Evaluating the Efficiency of Regulation in Matching Markets with Distributional Disparities*

Kei Ikegami† Atsushi Iwasaki‡ Akira Matsushita§
Kyohei Okumura¶

November 6, 2025

Abstract

Cap-based regulations are widely used to address distributional disparities in matching markets, but their efficiency relative to alternative instruments such as subsidies remains poorly understood. This paper develops a framework for evaluating policy interventions by incorporating regional constraints into a transferable utility matching model. We show that a policymaker with aggregate-level match data can implement a taxation policy that maximizes social welfare and outperforms any cap-based policy. Using newly collected data from the Japan Residency Matching Program, we estimate participant preferences and simulate counterfactual match outcomes under both cap-based and subsidy-based policies. The results reveal that the status quo cap-based regulation generates substantial efficiency losses, whereas small, targeted subsidies can achieve similar distributional goals with significantly higher social welfare.

JEL Classification Codes: C25; C78; H20; I14.
Keywords: Matching with Distributional Constraints; Discrete Choice; Assignment Problems; Two-sided Matching; Subsidies; Tax; Redistribution.

*We thank Alfred Galichon, Fuhito Kojima, Yeon-Koo Che, Paulo Somaini, and Adam Kapor for helpful comments and thoughtful discussions. We also acknowledge the participants of NYU econometrics seminar, JEA annual meeting, APIOC, IIOC, EC 2025, and empirical market design session at SITE conference.

†University of Tokyo, ikegamikei@e.u-tokyo.ac.jp.
‡University of Electro-Communications, atsushi.iwasaki@uec.ac.jp
§Kyoto University, amatsushita@i.kyoto-u.ac.jp
¶University of Wisconsin–Madison, kyohei.okumura@wisc.edu

1 Introduction

Outcomes in real-world matching markets often diverge from socially desirable allocations, particularly in terms of distributional equity. To address such imbalances, policymakers often restrict the number of matches for certain categories. Examples of such regulations include race-based affirmative action in U.S. college admissions aimed at promoting diversity (Ellison and Pathak, 2021), gender quotas in electoral systems designed to enhance female representation (Besley et al., 2017), and the residency market in Japan analyzed in Kamada and Kojima (2015), in which the government currently caps the number of matches in high-demand urban areas to redirect applicants to underserved rural areas.

While such cap-based regulations are widely used, monetary interventions such as taxes and subsidies offer a natural, yet comparatively underexplored, alternative. The potential advantage of monetary interventions lies in their ability to account for the intensity of preferences. A cap is a blunt instrument that may prevent high-surplus matches, whereas a targeted subsidy could, in principle, influence marginal participants without generating large welfare losses. It remains a central and unresolved question for both theory and policy whether this theoretical potential translates into significant efficiency gains, and under what conditions.

Answering these questions requires moving beyond the non-transferable utility (NTU) framework that dominates the literature on matching with distributional constraints. The NTU framework is ill-suited for policy analysis for two primary reasons. First, its reliance on ordinal preferences cannot readily accommodate interventions involving taxes or subsidies, or quantify their welfare consequences. Second, it typically treats transfers, such as wages, as exogenous, failing to capture how they endogenously adjust in response to policy interventions. These limitations underscore the need for a modeling approach toward more efficient and credible policy intervention.

In this paper, we develop a transferable utility matching model that incorporates *regional constraints*, defined as upper and lower bounds on the number of matches in each region. Within this framework, a system of taxes and subsidies can implement the matching that maximizes participants' welfare among all those satisfying the constraints. This welfare-maximizing outcome establishes a clear welfare benchmark against which other policies can be evaluated. We then provide a path to empirical implementation by embedding our model in the aggregate matching framework of Galichon and Salanié

(2021*a*). This enables us to identify model primitives from past match data, compute optimal subsidy policies, and conduct counterfactual simulations comparing their performance to existing cap-based regulations. Applying this framework to newly collected data from the Japan Residency Matching Program, our counterfactual simulations reveal that the current cap-based policy generates substantial efficiency losses, whereas small, targeted subsidies can achieve similar distributional goals with significantly higher social welfare.

Distributional concerns have played a central role in Japan's residency matching market, prompting various proposals from both practitioners and researchers. In 2005, Japan introduced a mandatory two-year residency program and adopted a centralized matching system known as the Japan Residency Matching Program (JRMP). However, geographic disparities in physician distribution have posed persistent challenges to equitable access to healthcare since the program's inception. In response, the JRMP began capping the number of residency positions in urban areas in 2008 to encourage matches in underserved regions.

The JRMP data highlight important limitations of the current cap-based policy and point to its inefficiency. Despite progressively tighter caps on urban positions, geographic imbalances in residency placements persist. Moreover, the introduction of regional caps has coincided with a rising rate of unmatched applicants. These patterns suggest the current policy is an inefficient instrument, motivating our development of a theoretical framework to study alternative interventions. Crucially, the data also reveal substantial salary variation, with less popular programs—particularly those in rural areas—offering higher compensation to attract medical students. This observation motivates our adoption of a transferable utility (TU) framework that endogenizes these transfers, allowing us to formally evaluate the efficiency of the status quo quantity restrictions against price-based policies like taxes and subsidies.

In our model, each position belongs to a region, and caps and floors on the number of matches in each region are exogenously determined. A policymaker aims to satisfy these constraints, but they may not be satisfied in an equilibrium matching formed by agents without intervention. If the policymaker has complete knowledge of market participants' preferences, she can induce the matching that maximizes social surplus among all matchings that satisfy the regional constraints via a taxation policy. The taxes and subsidies correspond to the Lagrange multipliers associated with the regional constraints

in the social surplus maximization problem. This result implies that the taxation policy strictly outperforms any cap-based policy, serving as a useful benchmark for evaluating the efficiency of alternative policies, as it achieves the constrained-optimal.

In practice, policymakers rarely have full knowledge of individual preferences and must infer them from past matching outcomes. To address this, we embed our framework in the model introduced by Galichon and Salanié (2021*a*), which enables identification of the primitives of the transferable utility model using only aggregate-level matching data. This allows us to compute the optimal taxation policy from observed data (Theorem 1) and to conduct counterfactual simulations to compare welfare across alternative interventions, including cap-based regulations (Proposition 1). These simulations also allow us to quantify the relative inefficiencies of different policy tools.

For the empirical analysis, we start by defining the transfer between matched pairs and then consider a measurement model for this transfer. To this end, we model the baseline utilities of both sides for each pair of agents. The transfer from the hospital to the matched resident represents the gap between this baseline utility and the equilibrium utility level. By aggregating this transfer across the pairs within each hospital, we construct an aggregate-level transfer. We then introduce a measurement model for this aggregate-level transfer, which is linear in observed monthly salary. Our estimation proceeds in two steps: first, we estimate the aggregate-level surplus split for each pair of a medical school and a hospital, following Galichon and Salanié (2021*a*); second, based on these first-step estimates, we recover the parameters in the baseline utilities and the measurement model.

The estimation results partly align with those in the existing literature while also presenting a departure from its assumptions. On the doctor's side, our estimates indicate that factors such as the distance between the hospital and the doctor's alma mater, as well as the hospital size, are significant. Furthermore, we find that the number of previous matches is an important determinant of doctors' preferences. These findings are consistent with the existing literature, which suggests that hospitals are horizontally differentiated from the perspective of doctors. On the hospital's side, we observe that hospitals exhibit horizontal preferences similar to those of doctors: doctors from distant regions are less preferred, in addition to considerations of quality. Such horizontal preference structure on the hospital side is not allowed in the existing literature, which hinders the identification of preferences from matching data.

Our empirical estimates inform a series of counterfactual simulations designed to quantify the inefficiency of existing regulations and evaluate policy alternatives. First, a comparison between the status quo and an unregulated benchmark reveals that the current urban caps, while effective at redirecting some residents, generate a substantial aggregate welfare loss. Second, to distinguish the cost of the distributional goal itself from the cost of the policy instrument, we simulate an optimal subsidy policy designed to achieve the same resident distribution as the current cap-based system. We find that a relatively small, targeted subsidy can meet this goal at a significantly higher level of social welfare. This result indicates that the vast majority of inefficiency stems not from the distributional constraint itself, but from the choice of regulatory instrument, the caps. Our findings therefore suggest that policymakers can achieve distributional goals more efficiently by replacing blunt quantity restrictions with carefully designed monetary interventions. Finally, we test the limits of non-monetary interventions by simulating the Flexible Deferred Acceptance mechanism, a sophisticated cap-setting algorithm proposed by Kamada and Kojima (2015). Our results show that even this advanced mechanism generates substantial welfare losses compared to an optimal subsidy policy. This finding reinforces our conclusion that quantity restrictions are an inherently inefficient instrument for achieving distributional goals.

Related literature Matching with constraints, initiated by Kamada and Kojima (2015), has attracted considerable attention across various fields due to its applicability to real-world settings (Abdulkadiroğlu and Sönmez, 2003; Ehlers et al., 2014; Kojima, 2012; Hafalir, Yenmez and Yildirim, 2013; Fragiadakis and Troyan, 2017). A key departure from their approach is our adoption of the TU framework in a matching market with distributional constraints.

Recent studies by Kojima, Sun and Yu (2020) and Jalota, Ostrovsky and Pavone (2025) investigate broad classes of distributional constraints, primarily focusing on the existence of stable outcomes. Yokote (2020) studies the design of optimal taxes and subsidies when the policymaker faces an interval constraint, which specifies the maximum and minimum number of workers to be hired by each firm. If a firm in their model is regarded as a region, their model and our model appear similar, but they differ in certain aspects. For instance, in our model, a distinction should be made between matching a doctor to a single hospital in a region and matching the same doctor to a different hospital

in the same region; however, such a distinction cannot be described in their models.[1] Our analysis focuses on a specific class of constraints, regional constraints, and studies the problem of a policymaker who designs a taxation policy to ensure that these constraints are satisfied. We also extend the model to account for unobserved heterogeneity and to accommodate empirical analysis using aggregate-level matching data.

Next, we highlight our contribution from the viewpoint of the empirical analysis of matching markets. The structural analysis of matching markets is widely accepted in many fields of economics, including labor economics and industrial organization, which also adopts a type of matching model to describe trade networks (Fox, 2018; Fox, Yang and Hsu, 2018). Notable methodological contributions are developed in Galichon and Salanié (2021*a*): the nonparametric identification result of the social surplus function in a transferable utility matching market, which is robust to distributional assumptions on unobserved heterogeneity, along with the corresponding estimators[2]. This study builds on the general framework of Galichon and Salanié (2021*a*), proposing an extended model that accommodates regional constraints. Our framework is thus applicable without assuming a specific distributional form for unobserved heterogeneity terms. Furthermore, diverging from Galichon and Salanié (2021*a*), we propose a formal estimation strategy that exploits a measurement of the transfer to quantify the parameter values in a monetary unit.

Lastly, we mention Agarwal (2015) as the closest empirical research to this study. Agarwal (2015) takes a non-transferable utility matching model to analyze the doctor-hospital matching market in the U.S. National Residency Matching Program (NRMP), where a centralized mechanism determines the matching, and salaries are determined almost exogenously. While Agarwal (2015) addresses the endogeneity of salaries using a control function approach, our study fully models the salary determination process. This difference reflects the variation in market environments: in Japan, the concentration in urban areas presents more severe issues, and salaries are used as a tool to attract more candidates. Furthermore, treating transfers endogenously allows us to recover preferences under counterfactual regulations and simulate the resulting matching outcomes. This feature is particularly useful for policymakers when evaluating potential regulatory interventions, which a non-transferable utility matching model, which treats transfers as

[1]This is analogous to the difference between the models for affirmative action such as Abdulkadiroğlu and Sönmez (2003) and the model of Kamada and Kojima (2015).

[2]The methodology in Galichon and Salanié (2021*a*) is also distinct compared to other methodologies dependent on distributional assumptions, such as the minimum score estimator proposed by Fox (2010, 2017); Fox and Bajari (2013) and the maximum likelihood estimator of Choo and Siow (2006).

exogenous, cannot offer.

1.1 Institutional Background of the JRMP

In 2004, Japan introduced a mandatory two-year residency training program. Under the new program, residents and hospitals are matched through a centralized mechanism called the Japan Residency Matching Program (JRMP). The JRMP is modeled after the National Resident Matching Program in the United States. Medical students and hospitals must register with the JRMP system and submit their respective preference rankings. Students are provided with information about hospitals, including hospital size, location, specifics of the training program, salary, and workload. Additionally, they can participate in job fairs and directly visit hospitals to obtain further details. Before submitting their preference lists, students are required to take examinations administered by each hospital they consider ranking. Once the preference lists are finalized, the JRMP employs the deferred acceptance algorithm to determine the matches. In 2023, the JRMP involved 10,202 students and 1,209 hospitals offering 10,895 positions. On average, each student listed 4.35 hospitals on their preference list. The algorithm successfully matched 87.9% of the students; specifically, 64.3% secured their first-choice hospitals, 16.3% their second choice, and 9.0% their third choice.[3]

Researchers and Japanese media have reported that the distributional disparity of physicians across regions has worsened following the introduction of the JRMP (Iizuka and Watanabe, 2016; Sakai et al., 2013; Endo, 2019). One primary factor contributing to this imbalance is that the centralized matching system allowed medical students to freely express their location preferences, leading to a concentration of applicants in non-university hospitals. Under the previous system, a majority of graduates would begin their residency at the university hospital affiliated with their medical school.[4] Many medical graduates were affiliated with university medical departments, which directed graduates to affiliated rural hospitals, maintaining a relatively balanced distribution. However, the JRMP effectively removed the obligatory connection between graduates and university departments, resulting in fewer graduates choosing positions in rural or underserved regions. Consequently, rural hospitals faced increasing difficulties recruiting

[3]Unmatched students may individually contact hospitals with vacancies or reapply through the matching process in the subsequent year.

[4]Although it was not mandatory, most medical students chose to undertake residency training after graduation.

physicians, exacerbating regional disparities in patient access to hospital services in these areas.

To mitigate these regional imbalances, the JRMP introduced regional caps starting in 2010. The process for setting regional caps involves several steps. First, the total number of residency positions nationwide is determined by multiplying the number of medical students by a fixed constant. This constant was approximately 1.22 in 2015 but is scheduled to be reduced gradually to 1.05 by 2025. After determining the nationwide cap, residency positions are allocated across prefectures using a formula that incorporates factors such as population size, medical school enrollment capacities, the current number of practicing doctors, geographic considerations, including physician density (the number of doctors per unit area), and populations in isolated or remote islands. This allocation method is intentionally designed to benefit underserved areas by disproportionately reducing the number of residency positions in urban regions. The rationale behind imposing regional caps is that by restricting positions available in urban areas and tightening overall capacity, more medical graduates will be encouraged or compelled to pursue training in rural regions, under the expectation that residents who train in these underserved areas are more likely to remain and practice there long-term, thereby addressing existing regional disparities.

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 between medical schools and hospitals, the characteristics of these institutions relevant to their preferences, and the salaries paid to residents during their internships. We begin by describing the data sources, present the descriptive statistics in Section 2.1, and discuss the empirical matching patterns observed in the market in Section 2.2.

The matching patterns between medical schools and hospitals, i.e., the numbers of matches between any given pair, are calculated based on the "Physician Registration Report," a certificate that summarizes the personal information of licensed physicians, including residents. This data source allows us to calculate the annual number of matches between specific medical schools and hospitals using information on their respective graduating institutions and training 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.[5] 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.[6]

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 of the last year as one of the covariates, we exclude the fiscal year of 2016. Panel A and Panel B in Table 1 summarize the results of JRMP. The environment of the matching market—characterized by the number of schools, students, hospitals, and available slots—remains stationary over these three years, with minimal entry or exit. The matching outcomes, such as the unmatched rate, also appear stable during this period. Based on these observations, we assume that the preference structures of schools and hospitals remain unchanged throughout the data period.

Despite the overall presence of unoccupied seats, 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. These figures suggest that rural areas, characterized by lower

[5]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' performance on practice exams and their success on university entrance exams.

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

Table 1. Environments and Outcomes of JRMP

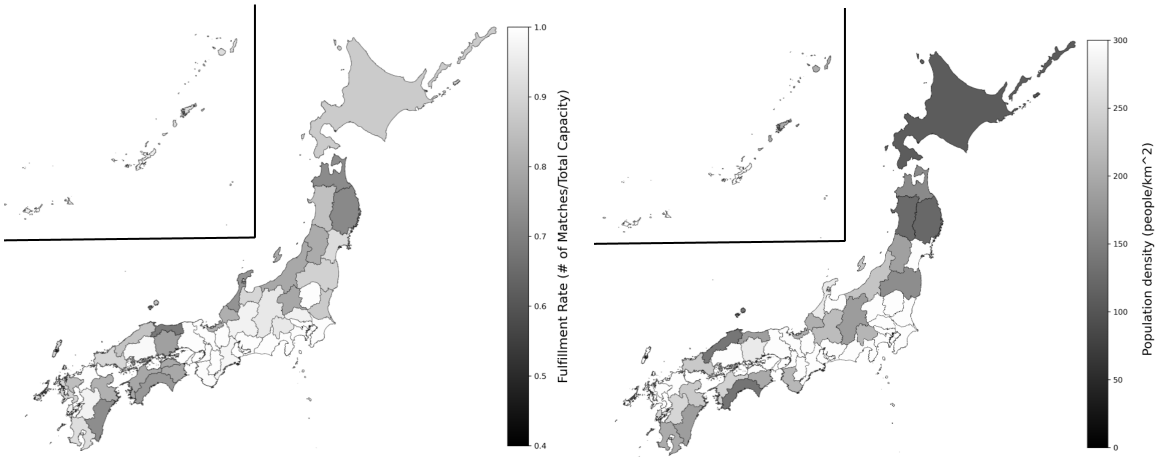
	2017	**2018**	**2019**
Panel A: Doctor side			
Number of schools	78	78	78
Number of students	9830	9916	9932
Number of matched students	8530	8369	8634
Number of unmatched students	1300	1547	1298
Unmatch rate (%)	13.22	18.48	15.03
Panel B: Hospital side			
Number of hospitals	1025	1025	1022
Number of total 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

population density, are less popular and experience lower fulfillment rates, even with the current tight caps on urban areas.

Table 2 presents descriptive statistics for the characteristics of medical schools and hospitals. Panel A summarizes the medical school variables. Medical schools are first categorized as either private or public. There are 51 public medical schools and 27 private ones. The T-score and graduation exam pass rate are used as proxies for student quality, with higher values indicating greater ability. Public medical schools tend to have higher average T-scores, suggesting that graduating from a public institution signals stronger student qualifications. This difference is statistically significant.

