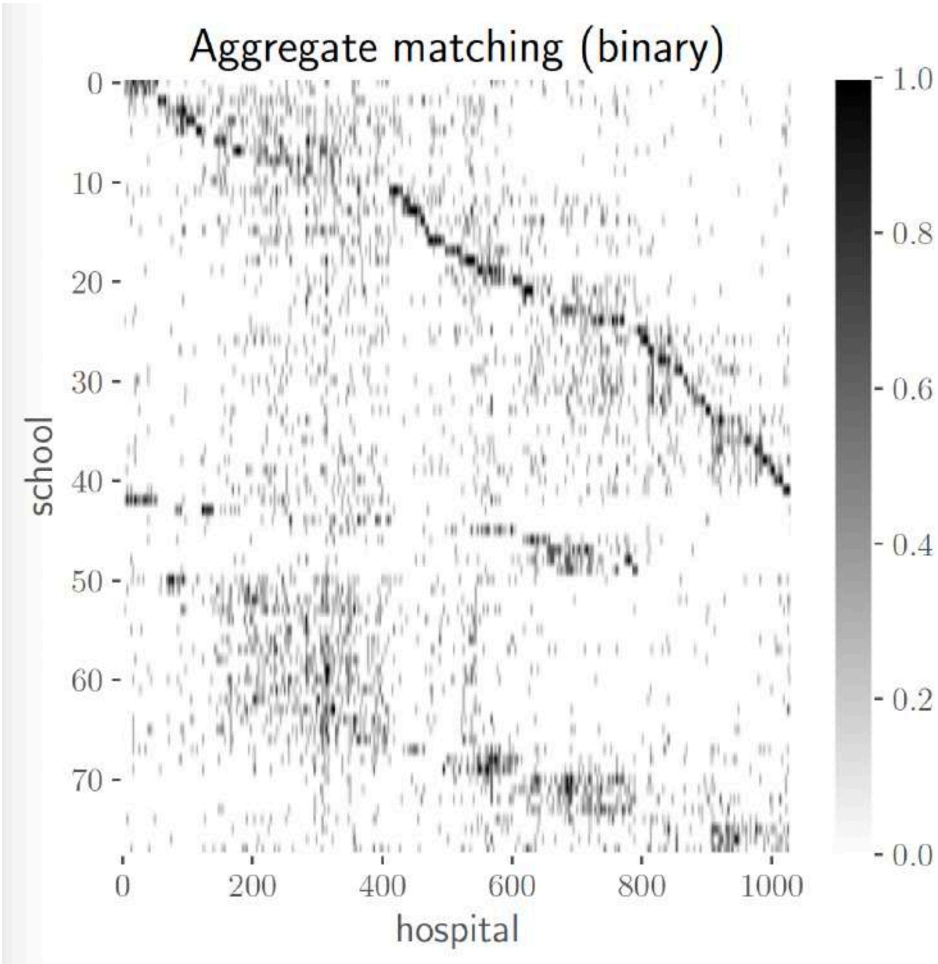


Figure 2. Binarized Matching Patterns between All Schools and Hospitals in 2017

Table 2. Summary Statistics of Covariates

	Count	**Mean**	**Std**	**Min**	**Max**
T score					
Private University	27	64.96	2.17	62	72
Public University	51	66.38	2.66	63	74
Exam pass rate					
Private University	27	0.93	0.05	0.79	1.00
Public University	51	0.93	0.03	0.82	1.00
Number of Beds					
Non University hospital	911	411.08	147.73	36	1097
University hospital	121	626.74	272.77	295	1379
Rural Hospital	684	417.99	155.92	36	1195
Urban hospital	348	472.48	217.84	38	1379

The upper half of Figure 2 (up to index 50) represents the matchings of public schools, while the lower half corresponds to private schools. Both public and private schools generally tend to match with nearby hospitals; however, this tendency is more pronounced for public schools, as reflected in the upward slope of the figure. Consequently, public school graduates are, on average, less likely to match with urban hospitals despite the higher average quality of public universities.

2.2 Strategic Nature of Salary

The matching patterns observed above cannot be attributed solely to the geographic preferences of hospitals and schools; another plausible explanation is that rural hospitals offer more attractive conditions to draw students from public universities.[12] In this section, we examine how salary functions as a compensating differential to attract more applicants to rural areas and explore evidence suggesting that salary should be treated as an endogenous or strategic variable in our model.

We conduct a series of regressions to investigate (i) which types of hospitals attract better students and (ii) which types of students are matched with better hospitals, similar to Agarwal (2015). Due to the unavailability of individual matching data, we first calculate the weighted averages of the characteristics of matched partners within each

[12] As discussed in Section 1.1, the government considers the excessive concentration of medical professionals in urban areas to be a serious issue, largely driven by student preferences for urban locations.

Table 3. Sorting between Residents and Programs from Program Viewpoint

	(1) Public	(2) T score	(3) Pass rate
ln(Beds)	0.0665***	7.060***	0.0990***
	(0.0225)	(1.242)	(0.0171)
ln(Wage)	0.00233	-3.459*	-0.0376
	(0.0364)	(1.974)	(0.0276)
University hospital	-0.282***	-1.704	-0.0150
	(0.0343)	(1.116)	(0.0155)
Urban	-0.110***	3.399***	0.0435***
	(0.0178)	(0.863)	(0.0120)
N	3096	3096	3096
FE	√	√	√

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

medical school and hospital. These aggregated characteristics are then regressed on the covariates of the medical schools and hospitals.

Table 3 presents the regression results on sorting from the perspective of hospitals. As depicted in Figure 2, urban hospitals have more students from private schools, as shown in column (1). Additionally, columns (2) and (3) indicate that hospitals located in urban areas and those with a larger number of beds are more likely to attract students from higher-quality schools. Given that urban and larger hospitals are inherently more desirable, positive assortative matching appears to be realized for these two measures in contrast to the case of indicator of public university.

Table 4 presents the regression results from the perspective of medical schools. Columns (1), (2), and (4) indicate that students from public schools are more likely to be matched with smaller hospitals in rural areas, where salaries tend to be higher than in urban hospitals. This finding supports the hypothesis that rural hospitals use higher salaries to attract more applicants, particularly those from nearby public universities.

Figure 3 provides further evidence for this hypothesis by visualizing hospital popularity, measured by the competition ratio, alongside monthly salaries on a choropleth map. As shown in Panel (a), applicants are concentrated in urban areas such as Tokyo and

Table 4. Sorting between Residents and Programs from Residents Viewpoint

	(1) Beds	(2) Univ. Hospital	(3) Urban Hospital	(4) ln Wage
Public University	-153.1***	-0.309***	-0.302***	0.136***
	(17.60)	(0.0334)	(0.0438)	(0.0250)
T score	2.935	-0.0145**	0.0427***	-0.00932**
	(3.056)	(0.00685)	(0.00888)	(0.00377)
Exam Pass Rate	-47.03	0.0945	0.0256	0.0563
	(58.28)	(0.149)	(0.114)	(0.0724)
N	234	234	234	234
FE	√	√	√	√

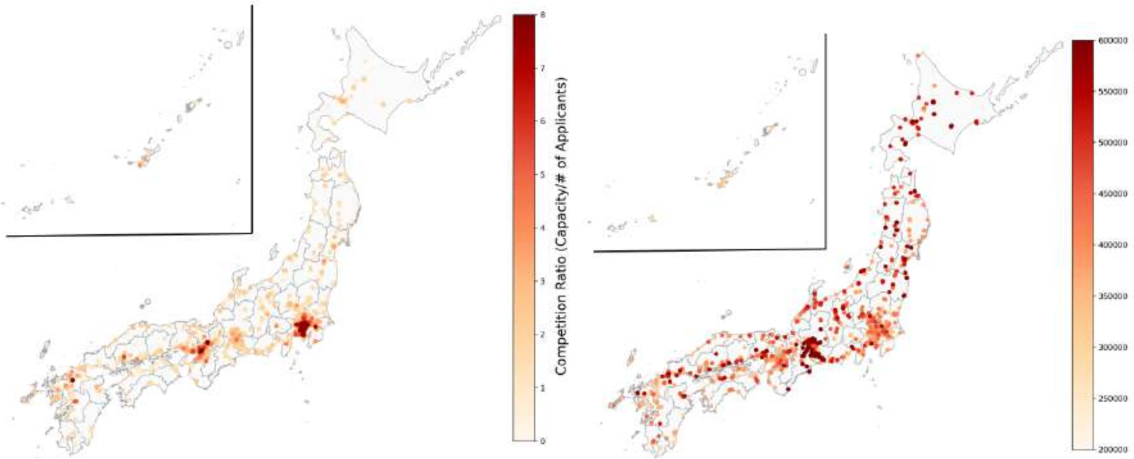
Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Osaka. In contrast, Panel (b) indicates that salaries in these urban areas are not particularly high. Instead, rural hospitals, which are less popular and appear lighter in Panel (a), tend to offer more competitive salaries. This observation suggests that salary serves as a key differentiator in this market, indicating a strategic use of compensation that deviates from the traditional non-transferable utility (NTU) matching market model.

3 Model

Stable outcome under a taxation policy We consider a two-sided matching market. Let I denote the set of doctors (medical students) and J the set of job slots owned by hospitals.[13] Each doctor $i \in I$ can be matched with at most one slot $j \in J$, and each slot can accommodate at most one doctor. If a doctor i is unmatched, they are paired with an outside option j_0. Similarly, an unmatched slot j is paired with an outside option i_0. A *matching* is represented by a 0-1 matrix $d = (d_{ij})_{i\in I, j\in J}$, where $d_{ij} = 1$ if and only if doctor i is matched with slot j. A matching d is *feasible* if each doctor is matched to exactly one slot or the outside option, and each slot is matched to exactly one doctor or the outside option: $\sum_{j\in J} d_{ij} \leq 1$ for all $i \in I$, and $\sum_{i\in I} d_{ij} \leq 1$ for all $j \in J$.

[13]Here, $j \in J$ refers to a job slot within a hospital. Each hospital can have multiple job slots, and aims to maximize the aggregate payoffs from these slots, implying that hospitals' preferences are *responsive* (Roth and Sotomayor, 1990).



(a) Competition Ratio (b) Monthly Salary (JPY)

Figure 3. Competition Ratio and Salary Differentials across Hospitals

There is a policymaker who faces an additional condition called regional constraints. There are L regions, denoted by $Z = \{z_1, z_2, \dots, z_L\}$, with each job slot j assigned to one region. Additionally, we define a special region z_0 that contains only the outside option j_0. With a slight abuse of notation, let $z\colon J \cup \{j_0\} \to Z \cup \{z_0\}$ be the mapping where $z(j)$ indicates the unique region to which job slot j belongs. Each region z has a cap and a floor, $\underline{o}_z \in \mathbb{R}_+$ and $\bar{o}_z \in \mathbb{R}_+ \cup \{\infty\}$, respectively. We say a feasible matching d satisfies *regional constraints* if it respects the caps and the floors: $\sum_{i \in I} \sum_{j \in z} d_{ij} \in [\underline{o}_z, \bar{o}_z]$ for each $z \in Z$. Throughout the paper, we assume that there exists at least one feasible matching that satisfies regional constraints.

Agents form a stable outcome à la Shapley and Shubik (1971). Without policy intervention, the realized matching may not meet the regional constraints. The policymaker can implement a taxation policy that influences the split of the joint surplus among agents to satisfy the regional constraints. When a doctor i and a slot j are matched, they generate an *(individual-level) net joint surplus* $\Phi_{ij} \in \mathbb{R}$. The tax $w_z \in \mathbb{R}$ is imposed on each match (i, j) in region z with negative taxes being interpreted as subsidies. We assume $w_{z_0} = 0$, i.e., no tax is imposed on the outside option. With taxation policy w, each matched pair divides the *gross joint surplus* $\Phi_{ij} - w_{z(j)}$ instead of the net joint surplus. [14] The stable outcome under a taxation policy is defined as follows:

[14]In principal, the policymaker may want to impose different amounts of taxes on distinct pairs within the same region. Although we exclude such possibilities in the definition of taxation policy, we can show that such a restriction is harmless in terms of social welfare: there is a welfare-maximizing taxation policy that imposes the same amount of tax on all the pairs in the same region.

Definition 1 (Stable outcome). *Given the matching market* (I, J, Z, z, Φ),[15] *a profile* $(d, (u, v))$ *of feasible matching* d *and equilibrium payoffs* (u, v) *forms a* stable outcome *under taxation policy* w *if it satisfies:*

1. *Individual rationality: For all* $i \in I$, $u_i \geq \Phi_{i,j_0}$, *with equality if* i *is unmatched. For all* $j \in J, v_j \geq \Phi_{i_0,j}$, *with equality if* j *is unmatched.*

2. *No blocking pairs: For all* $i \in I$ *and* $j \in J$, $u_i + v_j \geq \Phi_{ij} - w_{z(j)}$, *with equality if* $d_{ij} = 1$.

We say d *is a* stable matching *if there exists* (u, v) *such that* $(d, (u, v))$ *forms a stable outcome.*

Unobserved heterogeneity Let $X = \{x_1, x_2, \ldots, x_N\}$ represent the finite set of observable characteristics, or *types*, of doctors. Each doctor $i \in I$ has a type $x(i) \in X$. Similarly, let $Y = \{y_1, y_2, \ldots, y_M\}$ represent the finite set of observable characteristics of job slots, with each slot $j \in J$ having a type $y(j) \in Y$. Although agents with the same type are indistinguishable to the policymaker, there can be *unobservable heterogeneity*: doctors of the same type x or job slots of the same type y may generate different joint surpluses when matched. For convenience, we denote $i \in x$ if $x(i) = x$ and $j \in y$ if $y(j) = y$. We define x_0 and y_0 as special types representing the outside options i_0 and j_0, respectively, and let $X_0 = X \cup \{x_0\}$ and $Y_0 = Y \cup \{y_0\}$ include these outside options. The set of all type pairs is denoted by $T = X_0 \times Y_0 \setminus \{(x_0, y_0)\}$. We assume each job slot type $y \in Y$ belongs to a unique region, denoted by $z(y) \in Z$. Let n_x be the number of doctors with type x, and m_y be the number of job slots with type y.

Let μ_{xy} denote the number of matches between type-x doctors and type-y job slots, defined as $\mu_{xy} = \sum_{i \in x} \sum_{j \in y} d_{ij}$. An *aggregate-level matching* $\mu = (\mu_{xy})_{x \in X, y \in Y}$ is said to be *feasible* if it satisfies the population constraints $\sum_y \mu_{xy} = n_x$ and $\sum_x \mu_{xy} = m_y$ for each x and y. Furthermore, we say μ satisfies regional constraints if $\sum_{y \in z} \sum_{x \in X} \mu_{xy} \in [\underline{o}_z, \bar{o}_z]$ for each $z \in Z$.