Panel B summarizes the hospital-side variables. Hospitals are classified in two ways. First, based on their affiliation with a university, they are categorized as either university or non-university hospitals. Approximately 88.3% of hospitals have no university affiliation. University hospitals tend to have more beds on average, reflecting their typical role as regional flagship hospitals. The second classification is based on location. Hospitals situated in one of six prefectures—Tokyo, Kanagawa, Aichi, Kyoto, Osaka, and Fukuoka—where official caps on the number of matches are imposed under the JRMP, are defined as being in urban areas. The remaining hospitals are considered rural. Although this definition includes only six of the forty-seven prefectures in total, 33.7% of all hospitals fall into the urban category. Urban hospitals also tend to be larger in size.

We examine the total variation in salaries in the JRMP market. The latter half of



(a) Fulfillment Rates (b) Population Density

Figure 1. Fulfillment Rate and Population Density

Panel B 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.[7] For comparison, Table 1 of Agarwal (2015) reports 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. This variation reflects salary differentials across regions: on average, rural hospitals offer significantly higher monthly wages—by $430 (63,760 JPY)—than their urban counterparts. At the same time, affiliation with a university is associated with significantly lower pay, with university hospitals paying $692 (101,000 JPY) less on average.

2.2 Empirical Matching Pattern

In line with Agarwal (2015), we estimate a set of regressions to characterize the matching patterns between hospitals and students. Due to the unavailability of individual matching data, we first calculate the weighted averages of the characteristics of matched partners within each medical school and hospital. These aggregated characteristics are then regressed on the covariates of the medical schools and hospitals. We also regress the unmatch rate of each hospital and school on their characteristics to identify the types of students who are more likely to go unmatched in this market. We use random effects models to account for the correlation between unobserved error terms and observed co-

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

Table 2. Summary Statistics of Medical-School and Hospital Covariates

	Category	Count	Mean	Std	Min	Max
Panel A. Medical Schools						
T-score	Private	27	64.96	2.17	62	72
	Public	51	66.38	2.66	63	74
Panel B. Hospitals						
Number of beds	Rural	684	417.99	155.92	36	1195
	Urban	348	472.48	217.84	38	1379
	Non-university	911	411.08	147.73	36	1097
	University	121	626.74	272.77	295	1379
Monthly Salary (×100,000 JPY)	Total	1032	3.864	0.995	1.800	8.550
	Rural	684	4.079	0.984	1.800	8.550
	Urban	348	3.452	0.875	1.900	6.847
	Non-university	911	3.986	0.978	2.225	8.550
	University	121	2.976	0.566	1.800	4.720

* The summary statistics for monthly salary are based on the three-year average from 2017 to 2019 for each of the 1,032 hospitals.

variates in our panel matching data.[8] We compute cluster-robust standard errors where the cluster is set to the unit level: the hospital level for the hospital perspective and the school level for the school perspective.

Table 3 reports the regression results on matching patterns from the perspective of hospitals. The first and second column dependent variables represent the quality of matched students, so the positive correlations with hospital size shown in the first row are in line with expectations. The second row shows a notable finding about the role of urban location. It is natural to expect that urban hospitals attract higher-quality students—as urban areas are generally more preferred by students—this pattern is indeed observed for difficulty. However, urban hospitals are less likely to match with highly qualified students from public universities. This pattern suggests that public universities play a significant role in supplying physicians to rural areas. There is no statistically significant pattern in the distance between the matched students and their universities. As for the unmatch rate, urban hospitals exhibit fewer vacancies, as expected.

One straightforward explanation for the frequent matching between public universities and rural hospitals is geographic proximity: every prefecture hosts at least one public medical school. At the same time, higher salaries appear to serve as a strong incentive

[8]Since our main interest lies in the coefficients on time-invariant covariates, such as the urban hospital dummy and public university dummy, we do not use a fixed effects model.

Table 3. Matching from Hospital Viewpoint

	(1) Public	(2) Difficulty	(3) Distance	(4) Unmatch Rate
ln Beds	0.0576***	6.286***	9.081	-0.146***
	(0.0222)	(1.190)	(10.29)	(0.0227)
Urban Hospital	-0.113***	3.179***	2.837	-0.148***
	(0.0176)	(0.822)	(9.661)	(0.0150)
Observations	3072	3072	3072	3072
Year dummy	√	√	√	√

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

Table 4. Matching from School Viewpoint

	(1) Urban Hospital	(2) ln Wage	(3) Distance	(4) Unmatch Rate
Public University	-3.055***	1.335***	425.7	-1.416**
	(1.088)	(0.391)	(539.0)	(0.633)
Difficulty	0.0115	0.00452	-16.16**	-0.0220**
	(0.0146)	(0.00433)	(6.540)	(0.00956)
Observations	234	234	234	234
Year dummy	√	√	√	√

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

to retain talented students from public universities in rural areas after graduation. As shown in Table 4, which presents the regression results from the perspective of medical schools, graduates from public universities are more likely to match with rural hospitals and receive higher salaries. This finding supports the view that salary is a significant factor in students' preferences to sustain an equilibrium in which rural hospitals offer higher wages to retain talented graduates. Table 4 also shows that graduates from more competitive universities tend to match with nearby hospitals, suggesting that distance is an important factor in residents' choices. As expected, graduates from more prestigious universities are less likely to go unmatched.

3 Model

Matching market with regional constraints 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. 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 $j \in J$.

There is a policymaker who faces an additional condition called regional constraints. There are L regions, denoted by $Z = \{z_1, z_2, \ldots, z_L\}$, 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 $\bar{o}_z \in \mathbb{R}_+ \cup \{\infty\}$, where $\underline{o}_z < \bar{o}_z$ for each z. 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 at least one feasible matching exists that satisfies regional constraints.[9]

Stable outcome under taxation policy 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* $\Phi_{ij} \in \mathbb{R}$. Denote by Φ_{i,j_0} and $\Phi_{i_0,j}$ the payoffs $i \in I$ and $j \in J$ receive when unmatched, respectively. 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 $w = (w_z)_{z \in Z}$, each matched pair divides the *gross joint surplus* $\Phi_{ij} - w_{z(j)}$ instead of the net joint surplus. [10] The stable outcome under a

[9] Formally, we assume $\underline{o}_z \leq |\{j \colon j \in z\}|$ for each $z \in Z$, and $|I| \geq \sum_z \underline{o}_z$.

[10] In principle, the policymaker could set pair-specific taxes. We restrict our definition of a taxation policy to be uniform within each region; however, this is without loss of generality: the welfare-

taxation policy is defined as follows:

Definition 1 (Stable outcome)**.** *Given the matching market* (I, J, Z, z, Φ),[11] *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* $d_{ij} = 1$.

While it is reasonable to assume that participants form a stable outcome in many frictionless and decentralized environments, such as certain labor or marriage markets, the appropriateness of this assumption depends on the specific application. In Appendix A.1, we show that any stable outcome can be supported as an equilibrium of a stylized game that mirrors the JRMP's matching process.

A *cap-based policy* is represented by removing a subset of slots from the market. Let $J' \subseteq J$ denote the set of slots that remain after the policy is applied; this subset characterizes the cap-based policy. Operationally, this removal is equivalent to imposing an upper bound (a cap) on the number of slots. Let $\Phi' \coloneqq (\Phi_{ij})_{i \in I,\, j \in J'}$ denote the submatrix of Φ corresponding to the remaining doctors and slots. Given J', the agents form a stable outcome in the reduced market (I, J', Z, z, Φ') under the assumption that taxation policy is absent, that is, $w \equiv 0$.

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

maximizing taxation policy is, in fact, uniform, imposing the same tax on all pairs within a given region (see Section 4.1 for the related discussion and Appendix A.2 for the formal proof).

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

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 $z \in Z$.

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

4.1 Relative Efficiency of Caps and Taxation

To build intuition for our main analysis, this subsection begins with a stylized model that abstracts from unobserved heterogeneity. In this benchmark case, we provisionally assume that the policymaker knows all joint surplus values, Φ_{ij}. While this assumption is relaxed later, it clearly illustrates the properties of the *optimal taxation policy*, which is defined below, and its relationship to cap-based policies, which remain valid even in the presence of unobserved heterogeneity.

Given a matching d, the *social surplus* is defined as $\sum_{i,j} d_{ij} \Phi_{ij}$, the total net joint surplus generated by the matching d. Consider the following social surplus maximization problem subject to regional constraints:

$$
(P_0)\left|\begin{array}{lll}
\underset{d\in\{0,1\}^{I\times J}}{\text{maximize}} & \displaystyle\sum_{(i,j)\in I}\Phi_{ij}d_{ij}+\sum_{i\in I}\left(1-\sum_{j\in J}d_{ij}\right)\Phi_{i,j_0}+\sum_{j\in I}\left(1-\sum_{i\in I}d_{ij}\right)\Phi_{i_0,j} & \\
\text{subject to} & \displaystyle\sum_{j\in J}d_{ij}\leq 1 & \forall i\in I,\\
& \displaystyle\sum_{i\in I}d_{ij}\leq 1 & \forall j\in J,\\
& \underline{o}_z\leq\displaystyle\sum_{j\in z}\sum_{i\in I}d_{ij}\leq\bar{o}_z & \forall z\in Z,
\end{array}\right.
$$

Let d^* denote a solution to (P_0); that is, d^* is the matching that maximizes social surplus subject to the regional constraints.

Given full knowledge of the joint surplus matrix Φ, there exists an *optimal* taxation policy that implements d^* as a stable outcome. This implies that, within our framework, the optimal taxation policy outperforms any cap-based policy: no cap-based policy that satisfies the same regional constraints can generate a higher social welfare than the optimal taxation policy.

The optimal taxes and subsidies correspond to the Lagrange multipliers associated with the binding regional constraints and can be computed via linear programming. (see Appendix A.2 for details.) The optimal subsidy, w_z^*, for a region z with a binding floor constraint can be interpreted as the *shadow price* of attracting one additional doctor to region z. Intuitively, consider a marginal doctor i who decides to relocate to region z in response to the subsidy. A new match in region z generates a net joint surplus of $\Phi_{ij}+w_z^*$, while the doctor's previous match in another region yielded $\Phi_{ij'}$. At the margin, these two values must be equal, i.e., $\Phi_{ij}+w_z^*=\Phi_{ij'}$. Therefore, under the optimal subsidy, the total social surplus, defined as the sum of gross surpluses generated by all matches, decreases by w_z^*.

A key feature of this optimal taxation policy is its simplicity: it prescribes a single tax or subsidy w_z^* per region, applied uniformly to all matches within that region. Despite the theoretical possibility of setting pair-specific taxes, the welfare-maximizing policy takes this uniform structure, which greatly facilitates implementation.

The following example illustrates that the relative efficiency of a cap-based policy, measured as the ratio of its social welfare to the optimal taxation benchmark, can take any value in $(0,1)$, depending on the model's parameters. This underscores the need for empirical analysis to quantify the actual welfare loss in any given application.

Example 1. Consider two doctors, i_1 and i_2; two regions, z_1 and z_2; and two slots, $j_1 \in z_1$ and $j_2 \in z_2$. Region z_2 is subject to a floor constraint $\underline{o}_{z_2} = 1$, while there are no other regional constraints. The value of outside options is normalized to zero, and the joint surplus matrix is given by

$$\Phi := \begin{array}{c} \\ i_1 \\ i_2 \end{array} \begin{array}{c} \begin{array}{cc} j_1 & j_2 \end{array} \\ \begin{bmatrix} a & \delta \\ b & -\varepsilon \end{bmatrix} \end{array},$$

where $a, b, \delta, \varepsilon > 0$ and $a \geq \delta + b + \varepsilon$.

No intervention The stable matching pairs i_1 with j_1, leaving i_2 unmatched. The resulting social welfare is a, but the regional constraint is violated.

Cap-based policy The policymaker meets the constraint by removing the urban slot, j_1. The reduced surplus matrix becomes

$$\Phi' := \begin{array}{c} \\ i_1 \\ i_2 \end{array} \begin{array}{c} \begin{array}{c} j_2 \end{array} \\ \begin{bmatrix} \delta \\ -\varepsilon \end{bmatrix} \end{array}.$$

In this case, i_1 is matched with j_2, yielding social welfare δ.

Optimal subsidy policy The policymaker retains both slots and subsidizes matches in region z_2 by ε. The resulting gross surplus matrix is

$$\begin{array}{c} \\ i_1 \\ i_2 \end{array} \begin{array}{c} \begin{array}{cc} j_1 & j_2 \end{array} \\ \begin{bmatrix} a & \delta + \varepsilon \\ b & 0 \end{bmatrix} \end{array}.$$

The matching i_1 to j_1 and i_2 to j_2 forms a stable outcome, achieving social welfare of $a - \varepsilon$.

Comparison The ratio $\delta/(a-\varepsilon)$ compares the social welfare under the cap-based policy to that under the taxation policy. This ratio can take any value in $(0, 1]$, while satisfying

the condition $a \geq \delta + b + \varepsilon$.[12]

This example also illustrates how the inefficiency of a cap-based policy arises. As a blunt instrument, the cap satisfies the rural constraint by eliminating the high-surplus urban slot, j_1. This forces the most productive doctor, i_1, into a significantly less-valued match, sacrificing the large surplus generated by the (i_1, j_1) pair. A targeted subsidy, in contrast, achieves the goal more efficiently. By making the undesirable rural match, (i_2, j_2), just palatable enough for the marginal doctor, i_2, it fulfills the constraint while leaving the most productive match in the market, (i_1, j_1), undisturbed.

4.2 Preliminary Results

To account for unobserved heterogeneity in preferences, this subsection reviews the framework of Galichon and Salanié (2021*a*), which links the individual-level objects, $(\Phi_{ij})_{ij}$, $(d_{ij})_{ij}$, $(u_i)_i$, and $(v_j)_j$, introduced in Section 3 to their aggregate-level counterparts required for our analysis.

Discrete Choice Representation For any pair (i, j) with doctor i of type x and slot j of type y, we assume that the individual-level joint surplus decomposes as

$$\Phi_{ij} = \Phi_{xy} + \varepsilon_{iy} + \eta_{xj},$$

where Φ_{xy} is the *aggregate-level joint surplus* and ε_{iy} and η_{xj} are independent error terms. For each type x and every doctor $i \in x$, the vector $(\varepsilon_{iy})_{y \in Y_0}$ is drawn from a distribution $P_x \in \Delta(\mathbb{R}^{|Y|+1})$. Similarly, for each type y and every slot $j \in y$, the vector $(\eta_{xj})_{x \in X_0}$ is drawn from a distribution $Q_y \in \Delta(\mathbb{R}^{|X|+1})$.

Assumption 1 (Independence)**.** *The error terms are independent across all doctors and slots.*

Assumption 2 (Additive Separability)**.** *There exists a matrix $(\Phi_{xy})_{(x,y)\in T}$ such that:*

1. *For every $x \in X$, $y \in Y$, $i \in x$, and $j \in y$,*

$$\Phi_{ij} = \Phi_{xy} + \varepsilon_{iy} + \eta_{xj};$$

[12]For example, the ratio equals 1 if $\delta = a - \varepsilon$, $b = \varepsilon$, and $\varepsilon \leq a/2$. In contrast, suppose that $\delta = \varepsilon$ and $b \leq a - 2\varepsilon$. As δ and ε approach zero, the ratio converges to zero.

2. For every $x \in X$ and $y \in Y$,

$$\Phi_{i,y_0} = \varepsilon_{i,y_0} \quad \textit{and} \quad \Phi_{x_0,j} = \eta_{x_0,j}.$$

We define the *aggregate-level utilities* U_{xy} and V_{xy}, which depend solely on types, as follows. For each $x \in X$ and $y \in Y$, let

$$U_{xy} \coloneqq \min_{i\,:\,x(i)=x} \{u_i - \varepsilon_{iy}\}, \quad V_{xy} \coloneqq \min_{j\,:\,y(j)=y} \{v_j - \eta_{jx}\}, \tag{1}$$

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

Lemma 1 (Galichon and Salanié (2021*a*))**.** *Let (u, v) be a payoff profile corresponding to 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.3. □

Large Market Approximation By Lemma 1, the social welfare on the doctor side, defined as the sum of the equilibrium payoffs, can be expressed as

$$\sum_{i \in I} u_i = \sum_{x \in X} n_x \left(\frac{1}{n_x} \sum_{i \in x} \max_{y \in Y_0} \{U_{xy} + \varepsilon_{iy}\} \right).$$

When the number of doctors of type x, denoted by n_x, is large, the average

$$\frac{1}{n_x} \sum_{i \in x} \max_{y \in Y_0} \{U_{xy} + \varepsilon_{iy}\}$$

is well approximated by its expected value: $\mathbb{E}_{\varepsilon_i \sim P_x} [\max_{y \in Y_0} \{U_{xy} + \varepsilon_{iy}\}]$. Thus, the social welfare on the doctor's side becomes

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

Similarly, if the number of slots of type y, denoted by m_y, is large, the social welfare on the hospital side is approximated by

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

Under the assumption that the cumulative distribution functions of the error terms are continuously differentiable, the Williams-Daly-Zachary theorem (McFadden, 1980) implies that

$$\frac{\partial G}{\partial U_{xy}}(U) = n_x \Pr(\text{doctor of type } x \text{ chooses a slot of type } y).$$

For sufficiently large n_x, the product $n_x \Pr(\text{doctor of type } x \text{ chooses a slot of type } y)$ provides a good approximation of μ_{xy}.

Assumption 3 (Smooth Distribution)**.** *For each x and y, the cumulative distribution functions P_x and Q_y are continuously differentiable.*

4.3 Design and Evaluation of Policies with Unobserved Heterogeneity

An aggregate-level matching market is characterized by the tuple

$$\mathcal{M} \coloneqq (X, Y, n, m, Z, z, \Phi, P, Q),$$

where $n \coloneqq (n_x)_x$, $m \coloneqq (m_y)_y$, $\Phi = (\Phi_{xy})_{xy}$, $P = (P_x)_x$, and $Q = (Q_y)_y$. The policymaker, facing unobserved heterogeneity, seeks (i) to determine the optimal taxation policy w that maximizes social welfare subject to regional constraints and (ii) to compute the matching and social welfare outcomes for various policies. For these purposes, we formulate optimization problems that depend solely on the aggregate-level surplus $(\Phi_{xy})_{x,y}$. An additional technical assumption on the error term distributions guarantees that G and H are strictly convex,[13] thereby ensuring a one-to-one correspondence between an aggregate-level matching μ and the aggregate-level utilities U and V given taxation policy w.

[13] See Appendix A.4 for details.

Assumption 4 (Full Support)**.** *For each x and y,* $\mathrm{supp}(P_x) = \mathbb{R}^{|Y_0|}$ *and* $\mathrm{supp}(Q_y) = \mathbb{R}^{|X_0|}$.