3.1 Discussion

Interpretation of joint surpluses The joint surplus includes not only the revenue generated by the match but also other potential gains such as experience and knowledge.

[15]The symbol z denotes the mapping from job slots to regions.

Thus, Φ_{ij} represents the value of all such potential gains, measured in a numeraire.

Validity of stable outcome It is reasonable to assume that participants form a stable outcome in various matching markets such as in frictionless decentralized matching markets (e.g., certain labor markets or marriage markets.[16]) Additionally, we can show that agents may form a stable outcome in a game where (i) hospitals first set wages, (ii) agents submit their preference lists after observing wages, and (iii) the matching is determined by the deferred acceptance algorithm, which is a good approximation of the JRMP (see Appendix A.5 for more details.)

Interpretation of types and regions Types $y \in Y$ and regions $z \in Z$ can be interpreted in various ways. For example, in our application, a type y may correspond to a hospital, and a region z may correspond to a district (e.g., a prefecture). In other contexts, a type could represent a subcategory of occupation (e.g., registered nurse, physician assistant), and a region could represent a broader occupational category (e.g., healthcare).

4 Theoretical Results

Discrete choice representation We here review a set of conditions and results, developed by Galichon and Salanié (2021*a*), that connect the individual-level objects $((\Phi_{ij})_{ij}, (d_{ij})_{ij}, (u_i)_i, (v_j)_j)$ introduced in Section 3 to the aggregate-level objects we develop in this section.

For a pair (i, j) with $i \in x$ and $j \in y$, we assume the individual-level joint surplus Φ_{ij} can be decomposed into the sum of the *aggregate-level joint surplus* Φ_{xy} and independent mean-zero error terms ε_{iy} and η_{xj}. [17] For each x and $i \in x$, error term $(\varepsilon_{iy})_{y \in Y_0}$ is drawn from the distribution $P_x \in \Delta(\mathbb{R}^{|Y|+1})$. Similarly, for each y and $j \in y$, error term $(\eta_{xj})_{x \in X_0}$ is drawn from distribution $Q_y \in \Delta(\mathbb{R}^{|X|+1})$.

Assumption 1 (Independence)**.** *The error terms are independent across all i and j and of mean-zero.* [18]

[16] See, for example, Roth and Sotomayor (1990) and Chiappori (2017) for textbook references.
[17] Under Assumption 1 and 2, Φ_{xy} is the average joint surplus conditional on $i \in x$ and $j \in y$.
[18] The mean-zero assumption here is without loss of generality. If not, we can always demean the error terms and redefine them.

Assumption 2 (Additive Separability)**.** *There is a matrix* $(\Phi_{xy})_{(x,y)\in T}$ *such that (i)* $\Phi_{ij} = \Phi_{xy} + \varepsilon_{iy} + \eta_{xj}$ *for each* $x \in X$*,* $y \in Y$*,* $i \in x$*, and* $j \in y$*, and (ii)* $\Phi_{i,y_0} = \varepsilon_{i,y_0}$*,* $\Phi_{x_0,j} = \eta_{x_0,j}$ *for each* $x \in X$ *and* $y \in Y$*.*

We define *aggregate-level utilities* U_{xy} and V_{xy}, which represent the equilibrium payoffs dependent solely on types, as follows: for each x and y, define

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

and $U_{x,y_0} = V_{x_0,y} \coloneqq 0$. The following lemma argues that the matching market can be seen as a bilateral discrete choice problem.

Lemma 1 (Galichon and Salanié (2021*a*))**.** *Let* (u, v) *be a payoff profile in a stable outcome. Under Assumption 2, for any doctor* $i \in I$ *and any slot* $j \in J$*, we have*

$$u_i = \max_{y\in Y_0} \left\{U_{x(i),y} + \varepsilon_{iy}\right\}, \quad v_j = \max_{x\in X_0} \left\{V_{x,y(j)} + \eta_{xj}\right\}.$$

Proof. See Appendix A.1. □

By Lemma 1, the social welfare, defined as the sum of the equilibrium payoffs, on the doctor side can be written as $\sum_{i\in I} u_i = \sum_{x\in X} n_x \cdot \frac{1}{n_x} \sum_{i\in x} \max_{y\in Y_0} \{U_{xy} + \varepsilon_{iy}\}$. When n_x is sufficiently large, the term $\frac{1}{n_x} \sum_{i\in x} \max_{y\in Y_0} \{U_{xy} + \varepsilon_{iy}\}$ can be approximated by $\mathbb{E}_{\varepsilon_i\sim P_x} [\max_{y\in Y_0} \{U_{xy} + \varepsilon_{iy}\}]$. We will assume from now on that this large market limit is a good approximation, so the social welfare on the doctor side becomes

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

Similarly, when m_y is sufficiently large for each y, the welfare on the hospital side is approximated by

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

Under the assumption that the CDFs of the error terms are continuously differentiable, we have (the Williams-Daly-Zachary theorem (McFadden, 1980)):

$$\frac{\partial G}{\partial U_{xy}}(U) = \frac{\partial}{\partial U_{xy}} \mathop{\mathbb{E}}_{\varepsilon_i\sim P_x} \left[\max_{y\in Y_0} \{U_{xy} + \varepsilon_{iy}\}\right] = \Pr(i \text{ chooses } y \mid i \in x).$$

When n_x is sufficiently large, $n_x \Pr(i \text{ chooses } y \mid i \in x)$ is a good approximation of μ_{xy}.

Assumption 3 (Large Market Approximation)**.** *For each $x \in X$, we approximate $\mathbb{E}_{\varepsilon_i \sim P_x}[\max_{y \in Y_0}\{U_{xy} + \varepsilon_{iy}\}]$ as $\frac{1}{n_x}\sum_{i \in x} \max_{y \in Y_0}\{U_{xy} + \varepsilon_{iy}\}$. We also approximate μ_{xy} as $n_x \Pr(i \text{ chooses type-}y \text{ slot} \mid i \in x)$. Similar conditions are assumed for type $y \in Y$.*

Assumption 4 (Smooth Distribution)**.** *For each x, y, CDF's P_x and Q_y are continuously differentiable.*

Policymaker's problem A tuple $\mathcal{M} := (X, Y, n, m, Z, z, \Phi, P, Q)$ characterizes an aggregate-level matching market, where $n := (n_x)_x$, $m := (m_y)_y$, $\Phi = (\Phi_{xy})_{xy}$, $P = (P_x)_x$, and $Q = (Q_y)_y$. Given $\mathcal{M}$, the policymaker aims to (i) compute the optimal taxation policy w that maximizes social welfare while respecting regional constraints, and (ii) compute the matching and social welfare for different taxation policies. For these goals, we will construct optimization problems that only requires knowledge of the aggregate-level surplus $(\Phi_{xy})_{x,y}$. To this end, we need an additional technical assumption on the error term distributions:

Assumption 5 (Full support)**.** $\operatorname{supp}(P_x) = \mathbb{R}^{|Y_0|}$ *and* $\operatorname{supp}(Q_y) = \mathbb{R}^{|X_0|}$ *for each x and y.*

This assumption guarantees that G and H are strictly convex,[19] so there is a one-to-one correspondence between an aggregate-level matching μ and aggregate-level utilities U (and similarly between μ and V.)[20] The optimization problem is now defined as follows:

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

[19] See Appendix A.3 for the proof.

[20]If G and H strictly convex, the Legendgre transform of G and H, denoted by G^* and H^*, are differentiable (Proposition D.14 of Galichon (2018).) Then, we have $\mu \in \frac{\partial G}{\partial U}(U)$ iff $U \in \frac{\partial G^*}{\partial \mu}(\mu)$ (Proposition D.13 of Galichon (2018).)

where $\mathcal{E}(\mu) \coloneqq -G^*(\mu) - H^*(\mu)$, and G^* and H^* are the Legendre-Fenchel transform of G and H, respectively.[21] We can show that (P) is a concave programming that maximizes social welfare subject to regional constraints (see Appendix A.4.) Its dual problem is

$$
\text{(D)} \quad \left| \quad
\begin{aligned}
&\underset{U,V,\bar{w}_z,\underline{w}_z}{\text{minimize}} && G(U) + H(V) + \sum_{z\in Z} \bar{o}_z \bar{w}_z - \sum_{z\in Z} \underline{o}_z \underline{w}_z && \\
&\text{subject to} && U_{xy} + V_{xy} \geq \Phi_{xy} - \bar{w}_{z(y)} + \underline{w}_{z(y)} && \forall (x,y) \in T, \\
& && \bar{w}_z \geq 0,\ \underline{w}_z \geq 0 && \forall z \in Z.
\end{aligned}
\right.
$$

The primal problem (P) has the optimal solution since its objective function is continuous and its feasible set is compact and non-empty. Due to strong duality,[22] the dual problem (D) also has the optimal value, which coincides with that of (P). Moreover, (P) and (D) have unique solutions (see Appendix A.2.) The following theorem claims that, given regional constraints $(\bar{o}_z, \underline{o}_z)_z$, the optimal taxation policy w and a tuple of aggregate-level matching and utilities (μ, U, V) are characterized by the solutions to (P) and (D).

Theorem 1. *Assume that Assumptions 1-5 are satisfied. Fix any aggregate-level matching market $\mathcal{M}$. Suppose that μ and $(U, V, \bar{w}, \underline{w})$ are the solutions to (P) and (D), respectively.*

1. *Let $w_z \coloneqq \mathbb{1}\{\bar{w}_z > 0\}\bar{w}_z - \mathbb{1}\{\underline{w}_z > 0\}\underline{w}_z$ for each z. Then, w is the optimal taxation policy, and (μ, U, V) are the corresponding aggregate-level matching and utilities.*

2. *Suppose that w is the optimal taxation policy. Then, we have $\bar{w}_z \coloneqq \mathbb{1}\{w_z > 0\}w_z$ and $\underline{w}_z \coloneqq -\mathbb{1}\{w_z < 0\}w_z$.*

Proof. See Appendix A.3. □

We can also compute a tuple of aggregate-level matching and utilities (μ, U, V) realized under taxation policy w for any w by solving a pair of optimization problems. The proof of the following proposition is similar to the one for Theorem 1, and so is omitted.

Proposition 1. *Assume that Assumptions 1-5 are satisfied. Fix any aggregate-level matching market $\mathcal{M}$. For any taxation policy w, the aggregate-level matching and utilities realized under w, denoted by $(\mu(w), U(w), V(w))$, is characterized by the solution to (P_w) and (D_w), defined as follows:*

[21] G and H are proper convex functions, so their Legendre-Fenchel transforms are well-defined.

[22] The constraints of the primal problem satisfy the weak Slater's condition (all functions are affine in μ.)

$$
(P_w) \quad \left| \begin{array}{lll} \underset{\mu \geq 0}{\textit{maximize}} & \displaystyle\sum_{(x,y)\in T} \mu_{xy}\left(\Phi_{xy} - w_{z(y)}\right) + \mathcal{E}(\mu) & \\ \textit{subject to} & \displaystyle\sum_{y\in Y_0} \mu_{xy} = n_x & \forall x \in X, \\ & \displaystyle\sum_{x\in X_0} \mu_{xy} = m_y & \forall y \in Y. \end{array} \right.
$$

$$
(D_w) \quad \left| \begin{array}{lll} \underset{U,V}{\textit{minimize}} & G(U) + H(V) & \\ \textit{subject to} & U_{xy} + V_{xy} \geq \Phi_{xy} - w_{z(y)} & \forall (x,y) \in T \end{array} \right.
$$

For taxation policy w, tuple $(\mu(w), U(w), V(w))$ defined in Proposition 1 is called an *aggregate equilibrium (AE)* under w. It is called the *efficient aggregate equilibrium (EAE)* when the corresponding w is the optimal taxation policy.

5 Empirical Strategy

In this section, we begin by mapping the primitives and equilibrium objects in the model described in Section 3 and 4 to their empirical counterparts in the doctor-hospital matching market. Next, we introduce the concept of transfers between matched pairs within the matching market. To define these transfers, we impose an additional structure on the composition of the net joint surplus. Finally, we outline our empirical strategy to estimate the structural parameters. Our estimation consists of two steps, which we explain in sequence. The first step, detailed in Section 5.2.1, follows Galichon and Salanié (2021*a*) to identify aggregate-level utilities from the observed aggregate matching. The second step, described in Section 5.2.2, utilizes a new moment condition that links the observed salary to the model-induced objects to estimate the parameter values.

We use the following notation: a doctor is denoted by i, and a job slot is denoted by j. Each doctor belongs to a medical school, and each slot is offered by a hospital, with s representing a school and h representing a hospital. We consider s and h as observable types of doctors and slots, using $s(i)$ to denote the medical school to which doctor i belongs, and $h(j)$ to denote the hospital offering slot j. The matching market operates over $T \in \mathbb{N}$ periods, with t denoting each observation period. Let Z denote the set of regions and $z(h)$ denote the region to which hospital h belongs, assuming $z(h)$ remains constant over time. The aggregate-level joint surplus at time t is denoted by Φ_{sht}. The unobserved part (error term) of doctor i's preference for hospital h is denoted by ε_{iht}, while the unobserved part of slot j's preference for school s is denoted by η_{sjt}.[23] The net joint surplus satisfies the equality: $\Phi_{ijt} = \Phi_{s(i)h(j)t} + \varepsilon_{ih(j)t} + \eta_{s(i)jt}$ for each $i \in I$, $j \in J$, and $t \in [T]$.[24]

Each year, the matching market is subject to a set of regional constraints on the hospital side. We denote the set of regional constraints at a given time period by $\mathcal{O}^t = \{[\underline{o}_z^t, \bar{o}_z^t]\}_{z \in Z}$. As a data-generating process, we assume that the stable outcome with optimal taxation is realized each year under $\mathcal{O}^t$. However, the matching $(d_{ijt})_{i,j,t}$ is not observable. Instead, the available data comprises the aggregate-level matching $(\mu_{sht})_{s,h,t}$, which is the number of matches between medical school s and hospital h at time t.

5.1 Transfer

To define a model object corresponding to the observed salary, we impose an additional structure on the net joint surplus. The *base utility* of i when matched with j at time t is the utility felt by i when matching with j net of transfer, and is denote by U_{ijt}^{base}. Similarly, the base utility of j when matched with i at time t is denoted by V_{ijt}^{base}. We assume that the net joint surplus is the sum of the base utilities of a doctor and a slot forming the match.

Assumption 6. $\Phi_{ijt} = U_{ijt}^{\text{base}} + V_{ijt}^{\text{base}}$.

Furthermore, in accordance with additive separability (Assumption 2), we re-interpret the i.i.d. error terms as the unobserved taste shocks of the agents on both sides of the market. In other words, we assume the following utility structure: define $U_{sht}^{\text{base}} := U_{ijt}^{\text{base}} - \varepsilon_{iht}$ and $V_{sht}^{\text{base}} := V_{ijt}^{\text{base}} - \eta_{sjt}$, so we have

$$U_{ijt}^{\text{base}} = U_{sht}^{\text{base}} + \varepsilon_{iht}, \quad V_{ijt}^{\text{base}} = V_{sht}^{\text{base}} + \eta_{sjt}.$$

We call U_{sht}^{base} and V_{sht}^{base} by *aggregate-level base utility*: they are a part of base utilities

[23]This notation can be confusing as slots themselves do not have preferences. We interpret this as follows: the admission office is composed of members with varying tastes for schools, such as a strong preference for the school from which they graduated. η_{sjt} reflects these differences in the committee members' preferences.

[24]$[T] := \{1, 2, \ldots, T-1, T\}$.

which is determined by the observable characteristics. As a direct implication of Assumption 2, 6 and (5.1), we have $\Phi_{sht} = U^{\text{base}}_{sht} + V^{\text{base}}_{sht}$. Note that the aggregate-level base utility U^{base}_{sht} can be different from the aggregate-level utility U_{sht} introduced in (4).

First, we define an individual level transfer in this marekt. Fix any period t and the taxation policy $w_t = (w_{zt})_z$. Consider a matched pair (i, j) with $h(j) = h$ and $z(h(j)) = z$ for some h and z. We define *individual-level transfer from hospital h to doctor i with a ratio α*, denoted by τ_{iht}, as follows:

$$\tau_{iht} \coloneqq u_{it} - \left(U^{\text{base}}_{sht} + \varepsilon_{iht} - \alpha w_{zt}\right).$$

Tax w_{zt} is levied on the matched pair of doctor i and hospital h. The doctor incurs fraction α of the tax; thus doctor's payoff *without transfer* were to be $U^{\text{base}}_{ijt} - \alpha w_{rt} = U^{\text{base}}_{sht} + \varepsilon_{iht} - \alpha w_{rt}$. In equilibrium, doctor i enjoys equilibrium payoff u_{it}, which could be different from U^{base}_{ijt}. We interpret the difference between equilibrium payoff and payoff without transfer as the individual-level transfer from the hospital side to the doctor side.

Now we define an *aggregate-level transfer* as the average of the individual-level transfer in a hospital h and denote it by ι_{ht}:

$$\iota_{ht} \coloneqq \frac{1}{|D(h)_t|} \sum_{i \in D(h)_t} \tau_{iht},$$

where $D(h)_t$ is the set of doctors matched with any slot of hospital h at time t, We can show the following identities: the aggregate-level transfer from a hospital is equal to the weighted average of the gap between aggregate-level utility and aggregate-level base utility. We use these identities as moment conditions to identify the aggregate-level base utility.

Proposition 2.

$$\iota_{ht} = \sum_s \omega_{sht}\left(U_{sht} - U^{\text{base}}_{sht}\right), \ \ \iota_{ht} = \sum_s \omega_{sht}\left(V^{\text{base}}_{sht} - V_{sht}\right)$$

where $\omega_{sht} = \frac{\mu_{sht}}{\sum_{s'} \mu_{s'ht}}$.

5.2 Estimation

Based on the observable characteristics of s and h, we have a set of variables related to the preferences: we use $X_{sht}^{U,\text{base}}$ as the variables for U_{sht}^{base}, and $X_{sht}^{V,\text{base}}$ as the variables for V_{sht}^{base}. We assume linear structure on both of the preferences:

$$U_{sht}^{\text{base}} = X_{sht}^{U,\text{base}\prime}\beta_U,\ V_{sht}^{\text{base}} = X_{sht}^{V,\text{base}\prime}\beta_V.$$

We define $\delta_{ht}^U \equiv \alpha w_{r(h)t}$ and $\delta_{ht}^V \equiv (1-\alpha)w_{r(h)t}$ as the levied tax on school side and hospital side in period t and treat these two levied taxes as parameters to be estimated. Our parameters of interest are the following: β_U, β_V, δ_{ht}^U for every pair of h and t, and δ_{ht}^V for every pair of z and t. We use θ to indicate the vector of these parameters: $\theta := \left(\beta_U, (\delta_{ht}^U)_{h,t}, \beta_V, (\delta_{ht}^V)_{h,t}\right)$.

Our estimation consists of the following two steps:

1. Estimate the aggregate-level utilities U_{sht} and V_{sht} for every t, and then
2. Estimate θ using the estimated aggregate-level utilities and the observed salaries.

We explain these two steps in order in the following subsections.

5.2.1 First step

Despite the nonparametric identification results obtained in Galichon and Salanié (2021a), we estimate the parametrized version of the aggregate-level utilities. This is because some pairs of school and hospital have zero matches in practice. Hence, in the first step, we use the moment matching estimator proposed in Galichon and Salanié (2021a) to estimate U_{sht} and V_{sht}.[25]

By formulating the aggregate matching outcome μ_{sht} as a realization of a Poisson distribution, it is possible to estimate the aggregate-level utilities by a Poisson regression with fixed effects. For a regressor in the Poisson regressions, we make a set of polynomials for some degree from X_{sht}^U and X_{sht}^V, which is denoted by X_{sht}^{poly}. We model the aggregate-level utilities as follows:

$$U_{sht} = X_{sht}^{\text{poly}\prime}\beta_U^{\text{poly}},\ V_{sht} = X_{sht}^{\text{poly}\prime}\beta_V^{\text{poly}}.$$

[25]Note that our estimation target is U_{sht} and V_{sht}. The difference from the case of Galichon and Salanié (2021a) is that we just need one side fixed effect.

We use $\hat{\beta}_U^{\text{poly}}$ and $\hat{\beta}_V^{\text{poly}}$ as the estimated coefficients attached with the polynomials. And we define the estimated aggregate-level utilities by $\hat{U}_{sht} \equiv X_{sht}^{\text{poly}\prime}\hat{\beta}_U^{\text{poly}}$ and $\hat{V}_{sht} \equiv X_{sht}^{\text{poly}\prime}\hat{\beta}_V^{\text{poly}}$. For the details of this estimator, the readers can refer Galichon and Salanié (2021*b*).

5.2.2 Second step

When we directly observe the values of ι_{ht} for all hospitals and periods, we can use (2) to construct an estimator of θ. By inserting the estimation results in the first step, we construct the following moment conditions for θ:

$$\sum_s \omega_{sht}\left(X_{sht}^{U,\text{base}\prime}\beta_U - \delta_{ht}^U\right) = \sum_s \omega_{sht}\hat{U}_{sht} - \iota_{ht},\ \forall\ h,t$$
$$\sum_s \omega_{sht}\left(X_{sht}^{V,\text{base}\prime}\beta_V - \delta_{ht}^V\right) = \sum_s \omega_{sht}\hat{V}_{sht} + \iota_{ht},\ \forall\ h,t$$

In Section 6, we show a Monte Carlo exercise adopting this approach to show how to recover the structural parameters.

However, we face a measurement problem in practice: we cannot observe the aggregate-level transfer ι_{ht}. Instead, we can only observe the realized salaries paid by hospitals every period. Here, we consider that the transfer in this market includes even non-monetary asset transfers. For example, the hospital accepts the risk of medical incidence by allowing the less-experienced medical interns to get more practice on the job. The workload in a hospital also comprises such unobserved transfers. Furthermore, we expect that the observed salaries correlate with these unobservable terms, which makes the identification more demanding.

For this problem, we introduce a measurement model to connect the observed salaries to ι_{ht}. Denoting the salary paid in a hospital h at time t by S_{ht}, we assume that both schools and hospitals have quasi-linear utilities with respect to monetary transfers:

$$\iota_{ht} = \gamma_{0,U} + \gamma_{1,U}S_{ht} + \psi_{ht}^U$$
$$-\iota_{ht} = \gamma_{0,V} + \gamma_{1,V}S_{ht} + \psi_{ht}^V.$$

Note that $\gamma_{1,V}$ is expected to be negative because the salary is the amount of money paid to the doctor from the hospital. ψ_{ht}^U is the unobserved transfer from the hospital to the matched doctors, and ψ_{ht}^V is the same transfer from the doctor to the hospital.

We need some instrumental variables that have an influence just on the salary in a hospital. As such instrumental variables, we use the characteristics of the surrounding hospitals as in Berry, Levinsohn and Pakes (1995). From our model, such variables clearly impact the aggregate-level transfer from a hospital in an equilibrium. By focusing on the same region, we expect that the unobserved transfers are not so different. Hence, we consider the BLP-type IVs to be valid in our case.

By combining (5.2.2) and (5.2.2), our estimating equation is as follows:

$$\sum_s \omega_{sht} \hat{U}_{sht} = \gamma_{0,U} + \gamma_{1,U} S_{ht} + \sum_s \omega_{sht} \left(X_{sht}^{U\prime} \beta_U - \delta_{ht}^U \right) + \psi_{ht}^U$$
$$\sum_s \omega_{sht} \hat{V}_{sht} = \gamma_{0,V} + \gamma_{1,V} S_{ht} + \sum_s \omega_{sht} \left(X_{sht}^{V\prime} \beta_V - \delta_{ht}^V \right) + \psi_{ht}^V.$$

By considering the weighted average of every variable in X_{sht}^U and X_{sht}^V as independent variables in the right hand side, the above equations are just linear equations in θ. We estimate these linear equations using the instrumental variables discussed above.

6 Monte Carlo Experiment

We start by describing the overall setting of Monte Carlo simulation. All the detail parameter values are left to Appendix B. There are 10 prefectures, numbered from 0 to 9, grouped into three regions: $\{0,1\} \in R_0$, $\{2,3,4,5\} \in R_1$, and $\{6,7,8,9\} \in R_2$. R_0 represents an urban area, and R_1 and R_2 are rural areas. The government worries about the inefficient supply of medical services in R_2 and tryies to satisy a lower bound in terms of number of matches in the region.

We have 20 hospitals in total. Each hospital is placed in one of the prefectures based on a multinomial distribution. Hospital characteristics are modeled dynamically. When we denote each hospital by h, each hospital's capacity, denoted by c_{ht}, starts with a Poisson distribution at time $t = 0$ and evolves over time through a stochastic process involving increments and decrements modeled by independent Poisson distributions. We use j to denote each slot in a hospital. Other hospital-specific characteristics like the number of beds are captured by a variable z_{ht}, which follows a normal distribution.

We have 200 doctors and they are distributed among the prefectures in a similar way of the hospitals. Each doctor belongs to a one of 20 medical schools, and the schools themselves are distributed among prefectures, also based on a multinomial distribution.

Each school has the equal split of the docotors in the same prefecture. The schools have characteristics like an average ability measures following a normal distribution. We use s to denote the school and i to denote a doctor.

We define the net joint surplus generated by a matching between a slot j and a doctor i at time t in the following way:

$$\Phi_{ijt} = \Phi_{sht} + \xi_{ijt},$$

where

$$\Phi_{sht} = U_{sht}^{\text{base}} + V_{sht}^{\text{base}},\ \xi_{ijt} = \varepsilon_{iht} + \eta_{sjt},$$

and

$$\begin{aligned}
U_{sht}^{\text{base}} &= \beta_{1,1}w_{1,ht} + \beta_{1,2}w_{2,ht} + \beta_2 \left|l_s - l_h\right| + \beta_3 1\{h \in R_1 \text{ or } h \in R_2\},\\
V_{sht}^{\text{base}} &= \gamma_1 x_{1,st} + \gamma_2 x_{2,st},\\
\varepsilon_{iht} &\sim Ex1,\ \eta_{sjt} \sim Ex1.
\end{aligned}$$

$|l_s - l_h|$ is a measure of the distance between school s and hospital h: in this simulation, we define this as the absolute value of the gap between the prefecture index. And the last term in U_{sht}^{base} captures the negative impact on the utility from living in rural areas. Note that this rural areas include R_1 which is not the target of the subsidy to assure the lower bound on the matching outcomes. We also use $U_{ijt} = U_{sh(j)t}^{\text{base}} + \varepsilon_{ih(j)t}$ and $V_{ijt} = V_{s(i)ht}^{\text{base}} + \eta_{s(i)jt}$ to denote the individual level preferences.

6.1 Simulation

We compute a stable outcome of an instance of the above market at one time period. The number of matches in each region is 94, 43, and 33. The number of unmatched doctors is 30 and the number of unmatches of slots is 46. Imagine that the government set an upper bound on R_0 to increase the number of matches in rural regions. When we set the upper bound on R_0 to 60, an equilibrium number of mathces are: 60, 45, and 37. The number of unmatched doctors is 58 and the number of unmatches of slots is 74. Under this regional constraint, the tax levied on the matchings in R_0 is 3.149.

Figure 4 depicts the scatter plots of several equilibrium objects when we set $\alpha = 0$, which implies that all the amount of tax is levied on hospital side. The left four panles

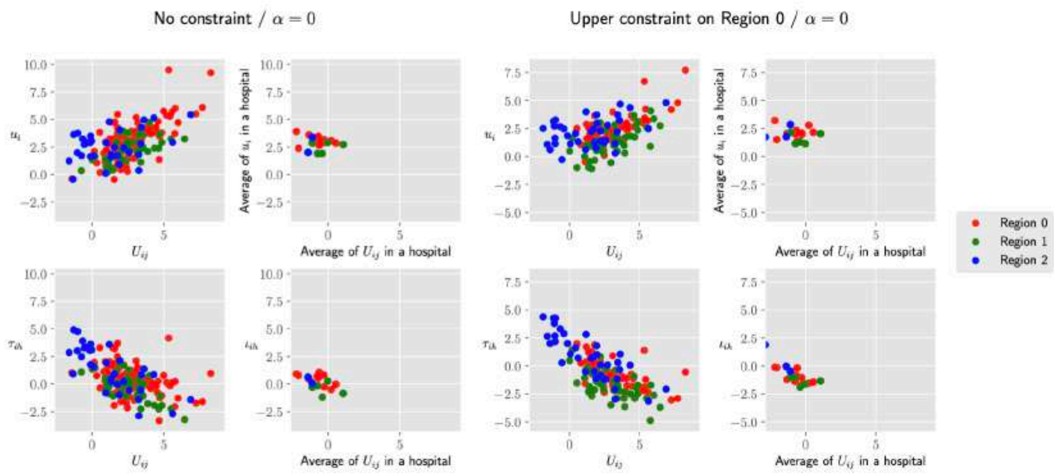


Figure 4. Simulated Stable Outcomes

are obtained when we set no regional constraint, and the right four panels are obtained when we set an upper bound on R_0 to 60. The four panels in each left and right half set of panels depict the same things for the case of no constraint and regional constraint. The first and thrid columns are the results in the stable outcome (with optimal tax) where each dot represents a match. The upper panels are the scatter plots of preference of doctor i has for slot j, U_{ij}, and the utilities attained in a stable outcome, u_i. The lower panels are the scatter plots of U_{ij}^{base} and the transfer in a stable outcome, τ_{ih}. The second and fourth columns represent aggregate level objetcs: the upper panels are the scatter plots of the average of U_{ij} and u_i among a matches in a hospital, and the lower panels are the scatter plots of the average of U_{ij} and aggregate-level transfer, ι_{ih}, from a hospital. In all scatter plots, a red marker represents a match or a hospital in R_0, a green marker for R_1 and a blue market for R_2.