We now define the following concave programming problem, which is a counterpart of the social surplus maximization problem (P_0) for the case without unobserved heterogeneity:

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

where

$$
\mathcal{E}(\mu) \coloneqq -G^*(\mu) - H^*(\mu),
$$

and G^* and H^* denote the Legendre-Fenchel transforms of G and H, respectively.[14] Its dual problem is given by

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

The primal problem (P) admits an optimal solution because its objective function is continuous and its feasible set is compact and nonempty. By strong duality,[15] the dual problem (D) attains the same optimal value. Moreover, both (P) and (D) possess unique solutions, as we show below.

The optimal value of (P) and (D) represents the social surplus generated under equilibrium matching μ. [16] The following theorem states that, given regional constraints

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

[15] The constraints of the primal problem satisfy the weak Slater condition, as they are affine in μ.

[16] For any w, we can first rewrite the dual problem of (P_0) using the aggregate level objects. We can show that (D_w) defined in the statement of Proposition 1 coincides with the dual problem under the large market approximation (see Section 1.4 of Galichon and Salanié (2021*b*) for more details.) Then, this statement follows from the second part of Theorem 1. See also Appendix B for more details about

$(\bar{o}_z, \underline{o}_z)_z$, the optimal taxation policy w^* together with the corresponding aggregate-level matching and utilities (μ, U, V) are characterized by the solutions to (P) and (D).

Theorem 1. *Assume that Assumptions 1–4 hold. Fix any aggregate-level matching market $\mathcal{M}$. Then,*

1. *There is a unique solution to (P) and (D), denoted by μ^* and $(U^*, V^*, \bar{w}^*, \underline{w}^*)$, respectively.*

2. *Define, for each z,*

$$w_z^* := \mathbb{1}\{\bar{w}_z^* > 0\}\bar{w}_z^* - \mathbb{1}\{\underline{w}_z^* > 0\}\underline{w}_z^*.$$

Then, w^* is the optimal taxation policy, and (μ^*, U^*, V^*) constitute the corresponding aggregate-level matching and utilities.

Proof. See Appendix A.4. □

We can also obtain the aggregate-level matching and utilities (μ, U, V) corresponding to any given taxation policy w by solving the following pair of optimization problems.

Proposition 1. *Assume that Assumptions 1–4 are satisfied. For any aggregate-level matching market $\mathcal{M}$ and any taxation policy w, the resulting aggregate-level matching and utilities, denoted by $(\mu(w), U(w), V(w))$, are characterized as the solutions to*

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

and

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

the large market approximation.

For a given taxation policy w, the tuple $(\mu(w), U(w), V(w))$ defined in Proposition 1 is referred to as an *aggregate equilibrium (AE)* under w. When w is optimal, the equilibrium is termed the *efficient aggregate equilibrium (EAE)*.

5 Empirical Strategy

We begin by mapping the primitives and equilibrium objects in the model described in Section 3 and 4 to their empirical counterparts. 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: school s and hospital h are 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. Let Z denote the set of regions and $z(h)$ denote the region to which hospital h belongs.

For a matching market of one year, the aggregate-level joint surplus is denoted by Φ_{sh}. The net joint surplus satisfies the equality: $\Phi_{ij} = \Phi_{s(i)h(j)} + \varepsilon_{ih(j)} + \eta_{s(i)j}$ for each $i \in I$ and $j \in J$. Note that $\varepsilon_{ih}, \eta_{sj}$ are independent error terms. The matching $(d_{ij})_{i,j}$ is not observable to the researcher. Instead, the available data comprises the aggregate-level matching $(\mu_{sh})_{s,h}$, which is the number of matches between medical school s and hospital h. We suppose that the aggregate-level matching in a year constitute the aggregate equilibrium under neither tax nor subsidy.

As described in Section 4.3, the sum of the aggregate-level utilities, which is aggregate-level joint surplus, is enough to simulate the counterfactual equilibrium and the corresponding optimal taxation policy according to any regional constraints. And, following Galichon and Salanié (2021*a*), we can identify the aggregate-level utilities of both sides by the realized aggregate-level matching given the distributions of the idiosyncratic error terms.

However, to evaluate these counterfactual objects in practice, they need to be quantified in monetary terms. To this end, Section 5.1 imposes additional structure on the net joint surplus, which lets us define individual-level transfer and link them to an aggregate-level transfer. Section 5.2 then parameterizes the primitives introduced in Section 5.1 and derives moment conditions for identification—both for aggregate transfers and for their measurement via monthly salaries. Finally, Section 5.3 details the estimation procedure based on these moment conditions.

5.1 Transfer

The *base utility* of i when matched with j is the utility achieved by i when matching with j net of transfer, and is denoted by U_{ij}^{base}. Similarly, the base utility of j when matched with i is denoted by V_{ij}^{base}. Under the transferable utility framework, the net joint surplus yielded by i and j is the sum of the base utilities of both sides, i.e., we have $\Phi_{ij} = U_{ij}^{\text{base}} + V_{ij}^{\text{base}}$. Furthermore, in accordance with additive separability (Assumption 2), we re-interpret the error terms as the unobserved taste shocks. As a part of base utilities which are determined by the observable characteristics, we introduce U_{sh}^{base} and V_{sh}^{base} as *aggregate-level base utility*. These are model primitives to be parametrized below.

Assumption 5. $U_{ij}^{\text{base}} = U_{sh}^{\text{base}} + \varepsilon_{ih}, V_{ij}^{\text{base}} = V_{sh}^{\text{base}} + \eta_{sj}$.

Suppose a stable outcome materializes in a matching market. Under Assumption 5 and by the definition of a stable outcome (Definition 1), for each pair (i, j) we can select a real number τ_{ij} such that

$$u_i - U_{ij}^{\text{base}} \;\geq\; \tau_{ij} \;\geq\; V_{ij}^{\text{base}} - v_j.$$

We refer to any such τ_{ij} as an *equilibrium transfer*. When i and j are matched, the equilibrium transfer is uniquely determined; when they are not matched, there are infinitely many admissible equilibrium transfers for that pair. By choosing one admissible value in each unmatched case, we obtain a set of equilibrium transfers associated with any stable outcome.[17]

We introduce aggregate counterparts of τ_{ij}. Specifically, define the *aggregate-level transfer* for hospital h, denoted by ι_h, as the average equilibrium transfer from a post at h to the doctors matched with those posts:

$$\iota_h \;\equiv\; \frac{1}{M_h} \sum_{(i,j):\, j \in h,\, i \text{ matches with } j} \tau_{ij},$$

where M_h is the number of such matched pairs at hospital h. Note that ι_h may not be directly observed in the salary paid by the hospital to matched doctors, because it includes

[17] By definition of the equilibrium transfer, $u_i = \max_j\{U_{ij}^{\text{base}} + \tau_{ij}\}$, $\quad v_j = \max_i\{V_{ij}^{\text{base}} - \tau_{ij}\}$. Given τ_{ij}, the surplus split (u_i, v_j) is generated by the participants' discrete choices; hence we interpret τ_{ij} as the equilibrium transfer. This formulation of transfer follows Choo and Siow (2006); Galichon, Kominers and Weber (2019).

unobserved transfers which reflect program-specific idiosyncrasies—such as mentorship quality, culture, curriculum, rotation structure, and other non-pecuniary amenities.

5.2 Parametrization and Moment Conditions

We use $X_{sh}^{U,\text{base}}$ and $X_{sh}^{V,\text{base}}$ to denote the variables composing U_{sh}^{base} and V_{sh}^{base}, respectively. X is used to denote the collection of them: $X = \left\{\left(X_{sh}^{U,\text{base}}, X_{sh}^{V,\text{base}}\right)\right\}_{sh}$. We assume linear structures: $U_{sh}^{\text{base}} = X_{sh}^{U,\text{base}\top}\beta_U$, $V_{sh}^{\text{base}} = X_{sh}^{V,\text{base}\top}\beta_V$. Our parameters of interest are β_U and β_V. We use θ to indicate the vector of these parameters: $\theta \coloneqq (\beta_U, \beta_V)$.

Imagine that ι_h is directly observed for every hospitals. Proposition 2 below states that the conditional expectation of aggregate-level transfer is linear in parameters θ and a term of equilibrium objects. Galichon and Salanié (2021*a*) show that the aggregate-level utilities, U and V, are identified by the aggregate-level matching. Hence, we can treat them as observed variables in the equalities once these are estimated and use the equalities as moment conditions for identifying θ.

Proposition 2. *Let* $\omega_{sh} = \frac{\mu_{sh}}{\sum_{s'}\mu_{s'h}}$. *Then,*

$$\begin{cases} \mathbb{E}\left[\iota_h - \sum_s \omega_{sh}\left(U_{sh} - X_{sh}^{U,\text{base}\top}\beta_U\right) \mid X, \mu\right] = 0, \\ \mathbb{E}\left[\iota_h - \sum_s \omega_{sh}\left(X_{sh}^{V,\text{base}\top}\beta_V - V_{sh}\right) \mid X, \mu\right] = 0. \end{cases} \tag{2}$$

Proof. See Appendix A.5. □

Remark 1. Proposition 2 implies that *selection bias* induced by conditioning on the realized matching is dealt with under aggregate-level model. Suppose we directly observe the individual-level transfer τ_{ij}, but only for matched pairs. Because of this selection, constructing moment conditions requires the conditional expectation $\mathbb{E}[\tau_{ij} \mid X, d]$, where d denotes the observed matching. Such conditional expectations are difficult to handle in practice: they require simulation and are generally nonlinear in the parameters. We exploit the identification results of Galichon and Salanié (2021*a*) to circumvent this problem. With aggregate-level matching data, even after conditioning on the observed matching, the moment conditions based on aggregate transfers are linear in the parameters and admit a closed-form representation as shown in (2).

In practice, ι_h is unobservable. We additionally introduce a measurement model to handle this issue: monthly salary paid from hospital to doctors plays a role of a measure

of this aggregate-level transfer. The salary paid by a hospital h is denoted by S_h. We assume both of doctors' and hospitals' aggregate-level transfer have quasi-linear structure with respect to monetary transfers, S_h, i.e.,

$$\begin{aligned} \iota_h &= \gamma_{0,U} + \gamma_{1,U} S_h + \psi_h^U \\ -\iota_h &= \gamma_{0,V} + \gamma_{1,V} S_h + \psi_h^V. \end{aligned} \tag{3}$$

Note that $\gamma_{1,V}$ is expected to be negative because the salary is the amount of money paid to the doctor by the hospital. I use γ to denote the set of parameters in the measurement model. ψ_h^U is the unobserved transfer from the hospital to the matched doctors, and ψ_h^V is the unobserved transfer from the doctor to the hospital.

The unobserved transfer is likely to be correlated with the observed salary. To avoid this endogeneity concern and identify the parameters in measurement model, we rely on instrumental variables, denoted by Z_h. For such IVs, we propose the characteristics of surrounding hospitals as in Berry, Levinsohn and Pakes (1995). These variables are *relevant*: they shift the competitive pressure for talent, thereby affecting the salary S_h chosen in the current market. They also satisfy *exclusion*: the unobserved transfer components ψ_h^U and ψ_h^V are hard to adjust in response to nearby-market conditions due to their high adjustment costs. In summary, the moment conditions for the measurement models are as follows:

$$\mathbb{E}\left[\psi_h^U Z_h\right] = \mathbb{E}\left[\psi_h^V Z_h\right] = 0, \quad \mathbb{E}\left[S_h Z_h\right] \neq 0. \tag{4}$$

5.3 Estimation

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 and γ using the estimated aggregate-level utilities and the observed salaries.

5.3.1 First Step

Despite the nonparametric identification results obtained in Galichon and Salanié (2021*a*), we estimate the parametrized version of the aggregate-level utilities. This is because some

pairs of schools and hospitals have zero matches in practice. Hence, in the first step, we use the moment matching estimator proposed in Galichon and Salanié (2021*a*) to estimate U_{sht} and V_{sht}.[18]

By formulating the aggregate-level matching 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^U_{sht} and X^V_{sht}, which is denoted by X^{poly}_{sht}. In other words, we model the aggregate-level utilities as follows: $U_{sht} = X^{\text{poly}\prime}_{sht}\beta^{\text{poly}}_U$, $V_{sht} = X^{\text{poly}\prime}_{sht}\beta^{\text{poly}}_V$. We use $\hat{\beta}^{\text{poly}}_U$ and $\hat{\beta}^{\text{poly}}_V$ as the estimated coefficients attached with the polynomials. And we denote the estimated aggregate-level utilities by $\hat{U}_{sht} \equiv X^{\text{poly}\prime}_{sht}\hat{\beta}^{\text{poly}}_U$ and $\hat{V}_{sht} \equiv X^{\text{poly}\prime}_{sht}\hat{\beta}^{\text{poly}}_V$.

Remark 2. (Interpretation of fixed effects) These fixed effects in the Poisson regression correspond to the Lagrange multipliers associated with a population constraints in the convex programming developed in Section 3.2 of Galichon and Salanié (2021*b*). Note that, while they take the form of fixed effects, they do *not* capture baseline utilities for each type—such as the value of remaining unmatched. As noted in Section 4.2, the average unmatched values are normalized to 0 in our model.

5.3.2 Second Step

By combining the moment conditions (2) and the measurement model (3), the estimating equations are specified as follows:

$$\sum_s \omega_{sht}\hat{U}_{sht} = \gamma_{0,U} + \gamma_{1,U}S_{ht} + \sum_s \omega_{sht}\left(X^{U,\text{base}\prime}_{sht}\beta_U\right) + \psi^U_{ht}$$
$$\sum_s \omega_{sht}\hat{V}_{sht} = \gamma_{0,V} + \gamma_{1,V}S_{ht} + \sum_s \omega_{sht}\left(X^{V,\text{base}\prime}_{sht}\beta_V\right) + \psi^V_{ht}.$$

When we take the weighted average of every variable in X^U_{sht} and X^V_{sht} as independent variables in the right-hand side, the above equations are linear in θ. We estimate these linear equations using the instrumental variables Z_h satisfying the conditions (4).

Specifically, Z_h is constructed by the characteristics of hospitals within a 10km radius around hospital h, we include: the number of university hospitals, the number of govern-

[18]For completeness, Appendix A.6 presents the relationship between the moment-matching estimator and the two-way fixed effects Poisson regression, complementing the discussion in Galichon and Salanié (2021*b*). Note that our estimation target is U_{sht} and V_{sht}. The difference from the case of Galichon and Salanié (2021*a*) is that we just need one side fixed effect.

ment hospitals, the average number of beds, the AE-based average distance to the alma maters of students who match to h, the AE-based average historical number of matches with those alma maters, and the AE-based average count of affiliations with those alma maters.

6 Empirical Results

In this section, we show the estimation results. In Section 6.1, we show the estimation results of our first step. In Section 6.2, we show the estimation results of our second step: the preference parameters of both sides of the market.

6.1 First Step

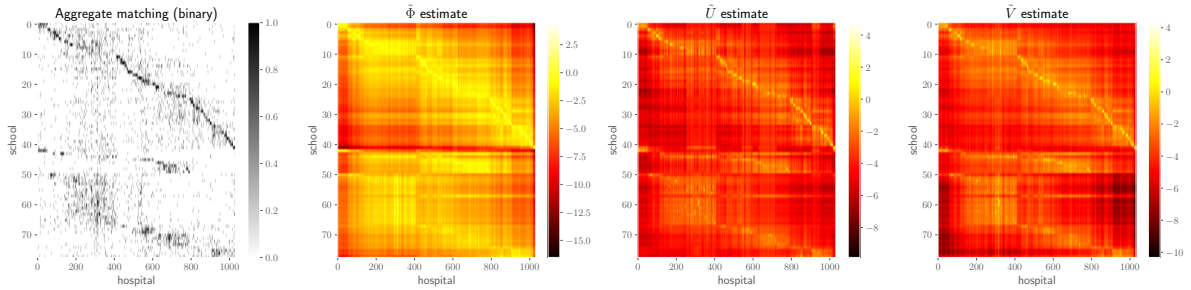
From 2017 to 2019, we estimate the aggregate-level utility of both sides separately. The degree of the polynomial approximating the aggregate-level utility is our tuning parameter. In this section, we show the results obtained when we choose three as the degree of polynomials, as this choice allows a more flexible functional form. In Appendix D.3, we show the results obtained when the degree of the polynomials is set to two.

Each panel in Figure 2 shows the observed matching pattern (leftmost figure), the estimated aggregate-level social surpluses (second from the left), the estimated aggregate-level utilities from the doctors' side (second from the right), and the estimated aggregate-level utilities from the hospitals' side (rightmost figure). All figures are heatmaps, where the vertical axis represents the indices of medical schools and the horizontal axis represents the indices of hospitals.[19] In the right three heatmaps, brighter colors indicate higher values.

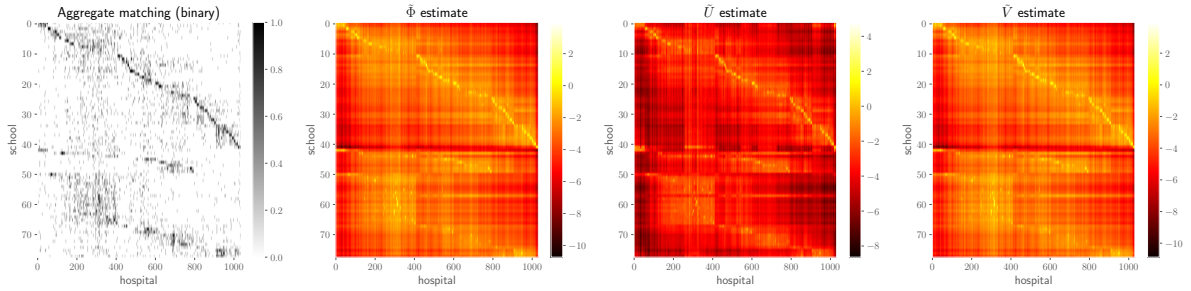
The visible pattern in the aggregate matching is well captured by our estimation. Specifically, all the heatmaps reflect the likelihood of matches between graduates from local public universities and nearby hospitals. As illustrated in the rightmost figures, even from the hospitals' perspective, graduates from closer medical schools provide higher aggregate-level utility. [20] This pattern can be interpreted as evidence of the importance of local knowledge: knowledge of the local medical environment is so critical that hospitals

[19]The way to set the index is described in Section 2.