As expected, the utility attained in a stable outcome is higher when a doctor can be matched with a preferred slot whereas the transfer decreases. This decrease is also reflected in a decrease in aggregate-level transfer from a hospital: when the average of U_{ij} in a macthes of a hospital increases, the aggregate-level transfers from the hospital decreases. The impact of a regional constraint on the aggregate-level transfers is clear: in the constrained region, R_0, they decrease under the constraint compared with the case of

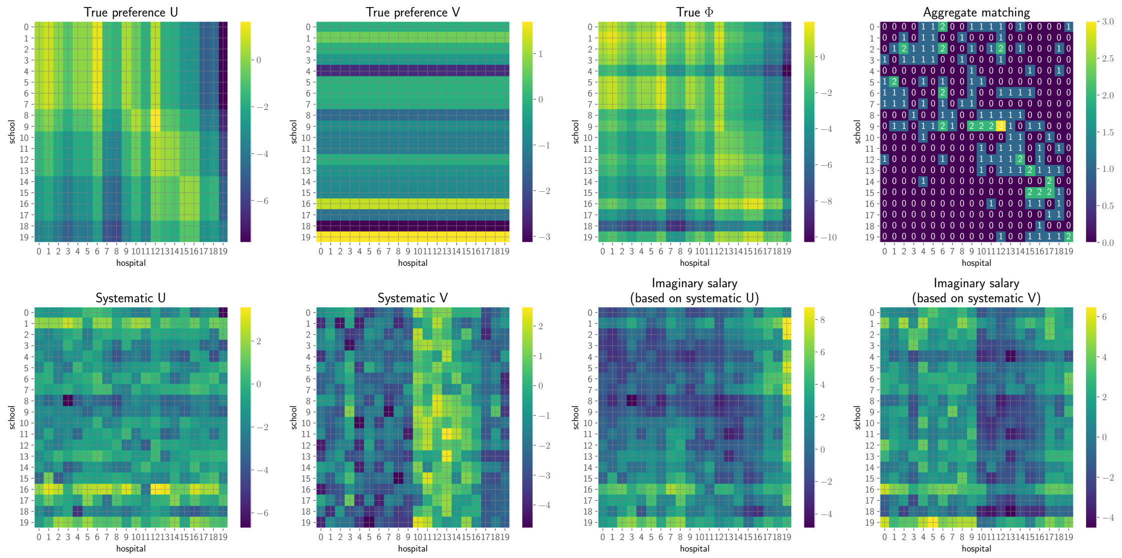


Figure 5. Aggregate Objects

no constraint. This is true in the level sense and the decrease is larger than the changes in other regions. Note that the changes in the aggregate-level transfers and their sizes depend on the value of α. For example, in the extreme case of $\alpha = 1$, the aggregate-level transfers in R_0 increases under the regional constraint. Hence, it is important to estimate the division of tax on hospital side and school side.

Hereafter, we set $\alpha = 0.2$. Figure 5 summarizes the aggregate-level objects computed based on the simulated stable outcome. In all the heatmaps, the horizontal axis represnts hospitals and the vertical axis represents schools. Aggregate matching is depicted in the upper panel in the most right column. The number annotated in each cell represents the number of matches between a hospital and a school. Aggregate-level utilities are computed following the definition stated in (4).

For the ease of argument, we name the gap between the aggregate-level utilities and the aggregate-level base utilities by *imaginary salary*: the imaginary salary from school is defined as $\chi^U_{sh} \equiv U_{sh} - U^{\text{base}}_{sh}$ and the same one from the hospital side is defined as $\chi^V_{sh} \equiv V^{\text{base}}_{sh} - V_{sh}$.[26] Th lower left two panels in Figure 5 show the imaginary salaries between schools and hospitals. The number of doctors in our simulation is 200, which is insufficient for approximating the market with infinite number of doctors. This makes

[26] This name is from the fact that the results of the following two are the same: (1) the agents in one side chooses the agent in the other side by comparing the sum of preference term, imaginary salary, and individual disturbance and (2) Aggregate matching outcome.

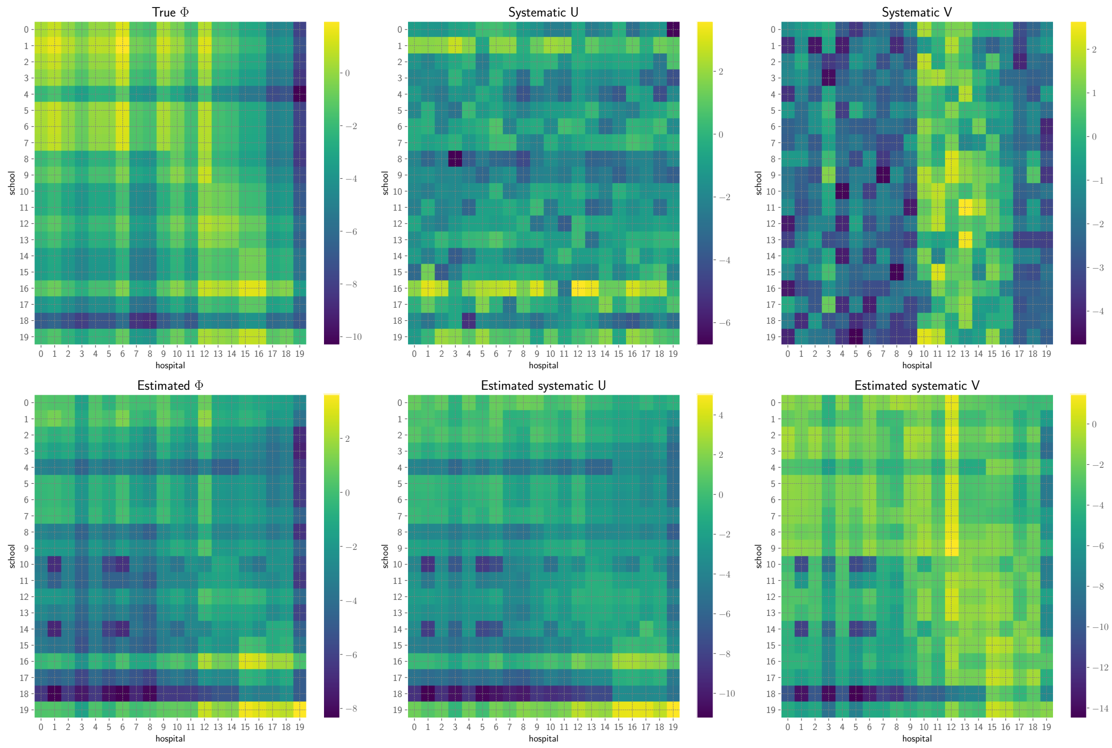


Figure 6. Estimation Results of the First Stage

the gap between the two imaginary wages computed based on U and V.[27]

6.2 Estimation

The estimation results in the first stage are depicted in Figure 6. The upper panels are the heatmap of the true values of Φ_{sh}, U_{sh}, and V_{sh}. They are the estimation targets. The lower panels are the estimation results for the corresponding upper panels. We set the degree of the polynomials to two. The estimated social surplus takes the similar patterns of the true social surplus whereas the estimated aggregate-level utilities show different patterns from the true values of them. These gaps are due to the incompleteness of polynomial approximations in equation (5.2.1). In practice, we handle this problem by inclduing non-linearly transformed base variables when making polynomial series.

For the second stage estimation, we simulate matching outcomes for two periods. In the first period, the government set the upper bound on region 0 to 80 and in the second period, the upper bound is changed to 60. Because, in this exercise, we assume that the true value of aggregate-level transfers are observable, we use the moment conditions

[27] We can show that these two must be equal in the inifite sample case.

Table 5. Estimates

Parameter	γ_1	γ_2	$\beta_{1,1}$	$\beta_{1,2}$	β_3	β_2	$\frac{\alpha}{T}\sum_t w_{0t}$	$\frac{1-\alpha}{T}\sum_t w_{0t}$
Estimate	0.151	-0.0384	0.907	-0.490	-0.527	-0.879	0.265	1.471
Standard Dev.	(0.178)	(0.220)	(0.0670)	(0.0496)	(0.0922)	(0.130)	(0.128)	(0.156)
True Value	1.00	-0.400	1.00	-0.500	-1.00	-1.00	0.378	1.510

(5.2.2) directly to construct a minimum distance estimator. We leave the detail of the construction of this estimator in Appendix B. In this exercise, we use the time average version of the moment conditions and so the tax term is just identified as the time average of the levied tax.

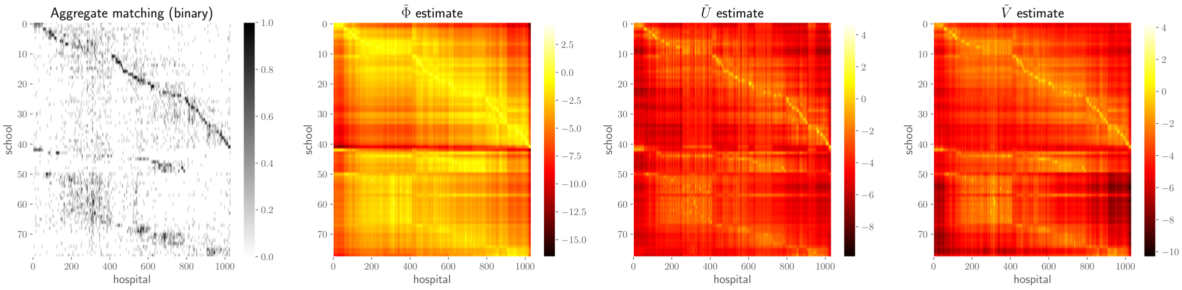
Table 5 summarizes the estimation results of the second stage. The first six columns are the structural parameters in equation (6). The last two columns are the average taxes levied on docotr side and hospital side. From these estimation results, the estimate of α is 0.153 whereas the true value is 0.2. Based on these estimates, we can conduct counterfactual analysis: for example, the taxes in the alternative regional constraints, the matching outcomes, and the salaries are obtained by solving the equilibrium.

7 Empirical Results

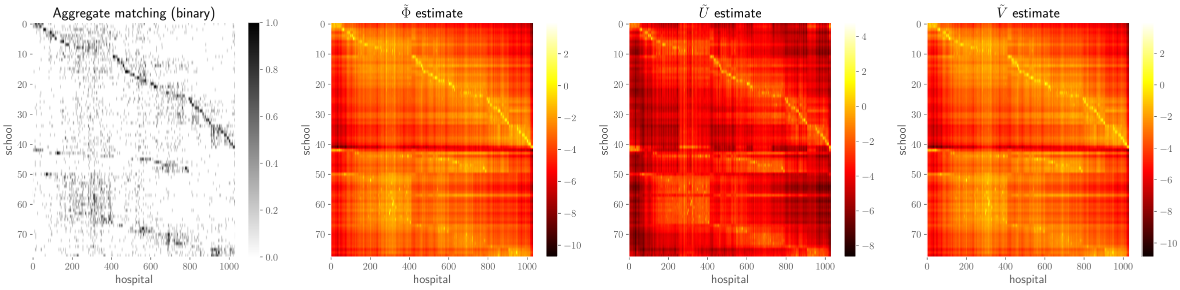
In this section, we show the estimation results based on the data we discussed in Section 2. This section is divided into the two subsections: in Section 7.1, we show the estimation results in our first step, the systematic utilities and the systematic social surplus, and then in Section 7.2, we show the estimation results in our second step, where we recover the preference parameters of both sides of the market. There is no qualitative gap between the two sets of results.

7.1 First Step

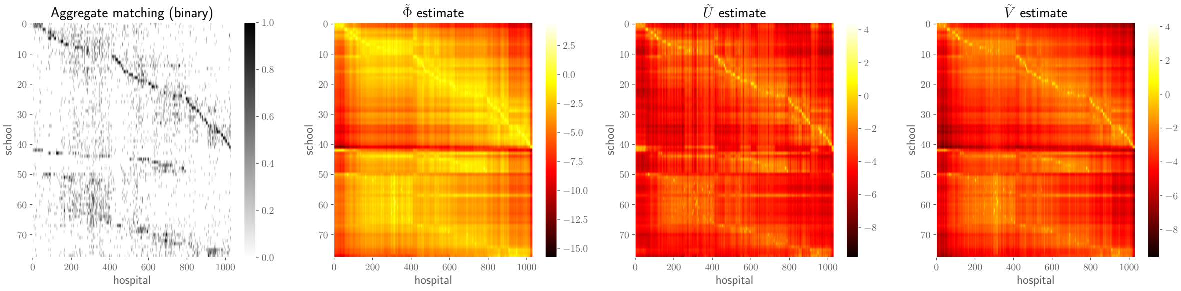
We follow the estimation strategy specified in Section 5.2.1. Remember that our estimation target is the systematic utilities of both sides, $\tilde{U}_{sht}$ and $\tilde{V}_{sht}$, and the systematic social surplus $\tilde{\Phi}_{sht}$. For a given year from 2017 to 2019, we estimate them separately. The degree of polynomial is our tuning parameter and we try two and three. In this section we show the results obtained when we choose three as the degree of polynomials as this choice allows the more flexible functional form. In Appendix C, we will show the results obtained when the degree of polynomials is set to two.



(a) 2017



(b) 2018



(c) 2019

Figure 7. Aggregate Matchings, Estimated Systematic Utilities and Estimated Social Surpluses.

Each panel in Figure 11 shows the aggregate matchings and estimation results in a year where the estimates are obtained when we set the degree of polynomial to three. In each panel, there are four sub figures: the column represents the hospital index and the row represents the university index. The left most figure shows the pairs of university and hospital which have at least one matching. The second left heatmap represents the estimated systematic social surplus between all the pairs of university and hospitals. The right two heatmaps are the same ones for estimated systematic utilities, $\tilde{U}$ and $\tilde{V}$. The brighter color represents the higher values in these three heatmaps.

The visible pattern in the aggregate matching is captured by our estimation. For example, we find many matchings in a downward sloping line in the upper half of the figure of the aggregate matching. This corresponds to the matchings between a public university in a prefecture and hospitals in the same prefecture. Every heatmaps of $\tilde{\Phi}$, $\tilde{U}$, and $\tilde{V}$ capture this pattern. In the lower half of the aggregate matching, we also find a downward sloping line. This represents the matchigns between private universities in a prefecture and hospitals in the same prefecture. From these observations, we expect that the distance between hospital and university matters so much when deciding the matching partner. This preference structure is true not only for university side but also for hospital side as you see in the right most figures[28]. This is because the knowledge about the local medical situation is so essential that the hospital also likes to hire local doctors.

7.2 Second Step

We follow the estimation strategy specified in Section 5.2.2. The target of the estimation is the preference parameters and the tax levied on each side. Based on the estimation, we answer the following two questions: (1) whether is the realized equilibrium AE or EAE? and (2) how much do doctors and hospitals evaluate the other's characteristics in monetary sense?

The first point is essential for the following counterfactual simulations. For all the data periods, the government put the regional constraints on the urban areas as we review in Section 2, whereas there was no official tax levied on such areas and instead the government arbitrarily adjust the capacities of the hospitals. Hence it is an empirical question if the realized matchings are the outcomes in the efficient aggregate equilibrium under the regional constraints or in the aggregate equilibrium with the adjusted capacities.

Remember that our estimating equation is (5.2.2). As the characteristics in the preference of doctor side, i.e. X^U_{sht}, we include the following hospital-specific variables: *logarithm of number of beds of a hospital*, which acts as the measure of the size and the quality of the hospital, *dummy variable of university hospital*, *dummy variable of governmental hospital*, and *dummy variable of urban area.* Furthermore, we include the following pairwise variables: *logarithm of distance*, *logarithm of number of previous matches*, and *dummy*

[28] Agarwal (2015) shows the distance matters for the doctor's decision and do not consider the preference heterogeneity with respect to the distance in the hospital side.

variable of affiliation relationship. As the characteristics in the preference of hospital side, i.e. X^V_{sht}, in addition to the pairwise variables, we include the following university-specific variables: *dummy variable of public university* and *a measure of prestige of a university*.

We consider two cases of already levied taxes: one is the case where the urban areas specified in Section 2 suffer from tax and the other one is the case where just Tokyo suffers from a tax. Because Tokyo is by far the largest city in Japan and the reduction of the realized number of matches in Tokyo is clear compared to the other urban areas, we treat Tokyo as another treatment. In (5.2.2), we include *dummy variable of urban areas* and *dummy variable of Tokyo* as variables in the right hand sides.

The estimation results based on the first step estimations obtained when the degree of polynomials is set to three is summarized in Table 6. Column 1 and 3 correspond to the case of OLS. Column 2 and 4 are the results obtained when we use BLP type IV estimator. First of all, by using the instrumental variables, the direction of the estimated coefficients of salary are alined with the expected signs. Hereafter, we focus on the results obtained using IV estimations. As we mention in the estimation results of the first step, distance between university and hospital negatively influences on both of the preferences of doctors and hospitals. As intuitive and anecdotally validated, the previous number of matches have the strong influence on the preferences. The more previous matches lowers the hurdle to apply for doctor side and the uncertainty about the quality is cleared from the point of view of hospital side. The quality measure for both sides are also impactful. From the doctor side, the number of beds of a hospital increases the utility obtained when matching with a hospital. And the hospital prefers the doctor from a public university, which is more difficult to enter.

Contrary to what is often said, we do not find the negative impact of a dummy variable of university hospital. We also find an unintuitive negative impact of prestige of university on the preference of hospital. Although it is true that the hospitals seek the doctors with higher ability, as evident in the positive impact of a dummy variable of public university, almost all hospitals cannot expect they hire such doctors. The negative coefficient of prestige captures this mind.

We find positive coefficients of Tokyo dummy and urban dummy in the preference of doctor side, and positive coefficient of urban dummy on hospital side. These estimates are considered as complex of two forces. If the preference itself depends on the urban or Tokyo dummy, the coefficients are expected to be positive. In addition to this effect,

Table 6. Estimation Result: Preference Parameters
Degree of Polynomials = 3

	(1) University	(2) University (IV)	(3) Hospital	(4) Hospital (IV)
Constant	-5.494***	-6.724***	1.194	1.954**
	(0.187)	(0.327)	(0.776)	(0.874)
Salary (million Yen)	0.574***	2.527***	0.634***	-1.780**
	(0.128)	(0.479)	(0.146)	(0.792)
Tokyo	0.0371	0.112**	0.0251	-0.0940
	(0.0376)	(0.0461)	(0.0618)	(0.0712)
urban	-0.0307	0.0572*	0.205***	0.125***
	(0.0264)	(0.0333)	(0.0346)	(0.0447)
log(Distance)	-0.438***	-0.436***	-0.409***	-0.373***
	(0.0150)	(0.0121)	(0.0194)	(0.0217)
log(Previous Match)	1.245***	1.229***	1.551***	1.560***
	(0.0305)	(0.0231)	(0.0411)	(0.0463)
Affiliation	0.380**	0.460***	-2.007***	-2.197***
	(0.166)	(0.111)	(0.186)	(0.205)
University hospital	-0.124*	-0.0127		
	(0.0726)	(0.0788)		
Govermental hospital	0.0723**	0.0132		
	(0.0298)	(0.0330)		
log(Beds)	0.744***	0.814***		
	(0.0303)	(0.0353)		
Public university			0.287***	0.289***
			(0.0569)	(0.0610)
Prestige			-1.660**	-2.774***
			(0.734)	(0.810)
N	2847	2627	2847	2627

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Table 7. Estimation Result: Tax Parameters
Degree of polynomials = 3

	(1) University	(2) University	(3) Hospital	(4) Hospital
Constant	-6.658***	-6.704***	1.875**	1.966**
	(0.314)	(0.327)	(0.848)	(0.874)
Salary (million Yen)	2.412***	2.519***	-1.579**	-1.810**
	(0.445)	(0.478)	(0.706)	(0.793)
Urban	0.0338	0.0429	0.114*	0.119*
	(0.0470)	(0.0496)	(0.0671)	(0.0702)
Urban × 2018	0.0639	-0.00972	0.0882	0.0805
	(0.0577)	(0.0644)	(0.0786)	(0.0878)
Urban × 2019	0.0752	0.0509	-0.0886	-0.0634
	(0.0575)	(0.0640)	(0.0821)	(0.0892)
Tokyo		-0.00704		-0.0740
		(0.0751)		(0.116)
Tokyo × 2018		0.265***		0.0282
		(0.102)		(0.144)
Tokyo × 2019		0.0903		-0.0940
		(0.102)		(0.155)
N	2627	2627	2627	2627
Other covariates	√	√	√	√
Tokyo × Year		√		√

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

these dummy variables capture the amount of taxes, i.e. δ_{ht}^{U} and δ_{ht}^{V} in (5.2.2). Hence, even if the estimated coefficients are positive or zeros, we cannot determine that there is no tax on urban areas just from this estimation results. In other words, we cannot determine whether the current situation is AE or EAE.

We take advantage of the fact that the regional constraints on urban areas are getting the more strict as time goes to clarify this point. As we explain in Section 2, the government lowers the upper bounds on the number of matches in the urban areas by 5% every year. Hence, if the urban areas have been alredy levied tax due to this upper bounds, the estimated coefficients on dummy variables of urban or Tokyo will decrease. Table 7 shows the estimation results based on IV estimation: where we include all the covariates in Table 6 and additioanly the interaction terms between dummy variables of urban and Tokyo and the dummy variables of each year. As found in every specifications for both sides, we do not find the decrease in the coefficients of dummy variables of urban areas. Furthermore, we do not find any positive impact of living in urban areas except for living in Tokyo in 2018. Based on these results, we conclude that the current situation is the outcome of the aggregate equilibrium under the reduced capacity.

Next, we evaluate the impact of the characteristics in a monetary unit. We take the fractions of the estimated coefficients of the covariates with respect to the coefficients of salary in the measurement model (5.2.2). The fractions and the standard errors for doctor side are shown in Table 8 and the same ones for hospital side are shown in Table 9. For doctor side, the estimates tell that the match with a hospital which is 10% faraway decreases the utility by from 0.017 to 0.018 million yen: this is about \$108. The number of previous matches, the affiliation relationship and the number of beds of a hospital play the positive influences: 10% increase in the number of previous matches improves the utility by from 0.049 to 0.051 million yes, which is about \$319, the hiring by affiliated hospitals increases the utility by from 0.182 to 0.184 million yen, which is about $\$1,169$, and 10% increase in the number of beds imrpoves the utility by from 0.032 to 0.034 million yen, which is about \$210.

For hospital side, distance and the previous number of matches play the similar roles: the doctors from 10% faraway univeristy decreaes the utility of hospital by from 0.021 to 0.024 million yen, which is about \$146, and 10% increase in the number of previous matches improves the utility of hospital by from 0.086 to 0.099 million yes, which is about from \$549 to \$632. The indicator of public university also has positive impact as

Table 8. University Preference Parameters (Unit: Million Yen)
Degree of Polynomials = 3

	(1)	(2)	(3)
Coefficient of Salary =	2.527	2.412	2.519
log(Distance)	-0.173***	-0.181***	-0.173***
	(0.03)	(0.03)	(0.03)
log(Previous Match)	0.486***	0.508***	0.487***
	(0.09)	(0.10)	(0.09)
Affiliation	0.182***	0.186**	0.182**
	(0.06)	(0.06)	(0.06)
University Hospital	-0.005	-0.001	-0.005
	(0.03)	(0.03)	(0.03)
Governmental Hospital	0.005	0.005	0.005
	(0.01)	(0.01)	(0.01)
log(Beds)	0.322***	0.338***	0.324***
	(0.06)	(0.06)	(0.06)
N	2627	2627	2627
Urban × Year		√	√
Tokyo × Year			√

Table 9. Hospital Preference Parameters (Unit: Million Yen)
Degree of Polynomials = 3

	(1)	(2)	(3)
Coefficient of Salary =	1.780	1.579	1.810
log(Distance)	-0.209*	-0.236*	-0.206*
	(0.10)	(0.11)	(0.10)
log(Previous Match)	0.877*	0.989*	0.862*
	(0.39)	(0.44)	(0.38)
Affiliation	-1.235*	-1.383*	-1.215*
	(0.52)	(0.59)	(0.50)
Public University	0.163*	0.188*	0.160*
	(0.08)	(0.09)	(0.07)
Prestige	-1.558*	-1.764*	-1.534*
	(0.66)	(0.74)	(0.64)
N	2627	2627	2627
Urban × Year		√	√
Tokyo × Year			√

expected: the premium of gradtuating from a public university is from 0.160 to 0.188 million yen, which is about from $\$1,022$ to $\$1,201$.

To grab the sizes of impacts, we compute the ratio of these coefficients to the systematic utility which is considered as the utility level attained in a match. Specifically, we first transform the systematic utilities into monetary unit based on the estimation results in the second stage estimation: for doctor side, we compute $U_{sht}^{money} \equiv \frac{\hat{\hat{U}}_{sht}-\hat{\gamma}_{0,U,t}}{\hat{\gamma}_{1,U}}$ and for hospital side, we compute $V_{sht}^{money} \equiv \frac{\hat{\hat{V}}_{sht}-\hat{\gamma}_{0,V,t}}{\hat{\gamma}_{1,V}}$. Note that the constants depend on period t, i.e. $\hat{\gamma}_{0,U,t}$ and $\hat{\gamma}_{0,V,t}$, because we include the dummy variables of every years. Then, we take the average of these transformed systematic utilities with respect to the periods and the institutions in the other side of the market: $\bar{U}_s^{money} \equiv \frac{1}{HT}\sum_{h,t} U_{sht}^{money}$ and $\bar{V}_h^{money} \equiv \frac{1}{ST}\sum_{s,t} V_{sht}^{money}$. $\bar{U}_s^{money}$ and $\bar{V}_s^{money}$ are measures of the expected utilities in the matching market computed for every universities and hospitals. Finally, we take the ratio of the estimated coefficients to these measures to grab the relative size of the coefficients. For the logarithm covariates, we compute the relative size of 10% changes of the covariates.

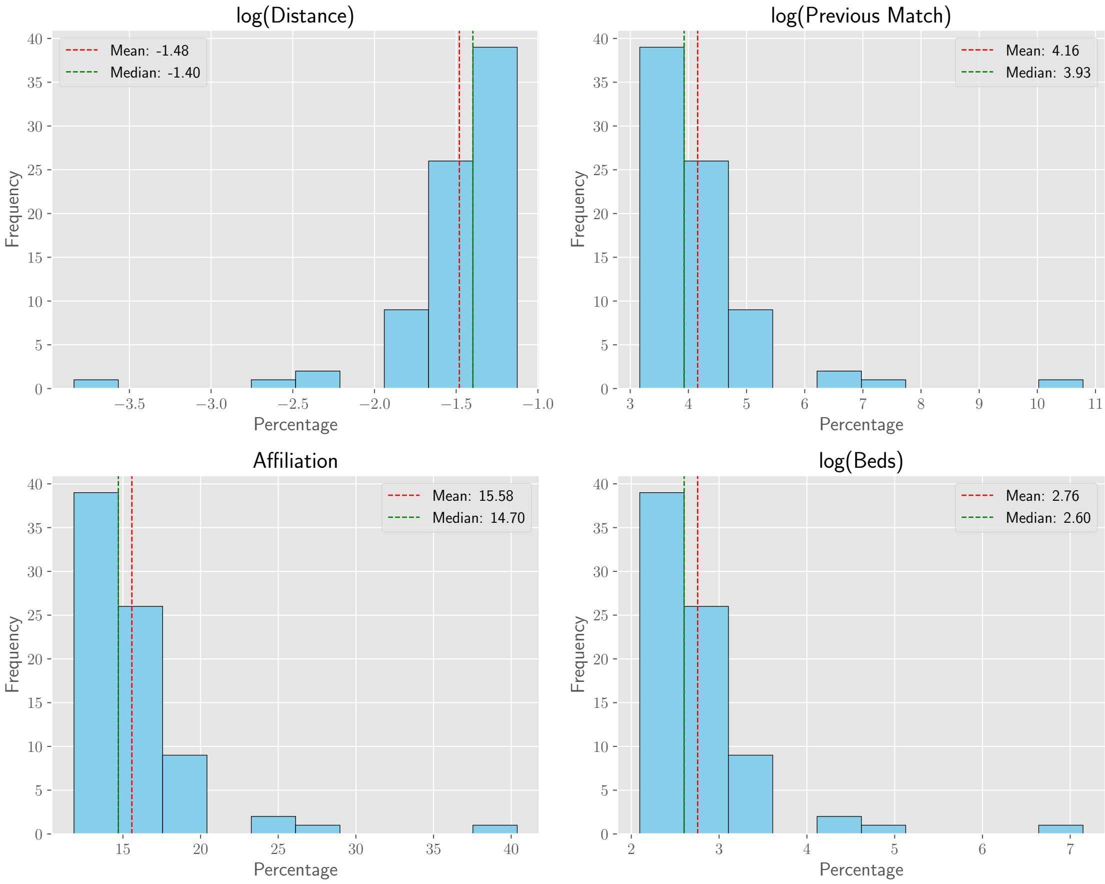


Figure 8. Relative Size of Coefficients in Doctor's Preference.