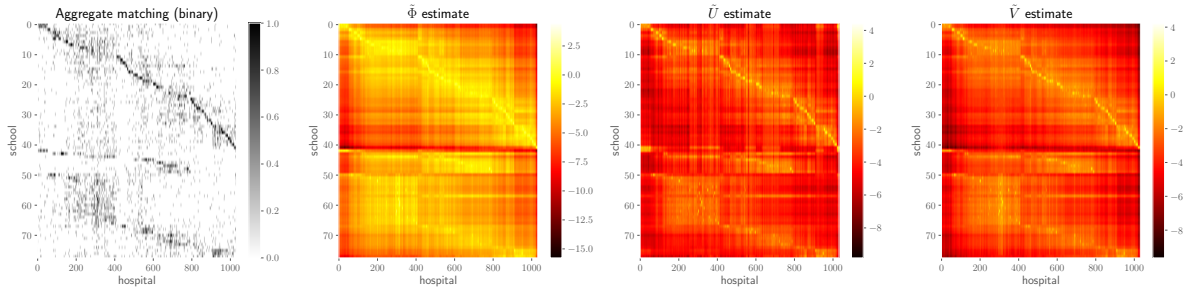
[20]In contrast, Agarwal (2015) highlights that distance influences doctors' decisions but does not account for preference heterogeneity concerning distance on the hospital side.



(a) 2017



(b) 2018



(c) 2019

Figure 2. Aggregate Matchings, Aggregate-level Utilities and Social Surpluses.

prefer to hire local doctors.

6.2 Second Step

Based on the estimation results from the first step, we calculate the marginal effects of covariates on the base utilities. We include the following hospital-specific variables as the characteristics in the preference of doctor side: *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*, *dummy variable of urban area*, and *dummy variable of Tokyo*.[21] Furthermore, we include the following pairwise variables: *logarithm of distance*, *logarithm of number of previous matches*, and *dummy variable of affiliation relationship*. As the characteristics in the preference of hospital side, in addition to the pairwise variables, we include the following university-specific variables: *dummy variable of public university*, *T-score of the entrance exam*, *dummy variable of urban areas*, and *dummy variable of Tokyo*.

The estimation results are shown in Table 5. Columns 1 and 3 correspond to the case of OLS. Columns 2 and 4 are the results obtained when we use BLP instruments for salary. The direction of the estimated coefficients of salary is aligned with the expected signs when we use instruments. We adopt the results obtained using IV estimations as our main estimation results. Although we cannot reject the null hypothesis that $|\gamma_{1,U}| = |\gamma_{1,V}|$ at the 5% significance level, we use different coefficients in the subsequent counterfactual analysis rather than assuming these two are equal.[22]

The distance between the university and the hospital negatively influences both the preferences of doctors and hospitals, which is aligned with the estimation results of the first step. Furthermore, as intuitively and anecdotally validated, the previous number of matches has a strong influence on the preferences. The more previous matches lower the hurdle to apply for the doctor side, and the uncertainty about quality is cleared from the hospital's perspective. The quality measure for both sides is also impactful. From the doctor's side, the number of beds in a hospital increases the utility obtained when matching with a hospital. The hospital prefers doctors from a public university, which is more competitive to enter. We find positive coefficients for the Tokyo dummy and the

[21] Tokyo is by far the largest metropolitan area compared to other urban regions and holds a unique status as the capital, which is why we included a dedicated dummy variable for it.

[22] The p-value of the test is 0.22.

Table 5. 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)
log(Beds)	0.744***	0.814***		
	(0.0303)	(0.0353)		
Public university			0.287***	0.289***
			(0.0569)	(0.0610)
T-score			-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 6. 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			√

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

Table 7. 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			√

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

urban dummy on the doctor's side, and a positive coefficient for the urban dummy on the hospital's side. [23]

Next, we evaluate the marginal effects in monetary units. For this purpose, we calculate the fractions of the estimated coefficients of the covariates relative to the coefficients of salary. The fractions and the standard errors for the doctor side are shown in Table 6, and the same ones for the hospital side are shown in Table 7. In both tables, we present results for three specifications that differ in whether they include interaction terms between year dummies and the urban and Tokyo dummy variables. The results remain qualitatively robust across these specifications.

For the doctor side, the estimates tell that the match with a hospital that is 10% far away 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 yen, which is about $319, the hiring

[23]If there is an implicit tax on the prefectures in the urban areas, this coefficient might be underestimated. In Appendix D.1, we exploit the tightening regional caps to check this existence and conclude that the current market does not face such implicit taxation.

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 improves 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 university decrease 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 a positive impact as expected: the premium of graduating from a public university is from 0.160 to 0.188 million yen, which is about from $\$1,022$ to $\$1,201$.

These marginal effects of the covariates are significant even compared to the aggregate-level utilities. In Appendix D.2, we calculate the ratio of the estimated marginal effects of the covariates to the aggregate-level utility. For many covariates, a 10% change in the covariates corresponds to approximately 1–5% of the aggregate-level utility on both sides.

7 Counterfactual Simulations

7.1 Decomposing the Welfare Loss of Cap-based Policy

We use our empirical estimates to conduct counterfactual simulations designed to evaluate the current regulatory framework against two alternative policies. This analysis enables us to quantify the welfare losses associated with the existing cap-based system and to evaluate the potential benefits of a more efficient, subsidy-based approach. By comparing these scenarios, we can decompose the sources of inefficiency and distinguish the welfare cost of the distributional goal itself from the cost of the policy instrument used to achieve it.

The first policy, *Artificial Caps* (AC), simulates the status quo implemented by the Japan Residency Matching Program (JRMP). We compute the aggregate equilibrium under the set of residency positions that were actually available in the JRMP, subject to regional caps.

The second, *No Caps* (NC), represents an unregulated welfare benchmark. This counterfactual assumes all residency positions are made available at each hospital's true ca-

pacity. To simulate this policy, we must infer the true capacity of each hospital, as this is unobservable under the current regulated regime. We define a hospital's true capacity as the maximum number of positions it reported to the JRMP between 2015 and 2023.[24] The welfare difference between the NC and AC scenarios measures the total efficiency loss created by the regional cap system.

The third policy, *Optimal Subsidy* (OS), provides a constrained-optimal benchmark. In this scenario, we remove all artificial caps and instead introduce subsidies designed to maximize social surplus subject to regional constraints. These subsidies are set to achieve a specific distributional goal: ensuring that designated rural prefectures receive at least as many residents as they did under the AC policy.[25] The optimal set of subsidies is then derived by solving the social surplus maximization problem (P) subject to these regional floor constraints.[26]

By construction, the three policies have a clear welfare ranking: the unconstrained NC policy is first-best, the OS policy is second-best, and the AC policy yields the lowest surplus. This framework allows for a clear decomposition of the total inefficiency. The comparison between the OS and AC outcomes isolates the welfare loss attributable to the choice of regulatory instrument (i.e., caps instead of subsidies). The remaining welfare difference, observed between the NC and OS outcomes, quantifies the inherent cost of the distributional goal itself.

Table 8 presents the simulation results for 2017. The findings for other years are qualitatively similar and are available in Appendix D.4. Unless otherwise specified, monetary values are in millions of JPY per month, calculated using the salary coefficients from specification (1) in Tables 6 and 7. For the OS policy, we report bounds for government revenue and total welfare, reflecting different possible incidences of taxes and subsidies between doctors and hospitals.

Comparing the AC and NC scenarios reveals the direct impact of the caps. Removing them increases social surplus by approximately 2.6 billion JPY and raises the match rate by 4.4 percentage points. This welfare gain, however, comes at the expense of three rural

[24]This definition provides a conservative estimate of each hospital's true capacity. As a result, our calculation of the total welfare loss from the cap system should be interpreted as a lower bound.

[25]By setting the floor for the number of residents equal to the outcome of the current policy, we adopt a conservative distributional objective. A more ambitious policy goal with higher floors would necessarily require a larger total subsidy. Thus, our estimate of the fiscal cost should be interpreted as a lower bound.

[26]We define rural prefectures as the 15 prefectures with the lowest physician match rates between 2017 and 2019, calculated as the total matches over the period divided by the sum of reported capacities.

Table 8. Comparison of Three Policies in 2017

Policy	AC (Artificial Caps)	NC (No Caps)	OS (Optimal Subsidy)
Artificial caps	Yes	No	No
Floor constraints	Yes	No	Yes
Subsidies	No	No	Yes
Match rate	0.868	0.912	0.912
Doctors' welfare	32795.8	33441.7	33444.5
Hospitals' welfare	29586.6	31578.1	31579.4
Government's revenue	0.0	0.0	$[-10.5, -7.4]$
Social welfare	62382.3	65019.8	$[65013.3, 65016.5]$
#(subsidized regions)	0	0	3
Average subsidy	0.000	0.000	-0.040
#(constraint violations)	0	3	0

* All welfare and revenue figures are expressed in units of 1 million JPY per month. Government revenue is positive when taxes are imposed on doctors and hospitals and negative when subsidies are provided. Doctors' and hospitals' welfare are scaled according to specification (1) in Table 6 and Table 7. We present the bounds of the government's net revenue, scaled by the coefficients on the doctor side and the hospital side, respectively. The social 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).

prefectures receiving fewer residents, confirming that the caps are binding and effective at redirecting physicians, albeit inefficiently.

The OS policy demonstrates that these distributional goals can be met with minimal welfare loss. The surplus under OS is substantially higher than under AC and is nearly identical to the unconstrained NC benchmark. This finding suggests that the vast majority of the welfare loss from the current policy originates from the choice of regulatory instrument—the caps themselves—rather than from the underlying distributional constraint.

Figure 3 illustrates the prefectural-level welfare differences between the OS and AC policies. The gains from adopting subsidies are positive for all prefectures and are largest in urban areas that face the tightest caps under the JRMP. This reinforces the conclusion that the inefficiency originates primarily from restricting matches in high-demand urban centers.

Finally, the OS policy appears fiscally practical. The required subsidy to achieve the distributional floor is approximately \$400 per month for each matched pair in the targeted rural prefectures, which represents 10–20% of a typical resident's salary. The total national cost of this subsidy program would not exceed \$100,000 per month, suggesting

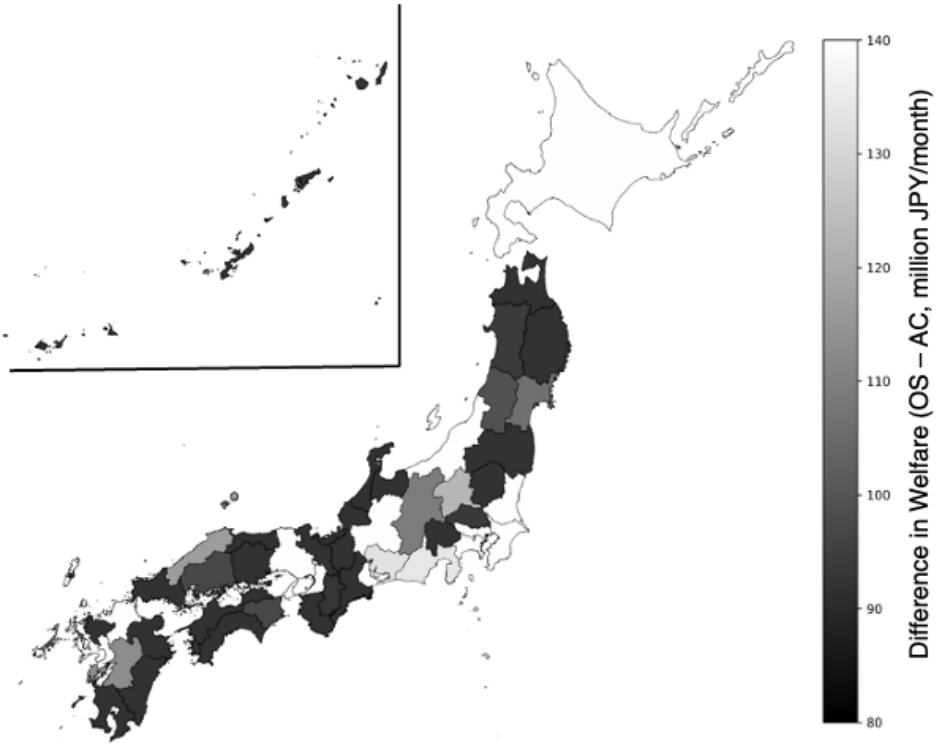


Figure 3. Difference in Welfare between OS and AC by Prefecture

that a shift to a more efficient subsidy-based policy is feasible.

Changes in matching patterns We compare the equilibrium matching patterns between AC, NC, and OS. Specifically, we replicate the regression analysis from Section 2.2 for every scenario and statistically test whether the coefficients on each covariate differ significantly. For these tests, we compute standard deviations of the estimates using cluster-robust standard errors at the unit level: the hospital level for the hospital perspective and the school level for the school perspective. The results from the hospital perspective are presented in Table 9, and those from the hospital perspective are shown in Table 10. In both tables, Panel A compares NC, where regional caps in urban areas are removed, with AC, which represents the current market equilibrium. Panel B compares OS, which allows subsidies to meet floor constraints in rural areas, with NC.

We begin by examining Panel A, which captures the changes in matching patterns that occur when the regional caps in urban areas are removed. According to Table 9, the newly available slots in urban hospitals are filled primarily by graduates from private universities, particularly those with lower entrance difficulty and located in more distant regions. While the unmatch rate in urban areas is already low, it decreases even further

Table 9. Coefficient Size Tests for Hospital Side

Panel A: NC vs AC	Public	Difficulty	Distance	Unmatch Rate
Urban	$<^{***}$	$<^{***}$	$>^{***}$	$<^{***}$
Panel B: OS vs NC	Public	Difficulty	Distance	Unmatch Rate
Urban	$>^{***}$	$>^{**}$	$<^{***}$	$>^{***}$

* Each cell compares coefficients across the two regimes indicated in the panel header. "$>$" (resp. "$<$") means the coefficient is larger (resp. smaller) under the first regime. Significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$ (two-sided test).

Table 10. Coefficient Size Tests for School Side

Panel A: NC vs AC	Urban Hosp.	Distance	Unmatch Rate
Public University	$<$	$<$	$<$
Difficulty	$<$	$<$	$<$
Panel B: OS vs NC	Urban Hosp.	Distance	Unmatch Rate
Public University	$>$	$<$	$<$
Difficulty	$>^{*}$	$<$	$>$

* Each cell compares coefficients across the two regimes indicated in the panel header. "$>$" (resp. "$<$") means the coefficient is larger (resp. smaller) under the first regime. Significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$ (two-sided test).

under the NC policy. From the school perspective shown in Table 10, although the differences are not statistically significant, the first row displays a pattern consistent with this observation. The negative correlation between distance and both public status and difficulty also suggests that public universities, especially those located in rural areas, have become even more prominent sources of residents for rural hospitals under this scenario.

Next, we turn to Panel B, which illustrates changes in matching patterns resulting from the introduction of subsidies in rural areas. As shown in Table 9, the changes in matching patterns from NC to OS (Panel B) are the exact opposite of those from AC to NC (Panel A) across all variables. This suggests that the subsidy effectively counteracts the influx of students into urban areas resulting from the removal of regional caps, encouraging them to remain in rural areas instead. The results from the school perspective in Table 10 are also consistent with this interpretation.

These changes in matching patterns help explain why the optimal subsidy required in our analysis is relatively small. The subsidy's primary role is to retain students from rural

universities who would otherwise move to urban hospitals after the removal of regional caps. Because distance is a significant factor in student preferences, matching with a local hospital is not a prohibitively costly alternative for these students. Therefore, a modest subsidy is sufficient to persuade enough students to choose a local hospital, thereby meeting the distributional goal at a low fiscal cost.

7.2 Inherent Inefficiency of Caps: An Evaluation of the Flexible DA Algorithm

To further assess the limitations of cap-based regulations, we conduct a second set of counterfactual simulations evaluating a state-of-the-art matching algorithm from the literature. The *Flexible Deferred Acceptance (FDA) algorithm*, a variant of the standard DA algorithm, adjusts hospital capacities to meet regional constraints dynamically, rather than relying on the pre-determined, fixed capacity reductions imposed by the current JRMP protocol. To distinguish whether the inefficiency of the current policy stems from a naive implementation or an inherent limitation of caps, this exercise compares the outcomes of the FDA algorithm with the three policies considered previously: Artificial Caps (AC), No Caps (NC), and Optimal Subsidy (OS).

This simulation exercise more closely mimics the institutional details of the JRMP. We generate individual-level preferences from which agents construct preference rankings, and then use a centralized algorithm to determine the final matches. Unlike the previous analysis, this approach computes individual-level outcomes without relying on the large-market approximation, and it determines matches using the DA algorithm rather than as a transferable-utility stable outcome. This allows us to check the robustness of our baseline findings from Section 7.1.

For the AC policy, we first estimate the aggregate-level utilities by solving the optimization problem (P) under the actual JRMP caps. We then generate individual doctor and hospital utilities by adding a random logit error to these aggregate utilities. Based on the resulting individual preferences, we construct preference lists and run the standard DA algorithm to determine the final matching under the JRMP caps. The NC policy simulation is analogous, except that we use each hospital's true capacity to compute the aggregate-level utilities.

For the OS policy, we set floor constraints requiring each designated rural prefecture

Table 11. Welfare Comparison of the Simulated Matchings in 2017

Scenario	AC	NC	OS	FDA
Algorithm	DA	DA	DA	FDA
Artificial Caps	Yes	No	No	No
Floor Constraints	Yes	No	Yes	Yes
Subsidies	No	No	Yes	No
Match rate	0.762	0.810	0.810	0.771
Doctors' welfare	35850.2	37176.3	37192.4	35979.4
Hospitals' welfare	29748.7	29174.1	29176.6	29643.0
Government's revenue	0.0	0.0	$[-17.0, -12.0]$	0.0
Social welfare	65598.9	66350.5	$[66352.1, 66357.1]$	65622.4
#regions violating lower bounds	0	4	3	0
#doctors required to meet lower bounds	0	15	12	0
#doctors matched with urban hospitals	3103	3196	3192	3016
#doctors matched with rural hospitals	1071	1170	1175	1209
Urban hospitals' welfare	6143.2	5758.4	5753.2	5954.1
Rural hospitals' welfare	2281.8	2227.0	2243.8	2447.8

* All welfare and revenue figures are expressed in units of 1 million JPY per month. Government revenue is positive when taxes are imposed on doctors and hospitals, and negative when subsidies are provided. Doctors' and hospitals' welfare are scaled according to specification (1) in Table 6 and Table 7. We present the bounds of the government's net revenue, scaled by the coefficients on the doctor side and the hospital side, respectively. The social welfare is the sum of doctors' welfare, hospitals' welfare, and the government's revenue. "#regions violating lower bounds" indicates the number of prefectures whose matched doctor count is less than that of AC. "#doctors required to meet lower bounds" indicates how many additional doctors must be matched with rural hospitals so that these prefectures exceed the matched doctor count in AC.

to receive at least as many residents as it did under the AC policy. We then solve the optimization problem (P) to compute the optimal subsidy and the corresponding aggregate-level utilities. Individual preferences are generated from these utilities, and the standard DA algorithm determines the match outcome.