Figure 8 depicts the histograms of the relative size of the coefficients in doctor's preference for the four covariates which have statistically significant influence in Table 8: the logarithm of distance, the logarithm of the number of previous matches, the dummy variable of affiliation, and the logarithm of the number of beds. In each panel, we show the mean and the median of the relative size of impacts. Although there is variation in the utility level among the universities, the distribution of the relative size of impacts has single peak and their means and the medians are not so different. The average of the relative size of impact of 10% change in distance amounts to 1.4% of the doctor's utility, the same one of the number of previous match amounts to 4.16%, and the same one of the number of beds amounts to 2.76%. On average, affiliation relationship amounts to about 15.58% of the doctors' average utillities.

Figure 9 shows the same histograms for the hospital side preference. We plot the historgrams of the four covariates which shows the statistically significance in Table 9:

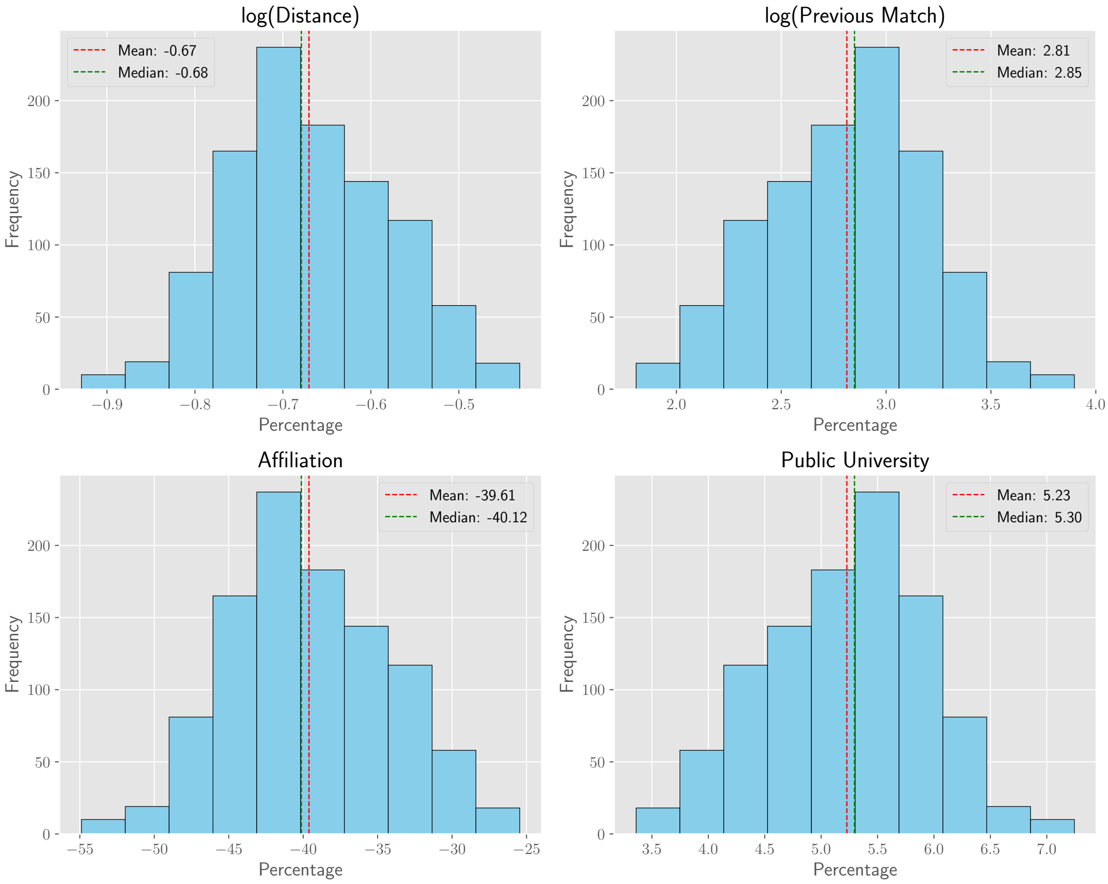


Figure 9. Relative Size of Coefficients in Hospital's Preference.

the logarithm of distance, the logarithm of the number of previous matches, affiliation relationship, and the indicator of the public university. As the average utilities of hospitals are larger than the ones of doctors in the monetary unit sense, the computed relative size of impacts are likely smaller than the values obtained in the case of doctors. The average of the relative size of impact of 10% change in distance amounts to 0.67% of the doctor's utility and the same one of the number of previous match amounts to 2.81%. On average, graduates from public university, which is usually an elite school, gives 5.23% increase in the utility of hospitals. Although the affiliation relationship gives the largest negative impact on the utility of hospitals, this estimate is not stable for the choice of the degree of polynomials as shown in Appendix C.

8 Counterfactual Simulations

While Proposition 1 establishes that the efficient aggregate equilibrium (EAE) is the welfare-maximizing equilibrium under the constraints, it does not quantify the magnitude of efficiency gain. In this section, we demonstrate how our optimal taxation policy can enhance welfare in the JRMP market.

As mentioned in the introduction, the JRMP currently imposes regional caps on the total number of matches within each prefecture to encourage the redistribution of medical residents to rural areas. Every year, the JRMP reduces hospital capacities proportionally to meet these regional upper bounds before the market starts, and then they run the standard deferred acceptance algorithm with those capacities. However, this artificial reduction does not necessarily reflect the actual demand by doctors, resulting in inefficient capacity allocation. Kamada and Kojima (2015) first addressed this issue, proposing a more efficient method for allocating regional caps.

In contrast, we take a different approach. Instead of imposing regional caps, we set lower bounds (floors) on the number of matches in rural areas and meet these thresholds by providing subsidies to both the matched residents and the hospitals in those regions.

To conduct this analysis, it is crucial to know the "actual" capacities of each hospital, which are likely higher than the capacities reported to the JRMP. We estimate each hospital's actual capacity by taking the maximum number of positions reported to the JRMP between 2015 and 2023.

8.1 AE, EAE, and AE without constraints

To assess the welfare improvements brought by the EAE, we first need to establish lower bounds on the number of matches in rural areas. For this purpose, we define "rural areas" as the 15 prefectures with the lowest match rates (calculated as the number of matches divided by the total caps in each prefecture) based on actual JRMP data from 2017 to 2019. We then set these lower bounds according to the total number of simulated matches in the aggregate equilibrium, which is computed using the estimated joint surplus $\hat{\Phi}$ as described in Section 7.

We compare the following three types of aggregate equilibria:

- **AE under the JRMP caps (AE)**: AE under the artificially reduced capacities.

This equilibrium best fits the actual situation of the JRMP. No taxation or subsidy is used.

- **Efficient aggregate equilibrium with lower bounds (EAE)**: EAE under the estimated actual capacities and the lower bounds on rural areas. Only subsidy is used.

- **AE under the estimated actual caps (AE w/o constraints)**: AE under the estimated actual capacities (no caps are set, unlike the current JRMP.) This AE does not necessarily satisfy the lower bounds on rural areas. No taxation or subsidy is used.

The comparison of these equilibria is presented in Table 10. All the values except match rates, #(subsidized regions), and #(constraint violations) are expressed in units of 1 million JPY per month, transformed by the salary coefficients of specification (1) in Table 8 and Table 9. The government's revenue varies depending on the proportion of taxes and subsidies collected from doctors versus hospitals; therefore, we display both the upper and lower bounds in the table. Since the total welfare is the sum of the doctors' welfare, the hospitals' welfare, and the government's revenue, we similarly display both the upper and lower bounds.

We observe that switching from AE to EAE results in a total welfare increase of 4 billion to 12 billion yen, with the match rate rising by approximately 5 percentage points. Meanwhile, the total subsidy required by the government amounts to no more than 18.9 million yen per month, which is significantly lower than the welfare gains achieved. What accounts for this substantial improvement in welfare? Our analysis shows that the welfare in the AE without constraints is very close to that in the EAE, despite a few constraint violations. This indicates that the expansion of capacities increases the available choice alternatives, thereby enhancing overall welfare.

Finally, Figure 10 illustrates the difference in social welfare between EAE and AE across prefectures. Welfare improves across all prefectures under EAE, with particularly notable gains in urban areas—the prefectures currently subject to regional caps under the JRMP. This further suggests that a significant portion of the inefficiency in the current JRMP policy stems from overly strict caps on urban areas.

Table 10. Welfare Comparison of AE, EAE, and AE without constraints

	AE	**EAE**	**AE w/o constraints**
2017			
Match rate	0.868	0.912	0.912
Doctors' welfare	82874.9	84514.3	84507.3
Hospitals' welfare	51788.4	56211.2	56209.0
Government's revenue	0.0	$[-10.5, -7.4]$	0.0
Total welfare	134663.3	$[140715.0, 140718.1]$	140716.3
#(subsidized regions)	0	3	0
Average subsidy	0.000	-0.040	0.000
#(constraint violations)	0	0	3
2018			
Match rate	0.844	0.896	0.895
Doctors' welfare	85736.8	86618.1	86606.8
Hospitals' welfare	53076.2	59992.0	59986.4
Government's revenue	0.0	$[-18.9, -13.3]$	0.0
Total welfare	138812.9	$[146591.3, 146596.8]$	146593.2
#(subsidized regions)	0	5	0
Average subsidy	0.000	-0.038	0.000
#(constraint violations)	0	0	5
2019			
Match rate	0.869	0.912	0.912
Doctors' welfare	84009.3	85300.2	85291.8
Hospitals' welfare	53255.4	58115.4	58114.9
Government's revenue	0.0	$[-9.4, -6.6]$	0.0
Total welfare	137264.7	$[143406.2, 143409.0]$	143406.6
#(subsidized regions)	0	2	0
Average subsidy	0.000	-0.042	0.000
#(constraint violations)	0	0	2

* All values except match rates, #(subsidized regions), and #(constraint violations) are expressed in units of 1 million JPY per month. The government's revenue is positive when taxes are imposed on doctors and hospitals and negative when subsidies are provided to them. The welfare of doctors and hospitals is scaled according to specification (1) in Table 8 and Table 9. We present the bounds of the government's net revenue, scaled by the coefficients on the doctor side and the hospital side, respectively. The total welfare is the sum of doctors' welfare, hospitals' welfare, and the government's revenue. #(constraint violations) counts the number of prefectures violating the lower bounds (among the 15 rural regions).

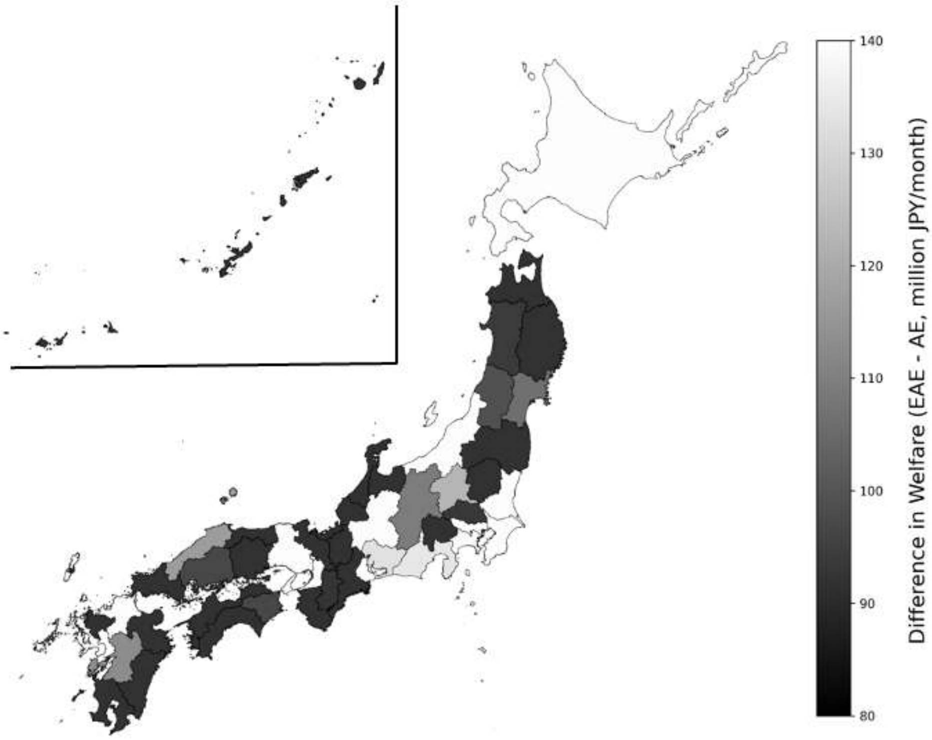


Figure 10. Difference in Welfare between EAE and AE by Prefecture

References

Agarwal, Nikhil. 2015. "An Empirical Model of the Medical Match." *American Economic Review*, 105(7): 1939–1978.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica*, 63(4): 841–890.

Besley, Timothy, Olle Folke, Torsten Persson, and Johanna Rickne. 2017. "Gender quotas and the crisis of the mediocre man: Theory and evidence from Sweden." *Am. Econ. Rev.*, 107(8): 2204–2242.

Chiappori, Pierre-André. 2017. *Matching with Transfers: The Economics of Love and Marriage.* Princeton University Press.

Ellison, Glenn, and Parag A Pathak. 2021. "The efficiency of race-neutral alternatives to race-based affirmative action: Evidence from Chicago's exam schools." *Am. Econ. Rev.*, 111(3): 943–975.

Endo, Hisao. 2019. “Ishi no Sotsugo Kenshu no Genjo to Kadai [The Current Status and Challenges of Postgraduate Medical Training for Doctors].” *Shakai Hosho Kenkyu*, 3(4): 476–491. [In Japanese].

Galichon, Alfred. 2018. *Optimal transport methods in economics.* Princeton University Press.

Galichon, Alfred, and Bernard Salanié. 2021*a*. “Cupid’s Invisible Hand: Social Surplus and Identification in Matching Models.” *Review of Economic Studies, forthcoming.*

Galichon, Alfred, and Bernard Salanié. 2021*b*. “Structural Estimation of Matching Markets with Transferable Utility.” *mimeo.*

Kamada, Yuichiro, and Fuhito Kojima. 2015. “Efficient Matching under Distributional Constraints: Theory and Applications.” *American Economic Review*, 105(1): 67–99.

Kitamura, Kiyoshi, and Yasushi Takagi. 2006. “Sotsugo Rinsho Kenshu Matching System [Postgraduate Clinical Training Matching System].” *Showa Ikaishi*, 66(2): 75–80. [In Japanese].

McFadden, Daniel. 1980. “Econometric Models for Probabilistic Choice Among Products.” *The Journal of Business*, 53(3): S13–S29.

Roth, Alvin E., and Marilda A. Oliveira Sotomayor. 1990. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis (Econometric Society Monographs).* Cambridge University Press.

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

A Omitted Proofs

A.1 Proof of Lemma 1

First, we will show the following lemma: [29]

[29] The following proof of Lemma 2 is almost identical to the proof of Proposition 1 of Galichon and Salanié (2021*a*).

Lemma 2. *Suppose that U and V are aggregate-level utilities given (u, v) (i.e., (4) holds). Suppose that $(d,(u,v))$ is a stable outcome. Then, we have $U_{xy}+V_{xy} \geq \Phi_{xy}+w_{z(y)}$ with equality when $\mu_{xy}>0$ for each $(x,y) \in T$.*

Proof. Fix any $(x, y) \in T$. First, we show $\Phi_{xy} - w_{z(y)} \leq U_{xy} + V_{xy}$. Suppose that $i \in x$ is matched with hospital $j \in y$. We have

$$
\begin{aligned}
u_i &= \max_{j\in J}\{\tilde{\Phi}_{ij} - v_j\} \\
&= \max_{y\in Y} \max_{j\in Y}\{\Phi_{ij} - w_{z(y)} - v_j\} \\
&= \max_{y\in Y} \max_{j\in Y}\{\Phi_{xy} - w_{z(y)} + \varepsilon_{iy} + \eta_{xj} - w_{z(y)} - v_j\} \\
&= \max_{y\in Y}\{\Phi_{xy} - w_{z(y)} + \varepsilon_{iy} + \max_{j\in y}\{\eta_{xj} - v_j\}\} \\
&= \max_{y\in Y}\{\Phi_{xy} - w_{z(y)} + \varepsilon_{iy} - V_{xy}\}
\end{aligned}
$$

Thus, for any $i \in x$, we have

$$
\begin{aligned}
u_i &= \max\left\{\max_{y\in Y}\{\Phi_{xy} - w_{z(y)} + \varepsilon_{iy} - V_{xy}\}, \varepsilon_{i,y_0}\right\} \\
&= \max_{y\in Y_0}\{\Phi_{xy} - w_{z(y)} + \varepsilon_{iy} - V_{xy}\}.
\end{aligned}
$$

Hence,

$$
\Phi_{xy} - w_{z(y)} \leq u_i - \varepsilon_{iy} + V_{xy}.
$$

By taking the infimum over $i \in x$, we have

$$
\Phi_{xy} - w_{z(y)} \leq U_{xy} + V_{xy},
$$

for each $x \in X$ and $y \in Y$.

Next, suppose that $\mu_{xy} > 0$. This implies that there exist $i \in x$ and $j \in y$ such that $d_{ij} = 1$. For this pair, we have $u_i + v_j = \Phi_{ij} - w_{z(y)}$. Suppose toward contradiction that $U_{xy} + V_{xy} > \Phi_{xy}$. By (4), we have $\Phi_{xy} - w_{z(y)} < u_i - \varepsilon_{iy} + v_j - \eta_{xj}$, and thus $\Phi_{ij} - w_{z(y)} < u_i + v_j$. A contradiction.

□

Proof of Lemma 1. Fix any type $x \in X$ and doctor $i \in x$. By definition of U_{xy}, we have

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

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

We want to claim that $u_i \le \max_{y \in Y_0} \{U_{xy} + \varepsilon_{iy}\}$. Suppose toward contradiction that there exists type $x \in X$ and doctor $i \in x$ such that

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

First, consider the case where i is matched with some hospital $j \in y$. Then

$$\begin{aligned} \Phi_{ij} - w_{z(y)} &= u_i + v_j \\ &> \left(\max_{y' \in Y_0} U_{xy'} + \varepsilon_{iy'} \right) + \left(\max_{x' \in X_0} V_{x'y} + \eta_{x'j} \right) \\ &\ge U_{xy(j)} + \varepsilon_{iy(j)} + V_{xy(j)} + \eta_{xj} \\ &\ge \Phi_{xy} - w_{z(y)} + \varepsilon_{iy(j)} + \eta_{xj} \quad (\because \text{ Lemma 2}) \\ &= \Phi_{ij} - w_{z(y)}. \end{aligned}$$

A contradiction. Next, consider the case where i is unmatched. Then

$$u_i = \Phi_{i,y_0} = \varepsilon_{i,y_0} > \max_{y \in Y_0} \{U_{xy} + \varepsilon_{iy}\} \ge \varepsilon_{i,y_0}.$$

A contradiction. Therefore, we have $u_i \le \max_{y \in Y_0} \{U_{xy} + \varepsilon_{iy}\}$ and hence $u_i = \max_{y \in Y_0} \{U_{xy} + \varepsilon_{iy}\}$. We can show $v_j = \max_{x \in X_0} \{V_{xy} + \eta_{xj}\}$ in a similar manner. □

A.2 Proof of the Uniqueness of the Solutions of (P) and (D)

We will show the following:

Lemma 3. *If G and H are strictly convex and differentiable,*[30] *(P) and (D) have unique solutions for any Φ.*

[30] This holds under Assumptions 1-5.

For (P), since G and H are differentiable (WDZ theorem), G^* and H^* are strictly convex (Proposition D.14 of Galichon (2018)). Therefore, the objective function of (P) is strictly convex in μ, which implies the uniqueness of the solution.

For (D), suppose toward a contradiction that there are two different optimal solutions $\alpha := (U, V, \bar{w}, \underline{w})$ and $\beta := (U', V', \bar{w}', \underline{w}')$. Note that $\gamma := \frac{1}{2}(\alpha + \beta)$ is also feasible. If either $U \neq U'$ or $V \neq V'$, then γ gives a strictly lower value due to the strict convexity of G and H, which contradicts the optimality of α and β.

Suppose that $U = U$ and $V = V'$. We must have $(\bar{w}, \underline{w}) \neq (\bar{w}', \underline{w}')$. Since G is strictly convex, we have $\frac{\partial G}{\partial U_{xy}} > 0$ for each (x, y). By the complementary slackness condition with respect to U_{xy}, we have $U_{xy} + V_{xy} = \Phi_{xy} + \bar{w}_{z(y)} - \underline{w}_{z(y)}$ and $U'_{xy} + V'_{xy} = \Phi_{xy} + \bar{w}'_{z(y)} - \underline{w}'_{z(y)}$ for each (x, y). Since $U = U'$ and $V = V'$, this implies that

$$\bar{w}_z - \underline{w}_z = \bar{w}'_z - \underline{w}'_z.$$

for each z. Since $\bar{o}_z > \underline{o}_z$, we must have $\bar{w}_z \underline{w}_z = 0$; otherwise, there exists $\varepsilon > 0$ such that $(U, V, \tilde{\bar{w}}, \underline{\tilde{w}})$, where $\tilde{\bar{w}}_z := \bar{w}_z - \varepsilon$ and $\underline{\tilde{w}}_z := \underline{w}_z - \varepsilon$ attains a strictly lower value. Similarly, we have $\bar{w}'_z \underline{w}'_z = 0$. However, these combined with (A.2) imply that $\bar{w} = \bar{w}'$ and $\underline{w} = \underline{w}'$: if $\bar{w}_z > 0$, then $\underline{w}_z = 0$, $\underline{w}'_z = 0$, and $\bar{w}_z = \bar{w}'_z$. If $\underline{w}_z > 0$, then $\bar{w}_z = 0$, $\bar{w}'_z = 0$, and $\underline{w}_z = \underline{w}'_z$. If $\bar{w}_z = \underline{w}_z = 0$, then $\bar{w}'_z = \underline{w}'_z = 0$. A contradiction.

A.3 Proof of Theorem 1

First, we show two lemmas used in the main proof.

Lemma 4. *Under Assumptions 1-5, G and H are strictly increasing and strictly convex.*

Proof. <u>G is strictly increasing.</u> 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

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

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

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

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

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

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

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 of full support,

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

holds. This implies that

$$\Big(\max_y \ s(U^1_{xy} + \varepsilon_{iy})\Big) + \Big(\max_y \ (1-s)(U^2_{xy} + \varepsilon_{iy})\Big) > \max_y \ sU^1_{xy} + (1-s)U^2_{xy} + \varepsilon_{iy}$$

occurs with strictly positive probability, and thus (A.3) $>$ (A.3) 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. $\square$

Proof of Theorem 1

Part 1: Suppose that w is an optimal taxation policy. For any taxation policy w, given $(\Phi_{ij})_{ij}$ and w, a stable outcome $(d,(u,v))$ is realized, and the corresponding μ, U, and V are defined. Then, μ is a solution to (P). Since u and v are a part of the stable outcome, U and V must satisfy $U_{xy} + V_{xy} \geq \Phi_{xy} - w_{z(y)}$ for each x and y, and the market clearing condition. Since G and H are strictly increasing and convex, the market clearing condition must hold at the optimum of (D). Thus, by the uniqueness of the solution (Lemma 3), the aggregate-level utilities U and V coincide with the ones in the optimal solution to (D). Therefore, for the optimal solutions to $(U, V, \bar{w}, \underline{w})$, we have $w_z = \mathbb{1}\{\bar{w}_z > 0\}\bar{w}_z - \mathbb{1}\{\underline{w}_z > 0\}\underline{w}_z$; otherwise it violates the uniqueness of $(U, V, \bar{w}, \underline{w})$. [31]

Part 2: Suppose that μ and $(U, V, \bar{w}, \underline{w})$ are the unique optimal solutions to (P) and (D), respectively. We want to claim that w, defined as $w_z = \mathbb{1}\{\bar{w}_z > 0\}\bar{w}_z - \mathbb{1}\{\underline{w}_z > 0\}\underline{w}_z$, is an optimal taxation policy. Suppose toward contradiction that there is another optimal taxation policy w'. Let μ' and (U', V') be the aggregate-level matching and utilities under w'. Then, μ' is the solution to (P); the market clearing condition and the uniqueness of the solution imply that U' and V' is a part of the optimal solutions of (D). Then, by the same argument as Part 1, w' must satisfy $w'_z = \mathbb{1}\{\bar{w}_z > 0\}\bar{w}_z - \mathbb{1}\{\underline{w}_z > 0\}\underline{w}_z$ for each z, and the uniqueness of the solution imply $w = w'$. A contradiction. $\square$

A.4 The objective function of (P) corresponds to social welfare

Recall the definition of the Legendre-Fenchel transforms:

$$G^*(\mu) := \begin{cases} \sup\limits_{U} \left\{ \sum\limits_{x \in X} \sum\limits_{y \in Y_0} \mu_{xy} U_{xy} - G(U) \right\} & \left(\forall x \in X, \sum_{y \in Y_0} \mu_{xy} \leq n_x \right) \\ \infty & \text{otherwise} \end{cases},$$

$$H^*(\mu) := \begin{cases} \sup\limits_{V} \left\{ \sum\limits_{y \in Y} \sum\limits_{x \in X_0} \mu_{xy} V_{xy} - H(V) \right\} & \left(\forall y \in Y,\ \sum_{x \in X_0} \mu_{xy} \leq m_y \right) \\ \infty & \text{otherwise} \end{cases}.$$

Observe that G^* and H^* are both continuous in μ on their effective domains. Suppose that (μ, U, V) forms an aggregate equilibrium. The market clearing condition states $\mu \in \partial G(U)$, where ∂G denotes the subgradient of G. By Proposition D.13 of Galichon (2018), we have $G(U) + G^*(\mu) = \sum_{x,y} \mu_{xy} U_{xy}$. Similarly, we have $H(V) + H^*(\mu) = \sum_{x,y} \mu_{xy} V_{xy}$. Thus, we have

$$\begin{aligned} \sum_{x,y} \mu_{xy} \Phi_{xy} + \mathcal{E}(\mu) &= \sum_{x,y} \mu_{xy} \Phi_{xy} + \left(G(U) + H(V) - \sum_{x,y} \mu_{xy} (U_{xy} + V_{xy}) \right) \\ &= G(U) + H(V), \end{aligned}$$

[31] Let $\bar{w}'_z := \mathbb{1}\{w_z > 0\} w_z$ and $\underline{w}'_z := -\mathbb{1}\{w_z < 0\} w_z$. Tuple $(U, W, \bar{w}', \underline{w}')$ is feasible in (D).

where the last equality holds since $U_{xy} + V_{xy} = \Phi_{xy}$ when $\mu_{xy} > 0$.

A.5 Another scenario for stable outcomes: deferred acceptance with endogenous wages

Consider the following dynamic game:

1. Hospitals set the amount of monetary transfer (henceforce *wage*) $t = (t_{ij})_{ij}$;

2. Doctors and hospitals submit their preference lists after observing wages t;

3. The matching is finalized by the standard deferred acceptance (DA) algorithm.

Let $\mu \colon I \cup J \to I_0 \cup J_0$ be a function such that $\mu(k)$ denotes the partner of agent k under μ. Denote the base utilities by $\tilde{u} = (u_{ij})_{ij}$ and $\tilde{v} = (v_{ij})_{ij}$ (see Section 5 for definition. Note that $\Phi_{ij} = u_{ij} + v_{ij}$.) Given t, doctor i's payoff under matching μ is $u_{i,\mu(i)} + t_{i,\mu(i)}$; job slot j's payoff is $v_{\mu(j),j} - t_{\mu(j),j}$.