Finally, for the FDA policy, individual utilities are identical to those in the AC scenario.[27] The matching is then determined by the FDA algorithm, which operates under regional caps calibrated to satisfy the floor constraints as in the OS scenario. The floor constraints requires each rural prefecture to receive at least as many residents as the number of matches under DA in the AC scenario.[28]

[27]Since hospital capacities are computed during the execution of the FDA algorithm, no definitive method exists for calculating the aggregate-level utilities prior to running the algorithm. We therefore use the utilities from the AC scenario.

[28]Specifically, a regional cap is initially set for each non-rural prefecture equal to the sum of the true capacities of the hospitals within it. We then find the smallest α such that if the regional caps of all prefectures are reduced uniformly by α%, the final matching produced by the FDA algorithm satisfies the floor constraints.

Table 11 reports the simulation results. Under the FDA policy, capacities in non-rural areas are reduced by 30% to achieve the floor constraints. Despite this substantial reduction, the FDA algorithm improves upon the current AC policy, increasing the total match rate by 0.9 percentage points and yielding higher total welfare. However, the welfare under FDA remains significantly lower than that under the unconstrained NC benchmark. In contrast, the OS policy achieves a total welfare level nearly identical to that of the NC policy. This result strengthens our earlier conclusion that a fundamental inefficiency is inherent in cap-based policies, at least in this application, and that monetary interventions offer a powerful alternative.

Furthermore, the results for the AC, NC, and OS policies are consistent with the findings in Section 7.1. Social welfare increases when the artificial caps are removed (NC vs. AC), and the OS policy nearly restores the first-best welfare level. This consistency across different simulation methodologies validates our baseline approach and reinforces the central findings of our analysis.

8 Conclusion

This study develops and implements a framework to evaluate the efficiency of regulations in matching markets with distributional disparities, focusing on the trade-offs between cap-based policies and monetary interventions. We build a transferable utility (TU) matching model that incorporates regional constraints, such as caps and floors, on the number of matches. By embedding this model within the aggregate matching framework of Galichon and Salanié (2021a), we provide a method to identify preference structures and conduct policy simulations using only aggregate-level matching data and an observable measure of transfers, such as salary.

Applying this framework to the Japan Residency Matching Program (JRMP), our empirical analysis reveals several key findings. First, the estimation results show that preferences are horizontally differentiated on both sides of the market; factors such as geographic distance and the history of previous matches are significant determinants of preferences for both doctors and hospitals. Second, our counterfactual simulations demonstrate that the current cap-based regulations, while redirecting some physicians to underserved areas, generate substantial welfare losses. We find that a small, targeted subsidy policy can achieve the same distributional outcome as the existing caps but at a

significantly higher level of social welfare. This indicates that the inefficiency stems not from the distributional goal itself, but from the choice of regulatory instrument. Finally, we demonstrate that even a more sophisticated, flexible Deferred Acceptance mechanism still yields considerable welfare losses compared to the optimal subsidy, underscoring the inherent limitations of employing quantity restrictions in this context.

The findings offer important policy implications. For markets like the JRMP, where geographic imbalances are a primary concern, monetary interventions appear to be a more efficient and fiscally practical tool than rigid caps. Our framework offers a practical path to designing such interventions, as its reliance on aggregate-level matching data and salaries is less demanding than approaches that require individual-level preference data. More broadly, the approach developed in this paper can be applied to other matching markets where transfers are present but granular data are scarce. Future research could extend this approach to analyze more complex distributional constraints or explore the long-run effects of such policies on physician career paths and settlement patterns.

References

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 2003. "School Choice: A Mechanism Design Approach." *American Economic Review*, 93(3): 729–747.

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." *American Economic Review*, 107(8): 2204–2242.

Choo, Eugene, and Aloysius Siow. 2006. "Who Marries Whom and Why." *Journal of Political Economy*, 114(1): 175–201.

Ehlers, Lars, Isa E. Hafalir, M. Bumin Yenmez, and Muhammed A. Yildirim. 2014. "School Choice with Controlled Choice Constraints: Hard Bounds versus Soft Bounds." *Journal of Economic Theory*, 153: 648–683.

Ellison, Glenn, and Parag A Pathak. 2021. "The efficiency of race-neutral alternatives to race-based affirmative action: Evidence from Chicago's exam schools." *American Economic Review*, 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].

Fox, Jeremy T. 2010. "Identification in matching games." *Quantitative Economics*, 1(2): 203–254.

Fox, Jeremy T. 2017. "Specifying a Structural Matching Game of Trading Networks with Transferable Utility." *American Economic Review*, 107(5): 256–60.

Fox, Jeremy T. 2018. "Estimating matching games with transfers." *Quantitative Economics*, 9(1): 1–38.

Fox, Jeremy T., and Patrick Bajari. 2013. "Measuring the Efficiency of an FCC Spectrum Auction." *American Economic Journal: Microeconomics*, 5(1): 100–146.

Fox, Jeremy T., Chenyu Yang, and David H. Hsu. 2018. "Unobserved Heterogeneity in Matching Games." *Journal of Political Economy*, 126(4): 1339–1373.

Fragiadakis, Daniel, and Peter Troyan. 2017. "Improving matching under hard distributional constraints: Improving matching under constraints." *Theor. Econ.*, 12(2): 863–908.

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*.

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

Galichon, Alfred, Scott Duke Kominers, and Simon Weber. 2019. "Costly Concessions: An Empirical Framework for Matching with Imperfectly Transferable Utility." *Journal of Political Economy*, 127(6): 2875–2925.

Ghouila-Houri, Alain. 1962. "Caractérisation des matrices totalement unimodulaires." *Comptes Redus Hebdomadaires des Séances de l'Académie des Sciences (Paris)*, 254: 1192–1194.

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

Hoffman, Alan J, and Joseph B Kruskal. 2010. "Integral boundary points of convex polyhedra." *50 Years of integer programming 1958–2008*, 49.

Iizuka, Toshiaki, and Yasutora Watanabe. 2016. "The impact of physician supply on the healthcare system: Evidence from japan's new residency program." *Health Econ.*, 25(11): 1433–1447.

Jalota, Devansh, Michael Ostrovsky, and Marco Pavone. 2025. "Matching with transfers under distributional constraints." *Games Econ. Behav.*, 152: 313–332.

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

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

Kojima, Fuhito, Ning Sun, and Ning Neil Yu. 2020. "Job matching under constraints." *Am. Econ. Rev.*, 110(9): 2935–2947.

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

Sakai, Rie, Wei Wang, Norihiro Yamaguchi, Hiroshi Tamura, Rei Goto, and Ichiro Kawachi. 2013. "The impact of Japan's 2004 postgraduate training program on intra-prefectural distribution of pediatricians in Japan." *PLoS One*, 8(10): e77045.

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

Yokote, Koji. 2020. "On optimal taxes and subsidies: a discrete saddle-point theorem with application to job matching under constraints." *Journal of Mechanism and Institution Design*, 5(1): 37–77.

A Omitted Proofs

A.1 Supporting Stable Outcomes as an Equilibrium

In this section, we show that any stable outcome (Definition 1) can be supported as an equilibrium of a stylized game that mirrors the JRMP's matching process. For expository simplicity, we assume that taxation policy is inactive, i.e., $w_z \equiv 0$.[29]

The game consists of three periods. In Period 1, each slot simultaneously announces a profile of transfers $t = (t_{ij})_{i,j}$, where $t_j = (t_{ij})_i$ is the transfer profile chosen by slot j. In Period 2, observing all the wage offers, doctors and slots submit their preference lists. In Period 3, the doctor-proposing deferred acceptance (DA) algorithm determines the final matching. The utility of doctor i being matched with a slot j is defined as follows: the *base utility*, which is the utility felt by i net of transfer, is denoted by U_{ij}. Given transfer offer t, when matched with slot j, doctor i's payoff is $U_{ij} + t_{ij}$. Similarly, the base utility of slot j when matched with doctor i is denoted by V_{ij}, and the payoff of j when matched with i given w is $V_{ij} - t_{ij}$. Note that $\Phi_{ij} = U_{ij} + V_{ij}$. We refer to this game as the *JRMP game*.

We assume that all agents submit the preference ranking truthfully given t in Period 2.[30] Given this assumption, the analysis of this game boils down to that of one-shot game played by the slots in Period 1.

The following proposition establishes the link between the stable outcome and the equilibrium of this game.

Proposition A.1. *For any stable outcome (d^*, u^*, v^*) of the matching market (I, J, Φ), there exists a Nash equilibrium of the JRMP game that results in the matching d^* and payoffs (u^*, v^*).*

Proof. Fix any slot j. Consider the following transfer offer profile $(t^*_{ij})_i$:

$$t^*_{ij} \coloneqq \begin{cases} u^*_i - U_{ij} & (d^*_{ij} = 1) \\ M_{ij} & (\text{o.w.}) \end{cases},$$

[29] If not, we can consider Φ_{ij} below as the gross joint surplus for any given w.

[30] Since the DA algorithm is strategy-proof for the proposing side, this means that we assume the truthful report from the other side. This is a common assumption when studying DA algorithms.

where $M_{ij} \in \mathbb{R}$ is sufficiently low for slot j to be unattractive to doctor i, i.e., $U_{ij} + M_{ij} < \min\{u_i^*, U_{i0}\}$ for each i.

Suppose that this transfer profile t^* is chosen in Period 1. If doctor i is matched with some slot j under the stable outcome (i.e, $d_{ij}^* = 1$), then the preference list submitted by doctor i is[31]

$$i\colon\ j \succ \emptyset \succ \cdots .$$

Since $d_{ij}^* = 1$, in the slot j's preference list, doctor i is placed above the outside option. Therefore, doctor i is not rejected in the first round of the doctor-proposing DA algorithm. If doctor i is unmatched under the stable outcome, he lists the outside option at the top of his preference list, and proposes to no slot in all rounds. Since no rejection occurs, the algorithm terminates in the first round, yielding the matching d^*. By construction of t^*, each agent enjoys the same payoff as in the stable outcome.

We now verify that no hospital has a profitable unilateral deviation in Period 1. Suppose toward contradiction that there exists slot j that can be strictly better off by matching with i' by setting $(\tilde{t}_{ij})_i$, where $d_{i'j}^* = 0$. For this to be a successful and profitable deviation, two conditions must be met:

$$v_j^* < V_{i'j} - \tilde{t}_{i'j},$$
$$u_{i'}^* \leq U_{i'j} + \tilde{t}_{i'j}.$$

This implies that

$$v_j^* + u_{i'}^* < \Phi_{i'j},$$

which violates the stability of (d^*, u^*, v^*). A contradiction. □

A.2 Results for Section 4.1

Below, without loss of generality, we normalize the outside option values to zero, i.e., $\Phi_{i,j_0} = \Phi_{i_0,j} \equiv 0$ for all $i \in I$ and $j \in J$.[32] The goal of this subsection is to prove the following result:

[31] If there is a tie between slot j and the outside option, the tie-breaking rule does not matter because it won't affect the equilibrium payoff. For simplicity, I assume that doctor i always places slot j at the top.

[32] This is possible because the optimal solution to (P_0) remains the same when we redefine Φ_{ij} as $\Phi_{ij} - \Phi_{i,j_0} - \Phi_{i_0,j}$ and ignore the terms in the objective function that are independent of d.

Theorem A.1. *Fix any market (I, J, Z, z, Φ) and regional constraints $(\underline{o}_z, \bar{o}_z)_z$. There exists taxation policy w^* such that the matching under the stable outcome under w^* is d^*.*

We first define an auxiliary problem in which we relax the integral constraint with respect to d so that d_{ij} can take any nonnegative real value:

$$(\mathrm{P}'_0) \begin{cases} \max\limits_{d \geq 0} & \sum\limits_{i,j} \Phi_{ij} d_{ij} & \\ \text{s.t.} & \sum_j d_{ij} \leq 1 & (i \in I) \\ & \sum_i d_{ij} \leq 1 & (j \in J) \\ & \sum_{j \in z} \sum_i d_{ij} \leq \bar{o}_z & (z \in Z) \\ & \sum_{j \in z} \sum_i d_{ij} \geq \underline{o}_z & (z \in Z) \end{cases}$$

For (P'_0) to be useful, it must yield a matching, or an integer solution. This is known to be the case without regional constraints. The following proposition states that it remains true even with regional constraints.

Proposition A.2. *Fix any matching market (I, J, Z, z, Φ) and regional constraints $(\underline{o}_z, \bar{o}_z)_z$. Assume that (P'_0) has at least one feasible solution. Then, (P'_0) has an integer optimal solution.*

Proof of Proposition A.2

Our proof utilizes the known results regarding total unimodularity:

Definition 2 (Total unimodularity)**.** *Let A be an integer matrix. A is* totally unimodular *if any minor principal is either -1, 0, or 1.*

First, we show the following lemma:

Lemma A.1. *The matrix that represents the set of constraints of (P'_0) is totally unimodular.*

Our proof of Lemma A.1 relies on the following fact about the total unimodularity.

Lemma A.2 (Ghouila-Houri (1962))**.** *An $m \times n$ integer matrix A is totally unimodular iff for each subset of rows $R \subseteq [m]$, there exists a partition R_1 and R_2 of R such that*

$$\forall k \in [n], \quad \sum_{i \in R_1} a_{ik} - \sum_{i \in R_2} a_{ik} \in \{-1, 0, 1\}$$

Proof of Lemma A.1. First, observe that the feasibility constraints (i.e., all the constraints except $x \geq 0$) of an instance of TU matching with regional constraints can be represented by $(|I|+|J|+2|Z|) \times (|I| \times |J|)$ matrix A (see also Example 2) such that

- Each row corresponds to either (1) agent $i \in I$, (2) agent $j \in J$, (3) an upper bound $\bar{o}_z$ for region $z \in Z$, or (4) a lower bound $\underline{o}_z$ for region $z \in Z$.
- Each column corresponds to an $(i, j) \in I \times J$ pair.
- The component of row $i \in I$ is 1 for column (i, j') for any $j' \in J$; the component is 0 otherwise.
- The component of row $j \in J$ is 1 for column (i', j) for any $i' \in I$; the component is 0 otherwise.
- Each region z has two corresponding rows: one is for upper bound $\bar{z}$ and another one is for lower bound $\underline{z}$.
  - The component of row $\bar{z}$ is 1 for column (i, j) such that $z(j) = z$; the component is 0 otherwise.
  - The component of row $\underline{z}$ is -1 for column (i, j) such that $z(j) = z$; the component is 0 otherwise.

We apply Lemma A.2 to prove that A is totally unimodular. Fix any subsets of rows R. Let $I_R \subseteq I$ be the rows corresponding to $i \in I$ contained in R. J_R, $\bar{Z}_R$, and $\underline{Z}_R$ are defined analogously.

We classify the rows in R by the following algorithm:

1. Let $R_1 \coloneqq \bar{Z}_R$ and $R_2 \coloneqq \emptyset$.
2. For each $\underline{z} \in \underline{Z}_R$, if $\bar{z} \in \bar{Z}_R$, then update $R_1 \leftarrow R_1 \cup \{\underline{z}\}$; otherwise, update $R_2 \leftarrow R_2 \cup \{\underline{z}\}$.
3. Let $J_R \coloneqq \{j(1), \ldots, j(T_J)\}$, where $T_J \coloneqq |J_R|$, $R_1(1) \coloneqq R_1$, and $R_2(1) \coloneqq R_2$. For each $t \in [T_J]$,
   (a) Define a row vector

$$r(t) \coloneqq \left(\sum_{i \in R_1(t)} a_{ik} - \sum_{i \in R_2(t)} a_{ik} \right)_k .$$

(b) If $r(1)_{(i,j(t))} = 1$ for some i, then $R_1(t+1) \coloneqq R_1(t)$ and $R_2(t+1) \coloneqq R_2(t) \cup \{j(t)\}$.

(c) Otherwise, $R_1(t+1) \coloneqq R_1(t) \cup \{j(t)\}$ and $R_2(t+1) \coloneqq R_2(t)$.

4. Let $R_1 \coloneqq R_1(T_J)$ and $R_2 \coloneqq R_2(T_J) \cup I_R$. Return R_1 and R_2.

We will show why the algorithm above works. First, under the hierarchical regional constraints, each component of $r(1)$ is either 0 or 1. By the construction of Step 3, all the components of $r(T_J)$ are either 0 or 1. Let

$$r \coloneqq \left(\sum_{i \in R_1} a_{ik} - \sum_{i \in R_2} a_{ik} \right)_k ,$$

where R_1 and R_2 are the output of Step 4. By construction, for any column k, $\sum_{i \in I_R} a_{ik} \in \{0, -1\}$. Therefore, we have

$$\forall k \in [n], \quad \sum_{i \in R_1} a_{ik} - \sum_{i \in R_2} a_{ik} \in \{-1, 0, 1\}.$$

By Lemma A.2, this implies that A is totally unimodular. □

By the following well-known result, we can conclude that (P'_0) has an integer optimal solution:

Lemma A.3 (Hoffman and Kruskal (2010))**.** *A is totally unimodular iff, for any $b \in \mathbb{Z}^m$, $P \coloneqq \{x \in \mathbb{R}^n \colon Ax \leq b, x \geq 0\}$ is an integral polyhedra, i.e., all the faces includes an integer vector. If P is bounded, this is equivalent to that the components of all vertices of P are integers.*

This completes the proof of Proposition A.2. □

Example 2 (TU matching with regional constraints)**.** Let $I \coloneqq \{i_1, i_2\}$, $J \coloneqq \{j_1, j_2, j_3\}$, $Z \coloneqq \{z_1, z_2\}$, $z(j_1) = z(j_2) = z_1$, and $z(j_3) = z_2$. The set of constraints can be written

as $A_0 x \leq b$ by defining A_0, x, and b as follows:

$$
A_0 \coloneqq \begin{array}{c} \\ i_1 \\ i_2 \\ j_1 \\ j_2 \\ j_3 \\ \bar{z}_1 \\ \bar{z}_2 \\ \underline{z}_1 \\ \underline{z}_2 \\ (i_1,j_1) \\ (i_1,j_2) \\ (i_1,j_3) \\ (i_2,j_1) \\ (i_2,j_2) \\ (i_2,j_3) \end{array}
\begin{array}{c}
\begin{array}{cccccc} (i_1,j_1) & (i_1,j_2) & (i_1,j_3) & (i_2,j_1) & (i_2,j_2) & (i_2,j_3) \end{array} \\
\left[\begin{array}{cccccc}
1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 \\
1 & 1 & 0 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 \\
-1 & -1 & 0 & -1 & -1 & 0 \\
0 & 0 & -1 & 0 & 0 & -1 \\
-1 & & & & & \\
 & -1 & & & & \\
 & & -1 & & & \\
 & & & -1 & & \\
 & & & & -1 & \\
 & & & & & -1
\end{array}\right]
\end{array}
$$

$$x \coloneqq (x_{i_1 j_1}, x_{i_1 j_2}, x_{i_1 j_3}; x_{i_2 j_1}, x_{i_2 j_2}, x_{i_2 j_3})^\top$$

$$b \coloneqq (1, 1, 1, 1, 1, \bar{o}_{z_1}, \bar{o}_{z_2}, \underline{o}_{z_1}, \underline{o}_{z_2}, 0, 0, 0, 0, 0, 0)^\top$$

The last $|I| \times |J|$ rows corresponds to non-negativity constraints $d_{ij} \geq 0$. Note that $[B\ I]$ is totally unimodular if B is totally unimodular. Thus, to show A_0 is totally unimodular,

it suffices to show that

$$A \coloneqq \begin{array}{c} \\ i_1 \\ i_2 \\ j_1 \\ j_2 \\ j_3 \\ \bar{z}_1 \\ \bar{z}_2 \\ \underline{z}_1 \\ \underline{z}_2 \end{array} \begin{array}{c} \begin{array}{cccccc} (i_1,j_1) & (i_1,j_2) & (i_1,j_3) & (i_2,j_1) & (i_2,j_2) & (i_2,j_3) \end{array} \\ \left[\begin{array}{cccccc} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ -1 & -1 & 0 & -1 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 \end{array}\right] \end{array}$$

is totally unimodular.

Proof of Theorem A.1

Consider the following dual problem of (P_0'):

$$(\mathrm{D}_0') \quad \left[\begin{array}{ll} \max\limits_{u,v,\bar{w},\underline{w}\geq 0} & \sum\limits_i u_i + \sum\limits_j v_j + \sum\limits_{z\in Z} \bar{o}_z \bar{w}_z - \sum\limits_{z\in Z_L} \underline{o}_z \underline{w}_z \\ \text{s.t.} & u_i + v_j \geq \Phi_{ij} - \bar{w}_{z(j)} + \underline{w}_{z(j)} \qquad (i\in I, j\in J) \end{array}\right.$$

Let d^* be an integer optimal solution to $(\mathrm{P}_0{}')$ (NB: this d^* is also the optimal solution to (P_0)) and let $(u, v, \bar{w}, \underline{w})$ be a solution to (D_0'). By a similar argument as in the standatd TU matching model that characterizes an stable outcome as a solution to the social welfare maximization problem and its dual problem, we can show that $(d^*, (u, v))$ is a stable outcome given Φ and w^*, where $w_z^* \coloneqq \bar{w}_z \mathbb{1}\{\bar{w}_z > 0\} - \underline{w}_z \mathbb{1}\{\underline{w}_z > 0\}$. Note that $\bar{w}_z > 0$ and $\underline{w}_z > 0$ cannot happen simultaneously due to the complementary slackness condition.

Corollary A.1. *Fix any cap-based policy characterized by $J' \subseteq J$ such that the resulting matching satisfies the regional constraints. The social welfare achieved under this policy is weakly less than the welfare attained under the optimal taxation policy w^*.*

Proof. Fix any J'. The social welfare with J' cannot be better than the optimal value of

(P_0)) with additional constraints $d_{ij} = 0$ for any $i \in I$ and $j \in J'$. □

A.3 Proof of Lemma 1

First, we will show the following lemma: [33]

Lemma A.4. *Suppose that U and V are aggregate-level utilities given (u, v) (i.e., (1) 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

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

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 $y \in Y$.

[33] The following proof of Lemma A.4 is almost identical to the proof of Proposition 1 of Galichon and Salanié (2021*a*).

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 (1), 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

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

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

We want to claim that $u_i \leq \max_{y \in Y_0} \{U_{xy} + \varepsilon_{iy}\}$. 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

$$\begin{aligned} \Phi_{ij} - w_{z(y)} &= u_i + v_j \\ &> \left(\max_{y' \in Y_0} \ U_{xy'} + \varepsilon_{iy'} \right) + \left(\max_{x' \in X_0} \ V_{x'y} + \eta_{x'j} \right) \\ &\geq U_{xy(j)} + \varepsilon_{iy(j)} + V_{xy(j)} + \eta_{xj} \\ &\geq \Phi_{xy} - w_{z(y)} + \varepsilon_{iy(j)} + \eta_{xj} \quad (\because \text{ Lemma A.4}) \\ &= \Phi_{ij} - w_{z(y)}. \end{aligned}$$

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

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

A contradiction. Therefore, we have $u_i \leq \max_{y \in Y_0} \{U_{xy} + \varepsilon_{iy}\}$ 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.4 Proof of Theorem 1

First, we show the strict convexity of G and H.

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

Proof. G is strictly increasing. Take any $U^1, U^2 \in \mathbb{R}^{N\times M}$ such that $U^1 \geq U^2$ and $U^1 \neq U^2$. Then, $U^1_{xy} > U^2_{xy}$ holds for some $x \in X$ and $y \in Y$. Define an event A as follows:

$$A \coloneqq \left\{\varepsilon_i \colon \forall y' \neq y,\ U^1_{xy} + \varepsilon_{iy} > U^1_{xy'} + \varepsilon_{iy'},\ \text{AND } U^2_{xy} + \varepsilon_{iy} > U^2_{xy'} + \varepsilon_{iy'}\right\}.$$

Let $G_x(U) \coloneqq \mathbb{E}_{\varepsilon_i \sim P_x}[\max_y\{U_{xy} + \varepsilon_{iy}\}]$. Note that $G(U) = \sum_x n_x G_x(U)$.

Since P_x has full support, we have $\Pr(A) > 0$. Then,

$$\begin{aligned}
G_x(U^1) &= \mathbb{E}\left[\max_y \left\{U^1_{xy} + \varepsilon_{iy}\right\}\right] \\
&= \mathbb{E}\left[\max_y \left\{U^1_{xy} + \varepsilon_{iy}\right\} \mid A\right] \Pr(A) + \mathbb{E}\left[\max_y \left\{U^1_{xy} + \varepsilon_{iy}\right\} \mid A^c\right] \Pr(A^c) \\
&= \mathbb{E}\left[U^1_{xy} + \varepsilon_{iy} \mid A\right] \Pr(A) + \mathbb{E}\left[\max_y \left\{U^1_{xy} + \varepsilon_{iy}\right\} \mid A^c\right] \Pr(A^c) \\
&> \mathbb{E}\left[U^2_{xy} + \varepsilon_{iy} \mid A\right] \Pr(A) + \mathbb{E}\left[\max_y \left\{U^2_{xy} + \varepsilon_{iy}\right\} \mid A^c\right] \Pr(A^c) \\
&\geq \mathbb{E}\left[\max_y \left\{U^2_{xy} + \varepsilon_{iy}\right\} \mid A\right] \Pr(A) + \mathbb{E}\left[\max_y \left\{U^2_{xy} + \varepsilon_{iy}\right\} \mid A^c\right] \Pr(A^c) \\
&= G_x(U^2).
\end{aligned}$$

Since $U^1 \geq U^2$, we also have $G_{x'}(U^1) \geq G_{x'}(U^2)$ for all $x' \neq x$. Therefore, we have

$$G(U^1) = \sum_x n_x G_x(U^1) > \sum_x n_x G_x(U^2) = G(U^2).$$

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

$$\begin{aligned}
&sG(U^1) + (1-s)G(U^2) \\
&= \sum_x n_x \mathbb{E}\left[\left(\max_y \ s(U^1_{xy} + \varepsilon_{iy})\right) + \left(\max_y \ (1-s)(U^2_{xy} + \varepsilon_{iy})\right)\right] \qquad (5)\\
&\geq \sum_x n_x \mathbb{E}\left[\max_y \ \left\{sU^1_{xy} + (1-s)U^2_{xy}\right\} + \varepsilon_{iy}\right] \qquad (6)\\
&= G\Big(sU^1 + (1-s)U^2\Big)
\end{aligned}$$

holds, G is a convex function.

Now suppose $U^1 \neq U^2$. Then $U^1_{xy} \neq U^2_{xy}$ holds for some $x \in X$ and $y \in Y$. Without loss of generality, assume $U^1_{xy} > U^2_{xy}$. Define an event B as follows:

$$B \coloneqq \left\{\varepsilon_i \colon U^1_{xy} + \varepsilon_{iy} > \max_{y'\neq y} U^1_{xy'} + \varepsilon_{iy'} \text{ AND } \max_{y'\neq y} U^2_{xy'} + \varepsilon_{iy'} > U^2_{xy} + \varepsilon_{iy}\right\}.$$

Note that $\arg\max_y\{U^1_{xy} + \varepsilon_{iy}\}$ and $\arg\max_y\{U^2_{xy} + \varepsilon_{iy}\}$ are both singletons, and they are different under event B. Since P_x is of full support, we have $P(B) > 0$. This implies that

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

occurs with strictly positive probability, and thus (5) $>$ (6) 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. □

We will show the following, which corresponds to the first part of the theorem:

Lemma A.6. *Under Assumptions 1–4, (P) and (D) have unique solutions for any* Φ.

Proof. 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 concave in μ, which implies the uniqueness of the solution.

For (D), suppose toward a contradiction that there are two different optimal solutions $\alpha \coloneqq (U, V, \bar{w}, \underline{w})$ and $\beta \coloneqq (U', V', \bar{w}', \underline{w}')$. Note that $\gamma \coloneqq \frac{1}{2}(\alpha + \beta)$ 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 $(\bar{w}, \underline{w}) \neq (\bar{w}', \underline{w}')$. By the WDZ theorem and the full support assumption, we have $\frac{\partial G}{\partial U_{xy}} = \mu_{xy} > 0$ 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

$$\bar{w}_z - \underline{w}_z = \bar{w}'_z - \underline{w}'_z. \tag{7}$$

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 $(U, V, \tilde{\bar{w}}, \tilde{\underline{w}})$, where $\tilde{\bar{w}}_z \coloneqq \bar{w}_z - \varepsilon$ and $\tilde{\underline{w}}_z \coloneqq \underline{w}_z - \varepsilon$ attains a strictly lower value. Similarly, we have $\bar{w}'_z \underline{w}'_z = 0$. However, these combined with (7) 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. □

By the complementary slackness condition, for each z, only one of the following holds: $\bar{w}^*_z > 0$ and $\underline{w}^*_z = 0$, $\bar{w}^*_z = 0$ and $\underline{w}^*_z > 0$, or $\bar{w}^*_z = \underline{w}^*_z = 0$. If we define

$$w^*_z \coloneqq \mathbf{1}\{\bar{w}^*_z > 0\}\bar{w}^*_z - \mathbf{1}\{\underline{w}^*_z > 0\}\underline{w}^*_z,$$

then $w^* = (w^*_z)_z$ is the taxation policy, and we have

$$\Phi_{xy} - \bar{w}^*_{z(y)} + \underline{w}^*_{z(y)} = \Phi_{xy} - w^*_{z(y)},$$

which corresponds to the stability condition with the gross surplus under taxation policy w^*. Thus, the primal and dual problems jointly compute the equilibrium matching and corresponding social welfare under the taxation policy w^*. The triple (μ^*, U^*, V^*) then constitutes the corresponding aggregate-level matching and utilities. □

A.5 Proof of Proposition 2

Let $\omega_{sh} = \frac{\mu_{sh}}{\sum_{s'} \mu_{s'h}}$. Then,

$$\begin{aligned}
\iota_h &= \frac{1}{M_h} \sum_{(i,j):\, j \in h,\, i \text{ matches with } j} \tau_{ij} \\
&= \frac{1}{M_h} \sum_{j \in h} \sum_{i \in I} \tau_{ij} \mathbf{1}\left\{ j = \arg\max_{j' \in J} \left\{ U^{\text{base}}_{s(i)h(j')} + \tau_{ij'} + \varepsilon_{ih(j')}, \varepsilon_{i0} \right\} \right\} \\
&= \sum_{j \in h} \sum_{s} \sum_{i \in s} \tau_{ij} \frac{\mathbf{1}\left\{ j = \arg\max_{j' \in J} \left\{ U^{\text{base}}_{s(i)h(j')} + \tau_{ij'} + \varepsilon_{ih(j')}, \varepsilon_{i0} \right\} \right\}}{\sum_{s'} \mu_{s'h}} \\
&= \sum_{s} \omega_{sh} \frac{1}{\mu_{sh}} \sum_{j \in h} \sum_{i \in s} \tau_{ij} \mathbf{1}\left\{ j = \arg\max_{j' \in J} \left\{ U^{\text{base}}_{s(i)h(j')} + \tau_{ij'} + \varepsilon_{ih(j')}, \varepsilon_{i0} \right\} \right\}.
\end{aligned}$$

Now the inside of the sum is

$$\begin{aligned}
&\tau_{ij} \mathbf{1}\left\{ j = \arg\max_{j' \in J} \left\{ U^{\text{base}}_{s(i)h(j')} + \tau_{ij'} + \varepsilon_{ih(j')}, \varepsilon_{i0} \right\} \right\} \\
&= \left(u_i - \varepsilon_{ih} - U^{\text{base}}_{sh} \right) \mathbf{1}\left\{ u_i = U^{\text{base}}_{sh} + \tau_{ij} + \varepsilon_{ih} \right\} \\
&= \left(\max_{h'} \left\{ U_{sh'} + \varepsilon_{ih'} \right\} - \varepsilon_{ih} - U^{\text{base}}_{sh} \right) \mathbf{1}\left\{ u_i = U^{\text{base}}_{sh} + \tau_{ij} + \varepsilon_{ih} \right\} \\
&= \left(U_{sh} - U^{\text{base}}_{sh} \right) \mathbf{1}\left\{ u_i = U^{\text{base}}_{sh} + \tau_{ij} + \varepsilon_{ih} \right\}.
\end{aligned}$$

Hence,

$$\iota_h = \sum_{s} \omega_{sh} \left(U_{sh} - U^{\text{base}}_{sh} \right) \frac{1}{\mu_{sh}} \sum_{j \in h} \sum_{i \in s} \mathbf{1}\left\{ u_i = U^{\text{base}}_{sh} + \tau_{ij} + \varepsilon_{ih} \right\} = \sum_{s} \omega_{sh} \left(U_{sh} - U^{\text{base}}_{sh} \right).$$

Galichon and Salanié (2021*a*) show that U_{sh} for all pairs of school and hospital can be identified by the aggregate-level matching μ. Then, the conditional moment is

$$\begin{aligned}
\mathbb{E}\left[\iota_h \mid X, \mu \right] &= \mathbb{E}\left[\sum_{s} \omega_{sh} \left(U_{sh} - U^{\text{base}}_{sh} \right) \mid X, \mu \right] \\
&= \sum_{s} \omega_{sh} U_{sh} - \sum_{s} \omega_{sh} \left(X^{U,\text{base}\top}_{sh} \beta_U \right).
\end{aligned}$$

For the hospital side, the same argument can be applied to show the conditional

moment.

A.6 Derivation of two-way fixed effects Poisson regression of Galichon and Salanié (2021*b*)

We derive the estimator based on two-way fixed effects Poisson regression, which is proposed in Galichon and Salanié (2021*b*). The derivation is for the estimation of Φ and the same argument can be applied to the estimation of aggregate level utility, U and V, which is transformed into a single-side fixed effect Poisson regression.

Consider a Poisson regression with two-way FE. The observation is matching between type x and y: μ_{xy}. We assume that $\mu_{xy} \sim \text{Po}(\theta_{xy})$ where $\theta_{xy} = e^{\frac{\sum_k \lambda_k \phi_{xy}^k - u_x - v_y}{2}}$, $\mu_{x0} \sim \text{Po}(e^{-u_x})$, and $\mu_{0y} \sim \text{Po}(e^{-v_y})$.

The likelihood is as follows: note that we double-count the matchings when we construct the likelihood.

$$\prod_{xy} \frac{e^{\mu_{xy} \frac{\sum_k \lambda_k \phi_{xy}^k - u_x - v_y}{2}} e^{-e^{\frac{\sum_k \lambda_k \phi_{xy}^k - u_x - v_y}{2}}}}{\mu_{xy}!} \prod_{xy} \frac{e^{\mu_{xy} \frac{\sum_k \lambda_k \phi_{xy}^k - u_x - v_y}{2}} e^{-e^{\frac{\sum_k \lambda_k \phi_{xy}^k - u_x - v_y}{2}}}}{\mu_{xy}!}$$
$$\prod_{x} \frac{e^{\mu_{x0} u_x} e^{-e^{-u_x}}}{\mu_{x0}!} \prod_{y} \frac{e^{\mu_{0y} v_y} e^{-e^{-v_y}}}{\mu_{0y}!}$$

Then the log-likelihood function is written as follows:

$$\begin{aligned}
G(\lambda, u, v) \equiv & \, 2 \sum_{xy} \ln e^{\mu_{xy} \frac{\sum_k \lambda_k \phi_{xy}^k - u_x - v_y}{2}} e^{-e^{\frac{\sum_k \lambda_k \phi_{xy}^k - u_x - v_y}{2}}} \\
& + \sum_x \ln e^{-\mu_{x0} u_x} e^{-e^{-u_x}} + \sum_y \ln e^{-\mu_{0y} v_y} e^{-e^{-v_y}} \\
= & \, 2 \sum_{xy} \mu_{xy} \frac{\sum_k \lambda_k \phi_{xy}^k - u_x - v_y}{2} - e^{\frac{\sum_k \lambda_k \phi_{xy}^k - u_x - v_y}{2}} \\
& + \sum_x -\mu_{x0} u_x - e^{-u_x} + \sum_y -\mu_{0y} v_y - e^{-v_y} \\
= & \sum_{xy} \mu_{xy} \left(\sum_k \lambda_k \phi_{xy}^k - u_x - v_y \right) - 2 e^{\frac{\sum_k \lambda_k \phi_{xy}^k - u_x - v_y}{2}} \\
& - \sum_x \mu_{x0} u_x - \sum_x e^{-u_x} - \sum_y \mu_{0y} v_y - \sum_y e^{-v_y}
\end{aligned}$$

The minimization objective function is

$$F(\lambda, u, v) = \sum_x e^{-u_x} + \sum_y e^{-v_y} + 2\sum_{xy} e^{\frac{\sum_k \lambda_k \phi_{xy}^k - u_x - v_y}{2}} - \sum_{xy} \mu_{xy} \left(\sum_k \lambda_k \phi_{xy}^k - u_x - v_y \right) + \sum_x \mu_{x0} u_x + \sum_y \mu_{0y} v_y$$

As shown in Galichon and Salanié (2021*b*), minimizing F is equivalent to the moment matching estimator.

B Comparison between Aggregate Equilibria, Stable Outcomes, and Deferred Acceptance Outcomes

In this section, we clarify the relationships among aggregate equilibrium (AE), stable outcome, and the outcome under the Deferred Acceptance (DA) mechanism. In our model, agents face nonzero outside options: doctor i obtains utility ε_{i0} when unmatched, and slot j obtains utility η_{0j} when unmatched.

By construction, the social welfare achieved under the DA mechanism is lower than that under the stable outcome. Furthermore, we demonstrate that, in a suitably defined large market limit, the per-capita social welfare under the stable outcome and AE converges. These results are established through both theoretical analysis (Section B.1) and simulation (Section B.2).

B.1 Theory

Below, we claim that social welfare *per capita* (i.e., total surplus divided by the number of doctors and positions) under the stable outcome almost surely converges to the (modified version of) optimal value of AE. For notational simplicity, we assume $w_z \equiv 0$ below. For general taxation policy z, the same argument goes through with Φ_{ij} being replaced by the gross surplus.

Let $L \coloneqq |I| + |J|$. We consider the large market limit in which n_x and m_y tends to ∞ for all x, y with

$$\frac{n_x}{L} \to p_x, \quad \frac{m_y}{L} \to p_y \quad (n_x, m_y \to \infty). \tag{8}$$

Let $(u_i)_i, (v_j)_j$ denote the equilibrium payoff profiles under the stable outcome with non-zero-valued outside option values, which is the solution to the following dual problem (D_S):

$$(D_S) \begin{cases} \min\limits_{u,v} & \sum\limits_{i\in I} u_i + \sum\limits_{j\in J} v_j \\ s.t. & u_i + v_j \geq \Phi_{ij} \\ & u_i \geq \varepsilon_{i0} \\ & v_j \geq \eta_{0j} \end{cases}$$

The corresponding primal problem P_S is

$$(P_S) \begin{cases} \max\limits_{\mu} & \sum\limits_{ij\in IJ} \mu_{ij} \left(\Phi_{ij} - \varepsilon_{i0} - \eta_{0j}\right) + \sum\limits_{i\in I} \varepsilon_{i0} + \sum\limits_{j\in J} \eta_{0j} \\ s.t. & \mu \text{ is feasible} \end{cases}$$

Observe that the value of the objective function of (P_S) corresponds to social welfare under the stable outcome with non-zero-valued outside options since (u, v) satisfies the corresponding stability condition (the constraints in (D_S)) and

$$\begin{aligned} &\sum_{ij\in IJ} \mu_{ij} \left(\Phi_{ij} - \varepsilon_{i0} - \eta_{0j}\right) + \sum_{i\in I} \varepsilon_{i0} + \sum_{j\in J} \eta_{0j} \\ &= \sum_{ij\in IJ} \mu_{ij}\Phi_{ij} + \sum_{i\in I} \left(1 - \sum_{j} \mu_{ij}\right) \varepsilon_{i0} + \sum_{j\in J} \left(1 - \sum_{i} \mu_{ij}\right) \eta_{0j}, \end{aligned}$$

which means that the matched pair with $\mu_{ij} = 1$ yields joint surplus Φ_{ij} and the unmatched agents yields surplus ε_{i0} or η_{0j}.

We consider a version of the dual for AE:

$$\bar{V}(AE) \coloneqq \min_{U,V} \left\{ \sum_{x\in X} p_x \mathbb{E}\left[\max_{y\in Y_0} \{U_{xy} + \varepsilon_{iy}\}\right] + \sum_{y\in Y} q_y \mathbb{E}\left[\max_{y\in Y_0} \{V_{xy} + \varepsilon_{xj}\}\right] \right\}. \quad (9)$$

Note that the weights p_x and q_y are used instead of n_x and m_y in the current main text.

The social welfare per capita under the stable outcome with non-zero-valued outside

options is denoted by

$$\begin{aligned}\bar{V}(S) &\coloneqq \frac{1}{L}\left[\sum_{x\in X}\sum_{i\in x} u_i + \sum_{y\in Y}\sum_{j\in y} v_j\right] \\ &= \sum_{x\in X}\frac{n_x}{L}\frac{1}{n_x}\sum_{i\in x} u_i + \sum_{y\in Y}\frac{m_y}{L}\frac{1}{m_y}\sum_{j\in y} v_j.\end{aligned}$$

By Lemma 1, for any $i \in x$ and $j \in y$, we have

$$u_i = \max_{y\in Y_0}\{U_{xy} + \varepsilon_{iy}\}, \quad v_j = \max_{x\in X_0}\{V_{xy} + \varepsilon_{xj}\}. \tag{10}$$

By the SLLN and (10), we have

$$\frac{1}{n_x}\sum_{i\in x} u_i \to \mathbb{E}\left[\max_{y\in Y_0}\{U_{xy} + \varepsilon_{iy}\}\right], \ a.s. \quad (n_x \to \infty),$$

and

$$\frac{1}{m_y}\sum_{j\in y} v_j \to \mathbb{E}\left[\max_{y\in Y_0}\{V_{xy} + \varepsilon_{xj}\}\right], \ a.s. \quad (m_y \to \infty).$$

By (8), we have

$$\bar{V}(S) \to \bar{V}(AE), \ a.s. \quad (n_x, m_y \to \infty).$$

□

We compute (D_S) for the stable outcome in our simulation, and compare its social welfare per capita (i.e., the optimal value of (D_S) divided by L) with the social welfare per capita in AE computed by (9).

B.2 Simulation

We consider a matching market with four medical schools, denoted by $X = \{x_1, x_2, x_3, x_4\}$, and three hospitals, denoted by $Y = \{y_1, y_2, y_3\}$. Each medical school admits n students, and each hospital offers m available slots. The aggregate-level joint surplus generated by

a match between a university x and a hospital y is given by the following matrix:

$$(\Phi_{xy})_{xy} = \begin{pmatrix} 5 & 3 & 2 \\ 0 & 4 & 1 \\ 4 & 2 & 3 \\ -1 & 2 & 2 \end{pmatrix}.$$

Given the matrix $(\Phi_{xy})_{xy}$ and the population of doctors and hospital slots, we first compute the AE for this market. This computation yields the number of matches across all types as well as the aggregate-level utilities U_{xy} and V_{xy}. In addition, we obtain the welfare measures G and H for both sides.

At the individual level, the utility for a doctor i of type x and for a slot j of type y is given by

$$U_{xy} + \varepsilon_{iy} \quad \text{and} \quad V_{xy} + \eta_{xj},$$

respectively. Summing these, we define the individual-level social surplus as

$$\Phi_{ij} = U_{xy} + \varepsilon_{iy} + V_{xy} + \eta_{xj}.$$

Note that this individual-level utility is defined only for doctor–slot pairs; in cases of unmatch, a doctor i obtains utility ε_{i0} and a slot j obtains utility η_{0j}. We assume all unobserved terms $(\varepsilon_{iy})_{y \in Y_0}$ and $(\eta_{xj})_{x \in X_0}$ follow a Gumbel distribution.

Using these simulated values of utility and surplus, we compute two types of market outcomes: the stable outcome and the outcome from the deferred acceptance algorithm. For the stable outcome, we solve (D_S) to obtain the equilibrium. For the DA outcome, we first construct preference lists based on the computed utilities and then implement the standard DA algorithm to determine the matching.

We examine four market settings in which n and m grow, with the pairs given by

$$(n, m) \in \{(20, 40),\ (40, 60),\ (60, 80),\ (80, 100)\}.$$

For each market setting, we compute the equilibrium outcome 50 times and report the average results. Figure 4 compares the three market outcomes. First, AE closely approximates both the average welfare on each side and the matching outcomes observed

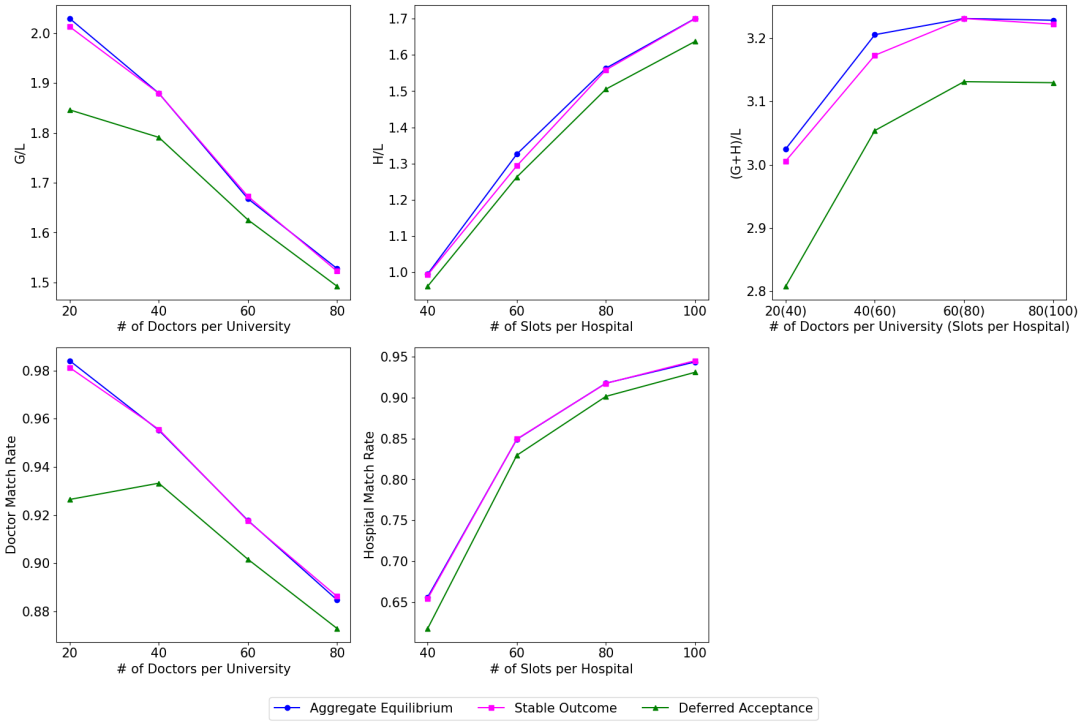


Figure 4. AE and individual market outcomes

Note: The upper panels present the welfare measures. The leftmost panel compares the average welfare on the doctors' side, the center panel compares the average welfare on the hospitals' side, and the rightmost panel compares the total social welfare, which is the sum of the values in the first two panels. The lower two panels compare the matching outcomes. The left panel shows the match rate on the doctors' side, defined as the ratio of matched doctors to the total number of doctors, and the right panel shows the match rate on the hospitals' side, defined as the ratio of filled slots to the total number of slots. The blue line represents AE, the magenta line represents the stable outcome, and the green line represents the DA outcome.

in the stable outcome, even when the numbers of doctors and slots are relatively small. Second, as expected, the average social welfare in the DA outcome is lower than that in the stable outcome. Nevertheless, as the market size increases, the welfare measures and matching outcomes under DA exhibit trends similar to those observed in AE and the stable outcome.

C Monte Carlo Simulation

We start by describing the overall setting of the Monte Carlo simulation. All the detailed parameter values are left to Appendix C.3. 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 is concerned

about the inefficient supply of medical services in R_2 and attempts to meet a lower bound in terms of the number of matches in the region.

We have a total of 20 hospitals. 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 to the hospitals. Each doctor belongs to one of 20 medical schools, and the schools themselves are distributed among prefectures, also based on a multinomial distribution. Each school has an equal split of the doctors in the same prefecture. The schools have characteristics such as average ability measures that follow 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

$$\Phi_{sht} = U^{\text{base}}_{sht} + V^{\text{base}}_{sht},\ \xi_{ijt} = \varepsilon_{iht} + \eta_{sjt},$$

and

$$\begin{aligned} U^{\text{base}}_{sht} &= \beta_{1,1} w_{1,ht} + \beta_{1,2} w_{2,ht} + \beta_2 \left| l_s - l_h \right| + \beta_3 1\{h \in R_1 \text{ or } h \in R_2\}, \\ V^{\text{base}}_{sht} &= \gamma_1 x_{1,st} + \gamma_2 x_{2,st}, \\ \varepsilon_{iht} &\sim Ex1,\ \eta_{sjt} \sim Ex1. \end{aligned} \tag{11}$$

$|l_s - l_h|$ 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^{base}_{sht} captures the negative impact on the utility from living in rural areas. Note that these rural areas include R_1, which is not the target of the subsidy to ensure the lower bound on the matching outcomes. We also use $U_{ijt} = U^{\text{base}}_{sh(j)t} + \varepsilon_{ih(j)t}$ and

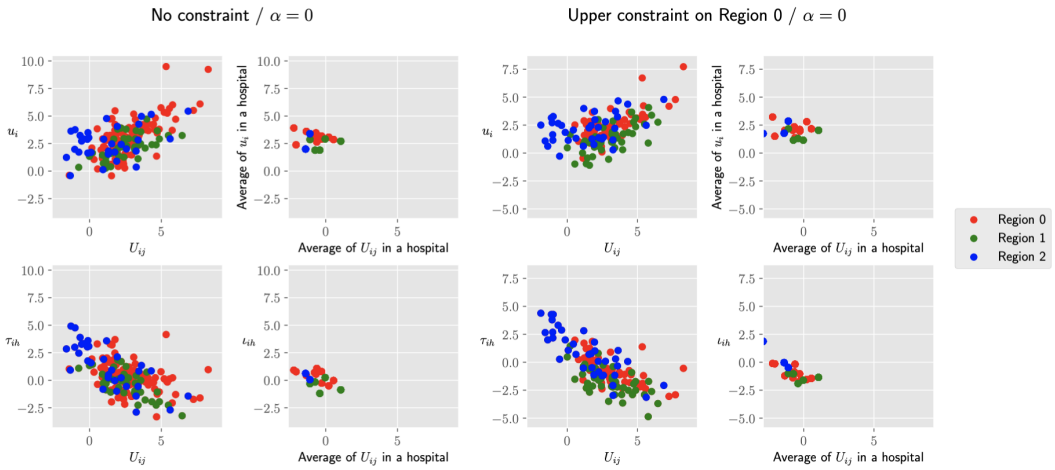


Figure 5. Simulated Stable Outcomes

$V_{ijt} = V^{\text{base}}_{s(i)ht} + \eta_{s(i)jt}$ to denote the individual level preferences.

C.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 unmatched slots is 46. Imagine that the government sets 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, the equilibrium numbers of matches are: 60, 45, and 37. The number of unmatched doctors is 58 and the number of unmatched slots is 74. Under this regional constraint, the tax levied on the matchings in R_0 is 3.149.

Figure 5 depicts the scatter plots of several equilibrium objects when we set $\alpha = 0$, which implies that the entire amount of tax is levied on the hospital side. The left four panels 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 third 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,

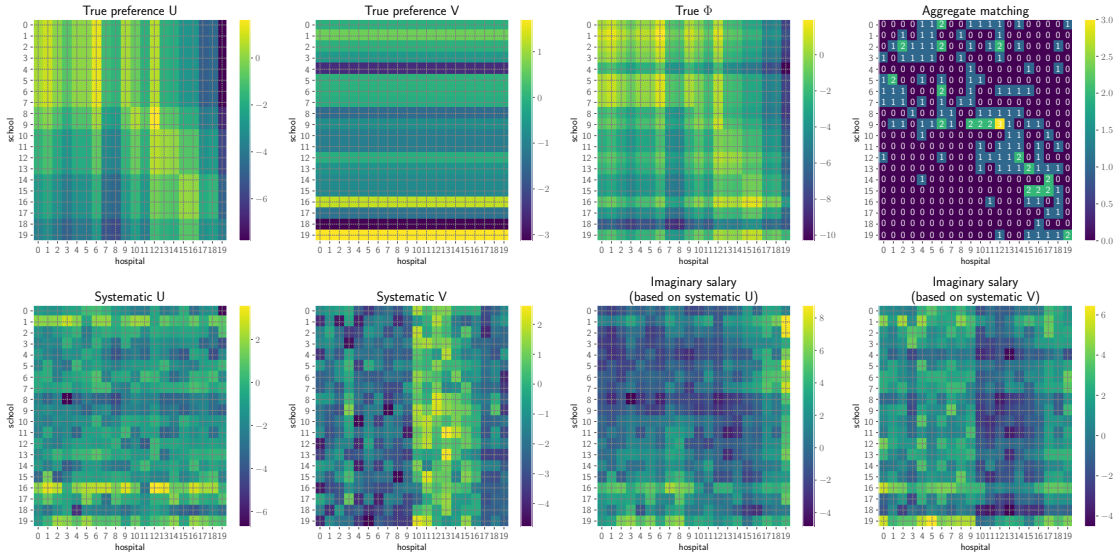


Figure 6. Aggregate Objects

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 objects: 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 marker 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 transfers from a hospital: when the average of U_{ij} in a hospital match increases, the aggregate-level transfers from the hospital decrease. 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 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 the hospital side and the school side.

Hereafter, we set $\alpha = 0.2$. Figure 6 summarizes the aggregate-level objects computed based on the simulated stable outcome. In all the heatmaps, the horizontal axis represents hospitals and the vertical axis represents schools. Aggregate matching is depicted in the

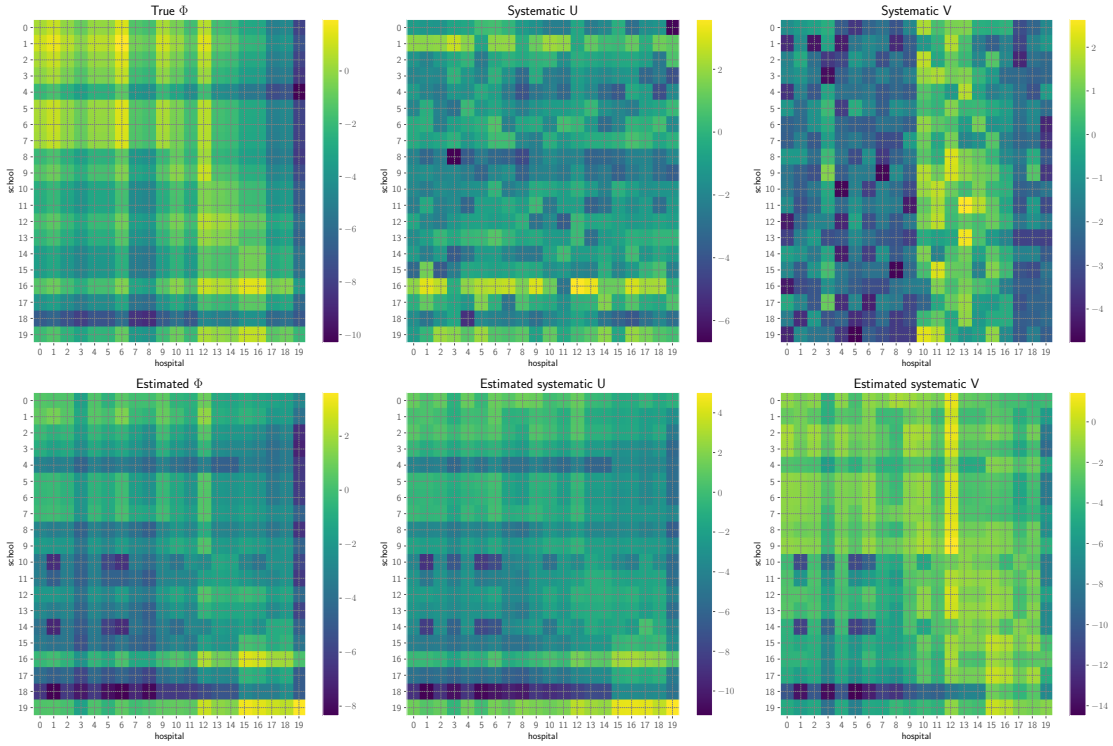


Figure 7. Estimation Results of the First Stage

upper panel in the rightmost 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 (1).

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 $\chi^V_{sh} \equiv V^{\text{base}}_{sh} - V_{sh}$.[34] The lower left two panels in Figure 6 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 an infinite number of doctors. This makes the gap between the two imaginary wages computed based on U and V.[35]

[34]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.

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

Table 12. 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

C.2 Estimation

The estimation results in the first stage are depicted in Figure 7. 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 display the estimation results corresponding to the upper panels. We set the degree of the polynomials to two. The estimated social surplus follows similar patterns to the true social surplus, whereas the estimated aggregate-level utilities show different patterns from the true values. These gaps are due to the incompleteness of polynomial approximations in equation (**??**). In practice, we handle this problem by including non-linearly transformed base variables when making polynomial series.

For the second-stage estimation, we simulate matching outcomes over 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 is observable, we use the moment conditions (**??**) directly to construct a minimum distance estimator. We leave the details of the construction of this estimator in Appendix C.3. In this exercise, we use the time-average version of the moment conditions, and therefore, the tax term is simply identified as the time-average of the levied tax.

Table 12 summarizes the estimation results of the second stage. The first six columns are the structural parameters in equation (11). The last two columns are the average taxes levied on the doctor side and the 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 a counterfactual analysis: for example, the taxes in the alternative regional constraints, the matching outcomes, and the salaries are obtained by solving the equilibrium.

C.3 Minimum distance estimator

We use the time average of both sides to construct the moment conditions (**??**). Note that, in this case, it is impossible to identify δ_{ht}^U and δ_{ht}^V for every t because the summation of them with respect to time determines the moment values. Hence, all we can identify is the average tax levied over the time periods. We define a function g to represent the moment conditions:

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

Our estimator is the minimum distance estimator where the moment condition is specified in (**??**). When we write the asymptotic variance of $\hat{\tilde{U}}_{sht}$ and $\hat{\tilde{V}}_{sht}$ 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 distribution of the estimator as follows, and the standard error can be obtained directly[36]. As the Poisson regression in the first step has an explicit form of S_t^U and S_t^V, we can directly compute the estimates of the standard errors for every parameter by inserting the estimated results.

Theorem C.1. *Under the regularity conditions, the asymptotic distribution of* $\hat{\theta}$ *is as*

[36]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, which is beyond the scope of this study.

follows:

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

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

D Additional analysis

D.1 Test of implicit tax on urban areas

Under the excessive competition for the slots in urban counties, it is possible that the surplus generated by matches in urban counties has already been diminished due to some external forces: for example, as the number of slots decreases, it becomes more difficult for residency programs to secure funding. Given this consideration, the marginal effect estimates for locations in Tokyo or other urban areas may be underestimated. Here, we examine whether matches in urban counties are subject to an implicit tax under the current market outcome.

Empirical strategy We define an individual level transfer as in the main analysis. Fix any period t and an implicit taxation policy $w_t = (w_{zt})_z$. 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 have the same moment conditions for this case as in the main analysis.

As in the main analysis, we model the base utilities as a linear function of observable characteristics. In addition to them, we define $\delta^U_{ht} \equiv \alpha w_{r(h)t}$ and $\delta^V_{ht} \equiv (1-\alpha) w_{r(h)t}$ as the levied implicit tax on school side and hospital side in period t and treat them as parameters to be estimated. Our parameters of interest are the following: β_U, β_V, δ^U_{ht} for every pair of h and t, and δ^V_{ht} for every pair of z and t. We use θ to indicate the vector of these parameters: $\theta \coloneqq \left(\beta_U, (\delta^U_{ht})_{h,t}, \beta_V, (\delta^V_{ht})_{h,t}\right)$.

For the estimation of the aggregate-level utilities are same as our main analysis. For the second step, we construct the following moment conditions for θ:

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

By adopting the same measurement model, the estimating equations are as follows:

$$\sum_s \omega_{sht}\hat{U}_{sht} = \gamma_{0,U} + \gamma_{1,U}S_{ht} + \sum_s \omega_{sht}\left(X^{U\prime}_{sht}\beta_U - \delta^U_{ht}\right) + \psi^U_{ht}$$
$$\sum_s \omega_{sht}\hat{V}_{sht} = \gamma_{0,V} + \gamma_{1,V}S_{ht} + \sum_s \omega_{sht}\left(X^{V\prime}_{sht}\beta_V - \delta^V_{ht}\right) + \psi^V_{ht}.$$

We estimate these linear equations using BLP-type IVs.

Estimation results We take advantage of the fact that the regional constraints on urban areas are getting the more strict as time goes to clarify the existence of implicit tax. 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 surplus in urban areas have been decreased due to the constraints, the estimated coefficients on dummy variables of urban or Tokyo will decrease over the years.

Table 13 shows the estimation results based on IV estimation: where we include all the covariates in Table 5 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

Table 13. 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

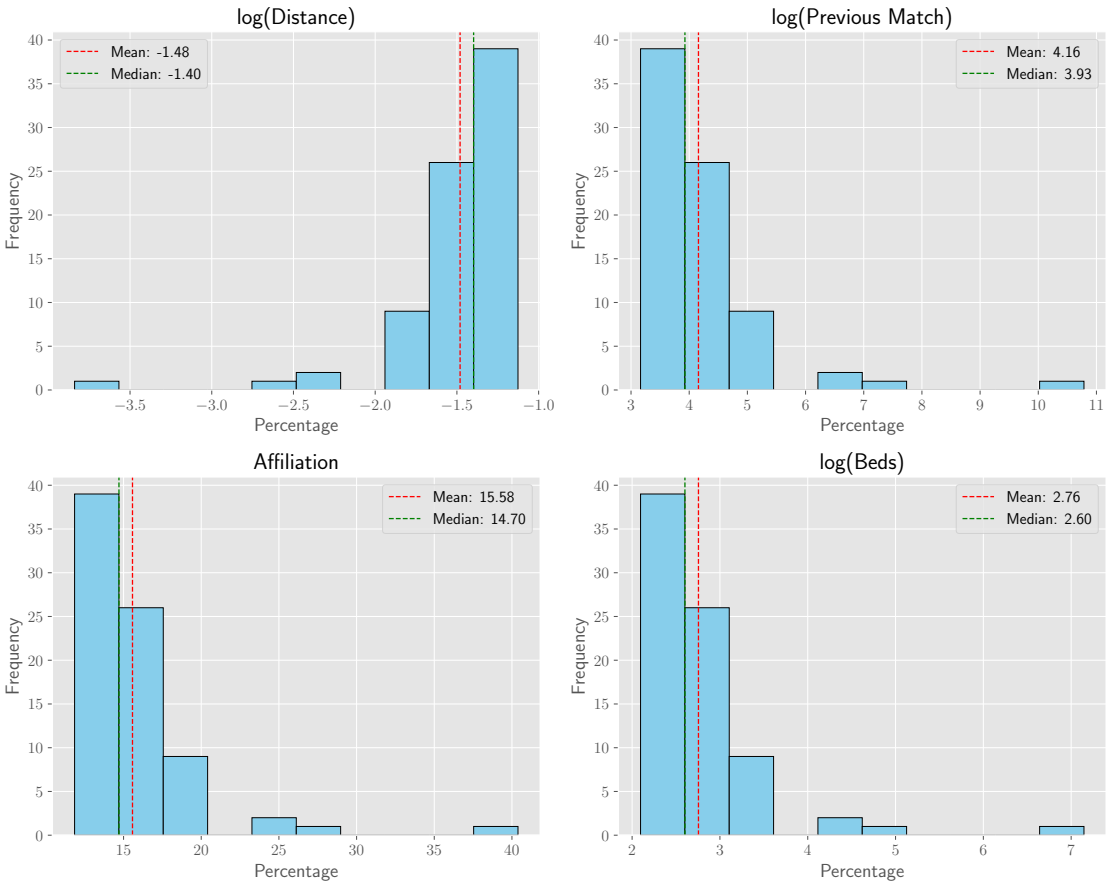


Figure 8. Relative Size of Coefficients in Doctor's Preference.

for living in Tokyo in 2018. Based on these results, we conclude that the current market does not suffer from any implicit tax and the market outcome is the aggregate equilibrium under the reduced capacities.

D.2 Relative impacts

To grab the sizes of impacts, we compute the ratio of these coefficients to the aggregate-level utility. Specifically, we first transform the aggregate-level utilities into monetary unit based on the estimation results in the second stage estimation: for doctor side, we compute $U_{sht}^{money} \equiv \frac{\hat{\bar{U}}_{sht} - \hat{\gamma}_{0,U,t}}{\hat{\gamma}_{1,U}}$ and for hospital side, we compute $V_{sht}^{money} \equiv \frac{\hat{\bar{V}}_{sht} - \hat{\gamma}_{0,V,t}}{\hat{\gamma}_{1,V}}$.[37] Then, we take the average of these transformed aggregate-level utilities with respect to the periods and the institutions in the other side of the market: $\bar{U}_s^{money} \equiv \frac{1}{HT} \sum_{h,t} U_{sht}^{money}$

[37]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.

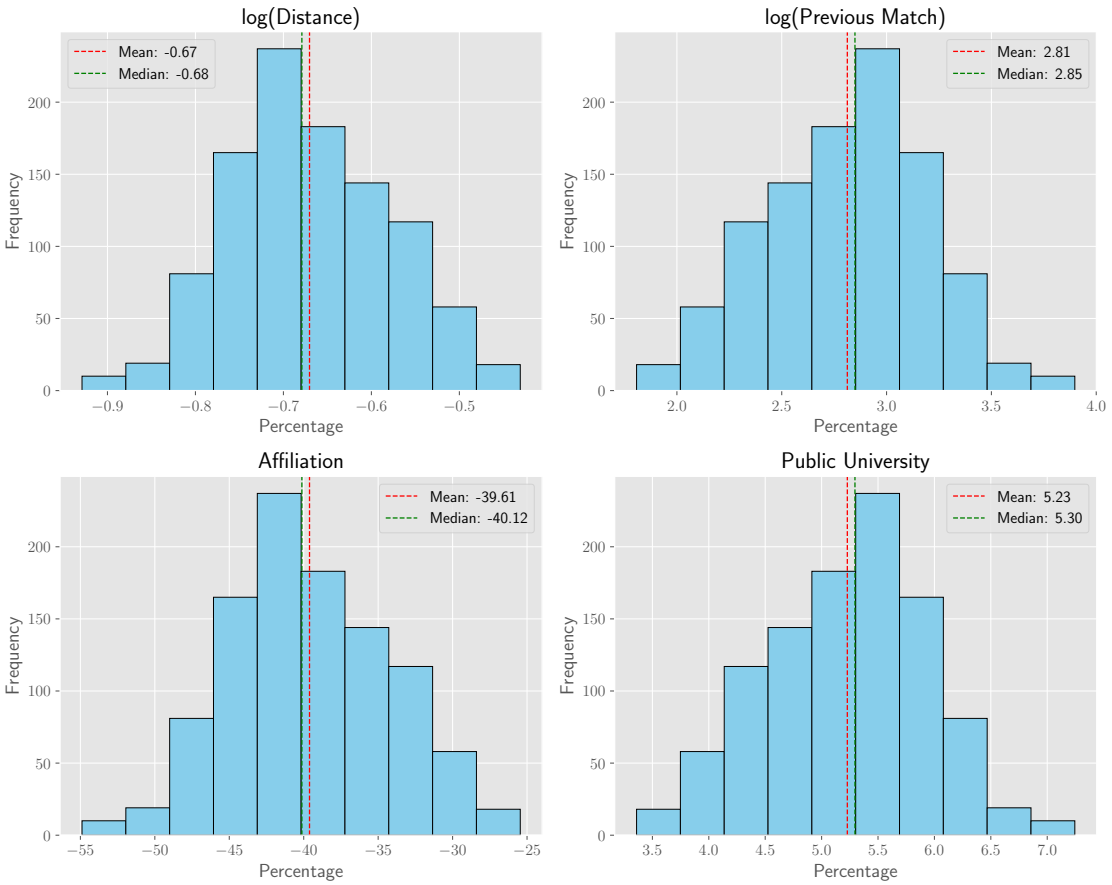


Figure 9. Relative Size of Coefficients in Hospital's Preference.

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.

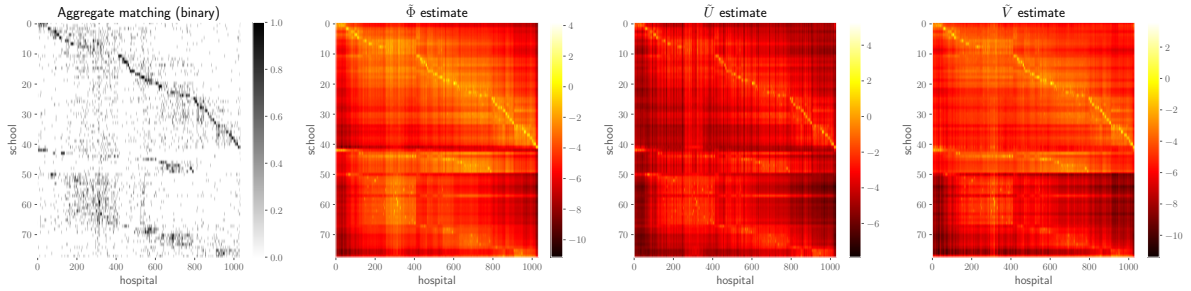
Figure 8 depicts the histograms of the relative size of the coefficients in doctor's preference for the four covariates which have statistically significant influence in Table 6: the logarithm of distance, the logarithm of the number of previous matches, the dummy variable of affiliation, and the logarithm 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.

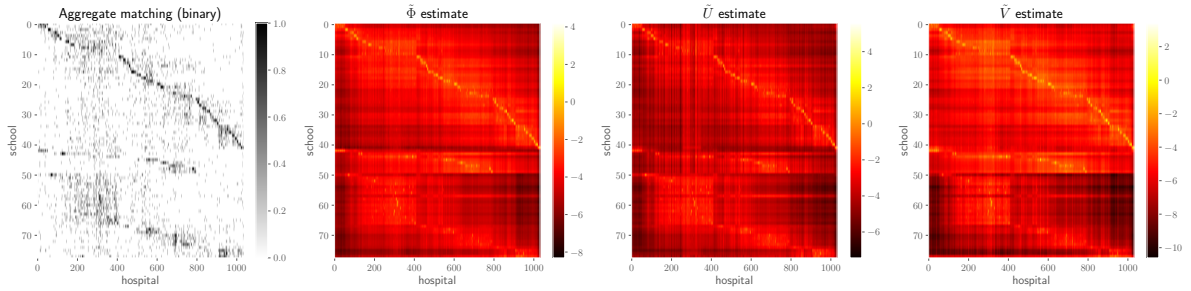
Figure 9 shows the same histograms for the hospital side preference. We plot the historgrams of the four covariates which shows the statistically significance in Table 7: the logarithm of distance, the logarithm of the number of previous matches, affiliation relationship, and the indicator of the public 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 D.3.

D.3 Results when the degree of polynomial is set to 2

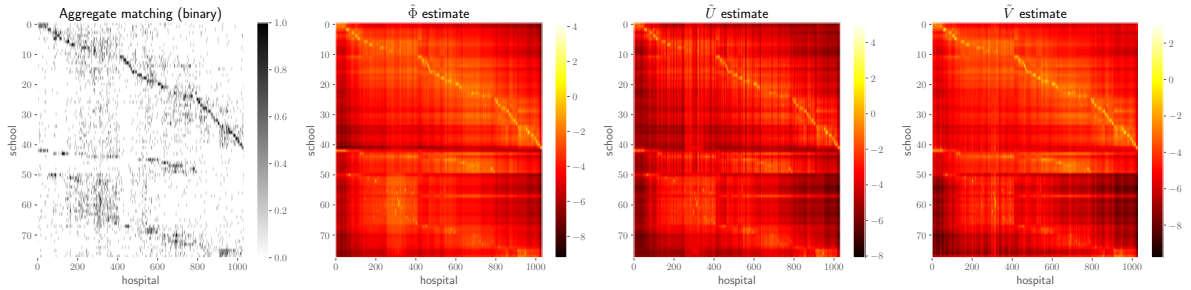
Here we show the empirical results obtained when we set the degree of polynomials in the first step to two. All the tables and figures listed here corresponds to the tables and figures shown in Section 6 and Appendix D.1. We do not find any qualittative difference in the main findings from the case where we set the degree of polynomials to three.



(a) 2017



(b) 2018



(c) 2019

Figure 10. Aggregate matchings, estimated systematic utilities and estimated social surpluses.

Table 14. 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

Table 15. 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

Table 16. 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		√	√
Tokyo × Year			√

Table 17. 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			√

D.4 Counterfactual simulations for the other years

Table 18. Comparison between Aggregate-level Equilibria

Policy	AC (Artificial Caps)	NC (No Caps)	OS (Optimal Subsidy)
Artificial caps	Yes	No	No
Floor constraints	Yes	No	Yes
Subsidies	No	No	Yes
2017			
Match rate	0.868	0.912	0.912
Doctors' welfare	32795.8	33441.7	33444.5
Hospitals' welfare	29586.6	31578.1	31579.4
Government's revenue	0.0	0.0	$[-10.5, -7.4]$
Total welfare	62382.3	65019.8	$[65013.3, 65016.5]$
#(subsidized regions)	0	0	3
Average subsidy	0.000	0.000	-0.040
#(constraint violations)	0	3	0
2018			
Match rate	0.844	0.895	0.896
Doctors' welfare	33928.3	34272.6	34277.1
Hospitals' welfare	30390.4	33700.2	33703.4
Government's revenue	0.0	0.0	$[-18.9, -13.3]$
Total welfare	64318.7	67972.8	$[67961.6, 67967.1]$
#(subsidized regions)	0	0	5
Average subsidy	0.000	0.000	-0.038
#(constraint violations)	0	5	0
2019			
Match rate	0.869	0.912	0.912
Doctors' welfare	33244.7	33752.2	33755.5
Hospitals' welfare	30406.1	32648.8	32649.1
Government's revenue	0.0	0.0	$[-9.4, -6.6]$
Total welfare	63650.8	66401.0	$[66395.2, 66398.0]$
#(subsidized regions)	0	0	2
Average subsidy	0.000	0.000	-0.042
#(constraint violations)	0	2	0

* 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 6 and Table 7. 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).