We assume that, in the second stage, all agents truthfully report their preference ranking to the mechanism. Since the DA algorithm is strategy-proof for one side, this is equivalent to assume that one of the two sides, say hospital-side, report their preferences always truth-telling.

Fix the job-slot-optimal stable outcome $(d, (u, v))$ given Φ.[32] For (i, j) matched under d (i.e., $d_{ij} = 1$), there exists t^*_{ij} such that $u_i = u_{ij} + t^*_{ij}$ and $v_j = v_{ij} - t^*_{ij}$. For (i, j) such that $d_{ij} = 0$, choose any t^*_{ij} that satisfy $u_{ij} + t^*_{ij} \leq u_i$ and $v_{ij} - t^*_{ij} \leq v_j$. Note that such t^*_{ij} must exists since, by stability, we have $u_{ij} + v_{ij} \leq u_i + v_j$.[33] Given t^*, in the second stage, each doctor i submits a preference ranking according to $u_{ij} + t^*_{ij}$, and each job slot j submits a preference ranking according to $v_{ij} - t^*_{ij}$. By construction, each agent prefers the partner matched under d most. Hence, the following lemma holds.

Lemma 5. *For any $i \in I$ and $j \in J$,*

$$v_j = \max_{i \in I_0} \left\{ v_{ij} - t^*_{ij} \right\}, \quad u_i = \max_{j \in J_0} \left\{ u_{ij} + t^*_{ij} \right\}.$$

We also assume that, if there is a tie, each job slot always place the partner under d on the top of the preference list in the second stage.

[32] The set of stable outcomes form a lattice. See Chapter 8 of Roth and Sotomayor (1990) for a textbook reference.

[33] $t^*_{ij} := \frac{1}{2}\left((v_{ij} - v_j) + (u_i - u_{ij})\right)$ works, for example.

Given this behavior in the second stage, we can show that setting t^* forms a NE in the first stage since $(d, (u, v))$ is the job-slot-optimal stable outcome.

Lemma 6. *Suppose that, in the last stage, all the agents truthfully report their preferences to the DA algorithm and breaks the tie in favor of the partner under d. Then, it forms a Nash equilibrium in the first stage that all the hospitals set t^* as their wages.*

Proof. Fix any job slot j. We will check if j can be strictly better off by choosing $(t_{ij})_i \neq (t^*_{ij})_i$. Suppose toward contradiction that it is possible under $t_j \coloneqq (t_{ij})_i$. Let $i \in I_0$ denote the partner of j under d.

First, we show that j should be matched with $i' \neq i$ with t_j: this is clear if $i = i_0$. If $i \in I$ and $i' = i$, it violates the job-slot-optimality of $(d, (u, v))$.

Suppose that j is matched with $i' \neq i$ with t_{ij}. Then, we have $v_{ij} - t_{ij} > v_j$ and $u_{ij} + t_{ij} \geq u_i$.[34] However, this implies that $\Phi_{ij} = u_{ij} + v_{ij} > u_i + v_j$, which violates the stability. A contradiction. □

The outcome of the gameplay described here coincides with the stable outcome $(d, (u, v))$. Moreover, given t^* set in the first stage, no agent has incentive to deviate from the truthful report.

B Detail Settings of Monte Carlo Simulation

B.1 Minimum distance estimator

We use the time average of both of the sides to construct the moment conditions (5.2.2). Note that, in this case, it is impossible to idenify δ^U_{ht} and δ^V_{ht} for every t because the summation of them with respect to the time determines the moment values. Hence, all we can identify is the average tax levied through the time periods. We define a function

[34] This inequality is actually strict due to the tie-breaking assumption.

g to represent the moment conditions:

$$g(\theta) \equiv \begin{pmatrix} \frac{1}{T}\sum_t \left(\sum_s \omega_{s1t}\left(X_{s1t}^{U\prime}\beta_U - \delta_{Ht}^U\right) - \sum_s \omega_{s1t}\hat{\hat{U}}_{s1t} - \iota_{1t}\right) \\ \vdots \\ \frac{1}{T}\sum_t \left(\sum_s \omega_{sHt}\left(X_{sHt}^{U\prime}\beta_U - \delta_{Ht}^U\right) - \sum_s \omega_{sHt}\hat{\hat{U}}_{sHt} - \iota_{Ht}\right) \\ \frac{1}{T}\sum_t \left(\sum_s \omega_{s1t}\left(X_{s1t}^{V\prime}\beta_V - \delta_{1t}^V\right) - \sum_s \omega_{s1t}\hat{\hat{V}}_{s1t} - \iota_{1t}\right) \\ \vdots \\ \frac{1}{T}\sum_t \left(\sum_s \omega_{sHt}\left(X_{sHt}^{V\prime}\beta_V - \delta_{Ht}^V\right) - \sum_s \omega_{sHt}\hat{\hat{V}}_{sHt} - \iota_{Ht}\right) \end{pmatrix}.$$

Our estimator is the minimum distance estimaor where the moment condition is specifed in (5.2.2). When we write the asymptotic variance of $\hat{\hat{U}}_{sht}$ and $\hat{\hat{V}}_{sht}$ by S_t^U and S_t^V, the optimally weighted minimum distance estimator is defined as follows:

$$\hat{\theta} \equiv \arg\min_{\theta}\ g'(\theta)S^{-1}g(\theta),$$

where

$$S = \begin{pmatrix} \frac{1}{T^2}\sum_t S_t^U & 0 \\ 0 & \frac{1}{T^2}\sum_t S_t^V \end{pmatrix}.$$

We can compute the asymptotic distribtuion of the estimator as follows and the standard error can be obtained directly[35]. As the Poisson regression in the first step has the explicit form of S_t^U and S_t^V, by inserting the estimated results, we directly compute get the estimates of the standard errors for every parameters.

Theorem 2. *Under the regularity conditions, the asymptotic distribution of $\hat{\theta}$ is as follows:*

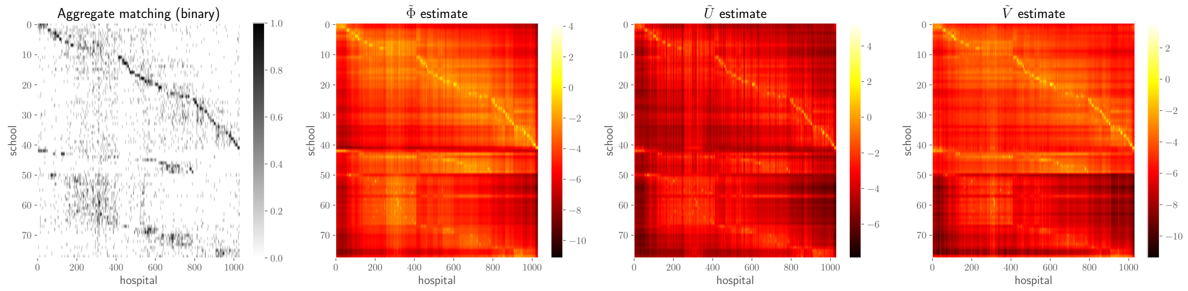
$$\sqrt{ST}\left(\hat{\theta} - \theta\right) \xrightarrow{d} N\left(0, \left(\Gamma' S^{-1}\Gamma\right)^{-1}\right),$$

where $\Gamma = \frac{\partial}{\partial\theta}g(\theta)$.

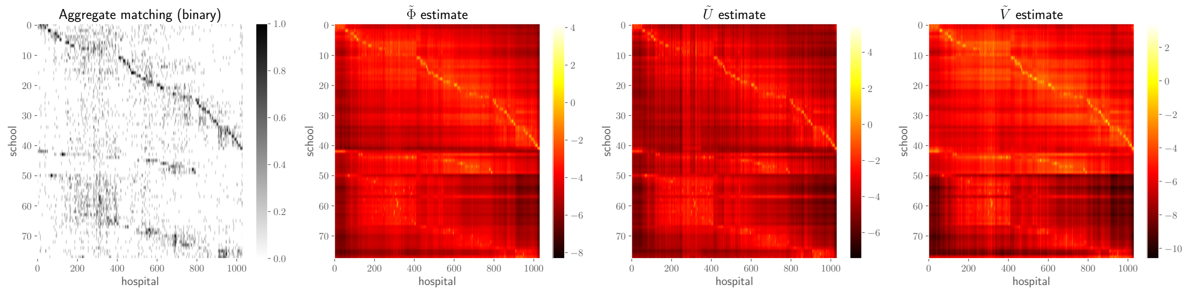
[35] We assume that the polynomial approximation regarding the systematic utility is correct. When there is a misspecification, we must treat the bias due to the misspecified model, whichi is beyond the scope of this study.

C Additional Empirical Results

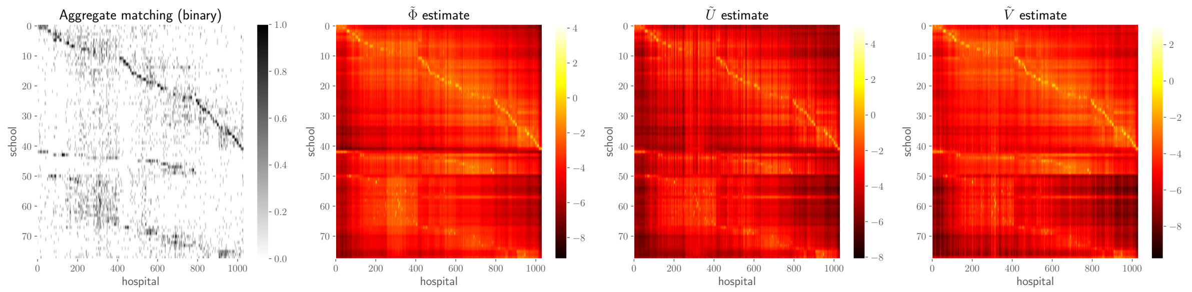
Here we show the empirical results obtained when we set the degree of polynomials in the first step to two. All the tables and figures listed here corresponds to the tables and figures shown in Section 7.



(a) 2017



(b) 2018



(c) 2019

Figure 11. Aggregate matchings, estimated systematic utilities and estimated social surpluses.

Note:

Table 11. Estimation Result: Preference Parameters
Degree of polynomials = 2

	(1) University	(2) University (IV)	(3) Hospital	(4) Hospital (IV)
Constant	-5.917***	-7.988***	1.485**	1.703**
	(0.225)	(0.429)	(0.705)	(0.787)
Salary (million Yen)	0.564***	4.129***	0.704***	-1.231*
	(0.143)	(0.627)	(0.142)	(0.718)
Tokyo	-0.129***	0.0102	0.100*	-0.00294
	(0.0492)	(0.0604)	(0.0524)	(0.0621)
urban	-0.102***	0.0524	0.252***	0.191***
	(0.0339)	(0.0436)	(0.0317)	(0.0418)
log(Distance)	-0.380***	-0.400***	-0.331***	-0.304***
	(0.0157)	(0.0158)	(0.0150)	(0.0172)
log(Previous Match)	1.583***	1.563***	1.663***	1.667***
	(0.0398)	(0.0302)	(0.0398)	(0.0449)
Affiliation	-0.488**	-0.431***	-2.676***	-2.827***
	(0.199)	(0.146)	(0.173)	(0.194)
University hospital	-0.199**	0.0127		
	(0.0800)	(0.103)		
Govermental hospital	0.0319	-0.0589		
	(0.0341)	(0.0433)		
log(Beds)	0.511***	0.628***		
	(0.0359)	(0.0462)		
Public university			0.182***	0.176***
			(0.0531)	(0.0573)
Prestige			-1.905***	-3.125***
			(0.668)	(0.727)
N	2847	2627	2847	2627

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Note:

Table 12. Estimation Result: Tax Parameters
Degree of polynomials = 2

	(1) University	(2) University	(3) Hospital	(4) Hospital
Constant	-8.028***	-7.928***	1.755**	1.725**
	(0.415)	(0.427)	(0.763)	(0.784)
Salary (million Yen)	4.298***	4.121***	-1.276**	-1.234*
	(0.589)	(0.625)	(0.643)	(0.719)
Urban	-0.100	-0.116*	0.106*	0.0755
	(0.0621)	(0.0649)	(0.0623)	(0.0646)
Urban × 2018	0.275***	0.247***	0.179**	0.239***
	(0.0763)	(0.0842)	(0.0720)	(0.0811)
Urban × 2019	0.217***	0.256***	0.0655	0.107
	(0.0761)	(0.0838)	(0.0749)	(0.0805)
Tokyo		0.0257		0.119
		(0.0982)		(0.101)
Tokyo × 2018		0.0976		-0.216*
		(0.133)		(0.123)
Tokyo × 2019		-0.145		-0.151
		(0.133)		(0.132)
N	2627	2627	2627	2627
Other covariates	√	√	√	√
Tokyo × Year		√		√

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Note:

Table 13. University preference parameters (unit: million Yen)
Degree of polynomials = 2

	(1)	(2)	(3)
Coefficient of Salary =	4.129	4.298	4.121
log(Distance)	-0.097***	-0.094***	-0.097***
	(0.01)	(0.01)	(0.01)
log(Previous Match)	0.378***	0.363***	0.379***
	(0.06)	(0.05)	(0.06)
Affiliation	-0.104**	-0.102**	-0.106**
	(0.04)	(0.04)	(0.04)
University Hospital	0.003	0.006	0.003
	(0.02)	(0.02)	(0.02)
Governmental Hospital	-0.014	-0.015	-0.014
	(0.01)	(0.01)	(0.01)
log(Beds)	0.152***	0.148***	0.153***
	(0.02)	(0.02)	(0.02)
N	2627	2627	2627
Urban × Year		$\surd$	$\surd$
Tokyo × Year			$\surd$

Note:

Table 14. Hospital preference parameters (unit: million Yen)
Degree of polynomials = 2

	(1)	(2)	(3)
Coefficient of Salary =	1.231	1.276	1.234
log(Distance)	-0.247	-0.238	-0.247
	(0.15)	(0.13)	(0.15)
log(Previous Match)	1.354	1.307*	1.352
	(0.79)	(0.66)	(0.79)
Affiliation	-2.296	-2.222*	-2.294
	(1.29)	(1.07)	(1.29)
Public University	0.143	0.136	0.141
	(0.09)	(0.07)	(0.09)
Prestige	-2.537	-2.451*	-2.522
	(1.37)	(1.15)	(1.36)
N	2627	2627	2627
Urban × Year		√	√
Tokyo × Year			√

Note: