Caps vs. Prices: Regulating Matching Markets with Distributional Disparities*

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

August 24, 2026

Abstract

Many matching markets rely on quantity-based instruments, such as caps, to achieve distributional objectives, but little is known about their efficiency relative to price-based instruments. We develop a matching model with transferable utility that incorporates regional constraints and characterize the taxes and subsidies that maximize surplus subject to these constraints. We then embed the model in an aggregate matching framework and show how the optimal policy can be constructed from aggregate match data. We apply the framework to the Japan Residency Matching Program, which limits urban residency positions to encourage training in rural areas. Relative to a benchmark without cap tightening, the policy increases rural placements but reduces social surplus by 3.23 billion yen per month. A targeted subsidy policy achieves the same placement gains in each rural prefecture while reducing social surplus by only 0.15 billion yen per month, recovering 95% of the loss under the current caps. Our results suggest that most of the efficiency loss comes from the choice of instrument rather than from the distributional objective itself.

JEL Classification Codes: C25; C78; H20; I14.
Keywords: Matching with Distributional Constraints; Market Design; Two-sided Matching; Discrete Choice; Subsidies; Taxes; Redistribution.

---

*We thank Matteo Camboni, Yeon-Koo Che, Julien Combe, Alfred Galichon, Adam Kapor, Fuhito Kojima, Marzena Rostek, Paulo Somaini, and Olivier Tercieux for helpful comments and thoughtful discussions. We also thank participants at the NYU econometrics seminar, JEA annual meeting, APIOC 2025, IIOC 2025, EC 2025, the empirical market design session at the SITE conference, UW–Madison theory seminar, and MATCH-UP 2026. This paper was previously circulated under the title "Evaluating the Efficiency of Regulation in Matching Markets with Distributional Disparities."

†Yokohama City University, `ikegami.kei.pr@yokohama-cu.ac.jp`.
‡University of Electro-Communications, `atsushi.iwasaki@uec.ac.jp`
§University of Tokyo, `amatsushita@e.u-tokyo.ac.jp`
¶University of Wisconsin–Madison, `kyohei.okumura@wisc.edu`

1 Introduction

Whether distributional objectives are better pursued through prices or quantities is one of the oldest questions in public economics. Quantity-based instruments, such as caps, floors, and reserves, restrict the set of admissible allocations, whereas price-based instruments, such as taxes and subsidies, preserve that set but alter the payoffs agents face, letting agents adjust their choices. Since Weitzman (1974), a central lesson of the literature has been that neither class of instruments generally dominates: their relative performance depends on the information and instruments available to the planner.[1]

We revisit this question in the context of matching markets, where distributional objectives are typically expressed as restrictions on the set of admissible matchings—for example, upper or lower bounds on the number of matches in designated regions (Kamada and Kojima, 2015). Much of the matching literature has pursued these goals through quantity-based instruments, incorporating the distributional requirements directly into the feasible set or the assignment rule. The design of price-based instruments has received substantially less attention. It remains unclear how such instruments should be designed to achieve a given distributional objective, whether they can implement the efficient outcome, and how their welfare performance compares with that of quantity-based policies.

This paper addresses these questions by developing a framework for designing and evaluating quantity- and price-based instruments in matching markets with distributional objectives. We first characterize the optimal tax-and-subsidy policy in a benchmark model with known preferences, and then show how to construct it from match data. Finally, we apply these results to newly collected data from the Japan Residency Matching Program (JRMP), comparing the resulting price-based instruments with the existing cap-based regulation through counterfactual simulations.

The JRMP provides a natural setting for this comparison. Since centralized matching began in 2004, the concentration of matched residents in urban areas and the associated physician shortages in rural areas have remained a central policy concern. In response, policymakers introduced caps on residency positions in the urban prefectures in 2010 and progressively tightened them after 2015 (see

[1] Contributions in this tradition compare prices and quantities under uncertainty (Weitzman, 1974), prices and rationing when consumers differ in needs and incomes (Weitzman, 1977), quantity controls and taxes in second-best environments (Guesnerie and Roberts, 1984), and price wedges and rationing as redistributive mechanisms under private information (Dworczak et al., 2021).

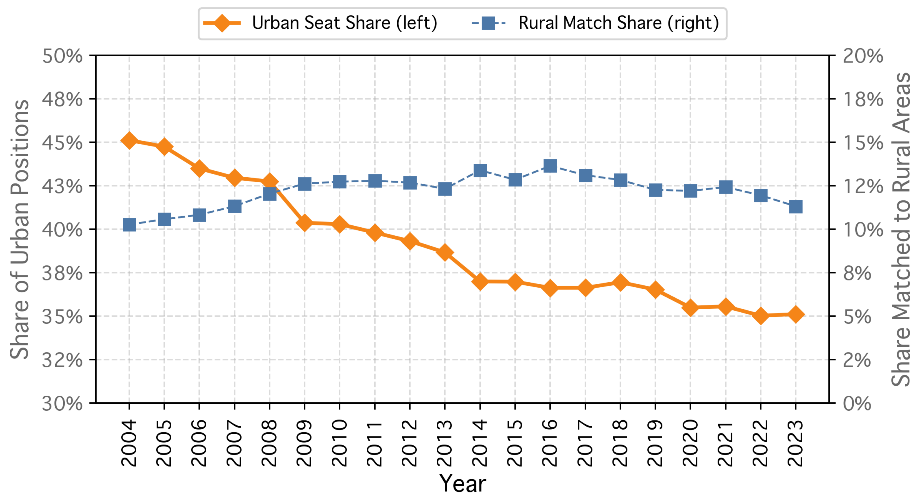


Figure 1. Trends in the share of urban residency positions and the share of applicants matched to rural areas.

Section 2.1 for institutional details). As a result, the share of positions in the urban prefectures fell, yet many rural prefectures gained little match share (Figure 1). Over the same period, the share of applicants unmatched in the centralized JRMP round rose from 4.7 to 9.5 percent. These patterns suggest that the caps have restricted urban matches without fully redirecting applicants into rural positions, raising substantial efficiency cost.

We study a matching model with transferable utility à la Shapley and Shubik (1971), in which a policymaker pursues regional distributional objectives. Each position belongs to a region, and caps and floors specify target ranges for the number of matches within each region. Absent policy intervention, the equilibrium matching need not satisfy these targets. We show that, when the policymaker observes participants' preferences, region-level taxes and subsidies implement the surplus-maximizing matching subject to these constraints; no policy that meets the targets can improve upon this constrained optimum. We then decompose the total policy loss—the surplus gap between the no-intervention benchmark and the cap-based policy—into a *distributional-target cost*, the surplus that must be sacrificed to meet the targets, and an *implementation cost*, the additional cost of achieving them through caps rather than prices.

In practice, policymakers rarely observe individual preferences. In the aggregate matching model of Galichon and Salanié (2022), however, the model primitives can be identified from aggregate matching outcomes under certain structural assumptions. We show that these model primitives suffice to compute the taxes and subsidies that implement the constrained optimum, and to evaluate the surplus under different counterfactual policies. These results show that the objects required for policy design and evaluation are identified from aggregate matching data, leaving open how to recover them from a finite sample and express them in monetary units.

We develop an estimation strategy for this model and apply it to the JRMP. We estimate preferences from school–hospital match counts, which identify the latent utilities that rationalize observed matching patterns. We then use resident salary data to convert these utilities into monetary units. The estimated model fits the observed matching patterns well, including the concentration of matches between nearby schools and hospitals. The estimates also reveal preference heterogeneity on both sides of the market: applicants value larger and geographically closer hospitals, while hospitals assign higher value to applicants from higher-ranked, public, and nearby medical schools.

Using the estimated JRMP model, we conduct several counterfactual policy exercises. We first compare the welfare performance of the status quo caps with that of the surplus-maximizing subsidy policy, holding fixed the distributional targets for a set of rural prefectures. Relative to the no-intervention benchmark, the status quo caps reduce social surplus by 3.23 billion yen per month, whereas the subsidy policy reduces it by only 0.15 billion yen per month. The subsidy policy therefore recovers 3.08 billion yen per month, about 95% of the surplus lost under the status quo caps. Most of the total policy loss is therefore implementation cost—the price of using caps rather than prices—and not the cost of the distributional targets themselves.

Two additional counterfactual exercises reinforce this conclusion. The most efficient cap-based policy we identify recovers only 6.7% of the surplus loss under the status quo cap. This suggests that the efficiency loss arises primarily from the use of quantity restrictions rather than from imperfect calibration of the existing caps. Moreover, we can construct a nearly budget-balanced tax-and-subsidy policy that preserves almost all of the surplus gains from the subsidy-only policy. This implies that the surplus gain via a price-based instrument is achievable even under a budget-balance requirement.

Related literature Our paper relates to the theoretical literature on matching under distributional constraints. Much of this literature studies non-transferable utility (NTU) environments, in which monetary transfers are unavailable as policy instruments. One strand studies affirmative action and other type-based constraints in school choice, including majority quotas, minority reserves, and type-specific lower and upper bounds, and analyzes how these restrictions interact with stability, efficiency, and incentives (Abdulkadiroğlu and Sönmez, 2003; Kojima, 2012; Hafalir et al., 2013; Ehlers et al., 2014). Our setting is more closely related to a second strand that studies aggregate distributional constraints across groups of institutions, including regional caps and more general distributional constraints (Kamada and Kojima, 2015; Fragiadakis and Troyan, 2017).

By contrast, relatively few papers study distributional constraints in matching models with transferable utility (TU). Recent papers by Kojima et al. (2020) and Jalota et al. (2025) examine the existence of equilibria under various constraints. The theory paper most closely related to ours is Yokote (2020), which studies welfare-maximizing taxes and subsidies under firm-specific interval constraints that impose upper and lower bounds on the number of workers hired by each firm (a "worker" and a "firm" in his model correspond to an "applicant" and "hospital" in ours).[2] Our framework instead allows regional constraints that specify lower and upper bounds on the number of matches across multiple hospitals. We further embed this framework in an empirical TU matching model with unobserved heterogeneity, allowing us to construct policies from aggregate match data.

Our paper also contributes to the empirical market design literature that uses estimated matching models to evaluate policies motivated by distributional concerns. In the closely related U.S. residency market, Agarwal (2017), using estimates from Agarwal (2015), compares financial incentives and capacity changes intended to promote rural training. Related work quantitatively evaluates distributional interventions in teacher-assignment markets: improving rural access to effective teachers in Peru (Bobba et al., 2024) and reducing regional inequality in teacher experience in France (Combe et al., 2025). Our contribution is a welfare decomposition that distinguishes the distributional target from the instrument used to achieve it. Holding the regional-placement target fixed across instruments, we construct taxes and subsidies that implement the surplus-maximizing matching subject to that target. This common benchmark lets us decompose the loss under a cap policy into the minimum surplus loss required to meet the target and the

[2] In Yokote (2020), each interval constraint applies to the number of workers hired by a single firm. If firms are interpreted as regions, his framework is related to ours, but an important difference remains: in our setting, matching an applicant to one hospital in a region is distinct from matching the same applicant to another hospital in that region, whereas this within-region differentiation is absent from his model. Thus, our model is not a special case of his. This distinction is analogous to that between the model of Abdulkadiroğlu and Sönmez (2003) and that of Kamada and Kojima (2015).

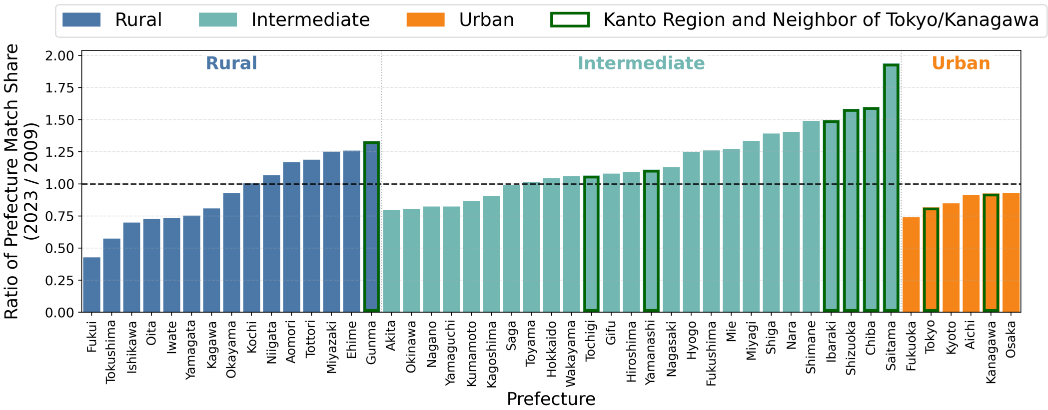


Figure 2. Ratio of prefectural match shares in 2023 relative to 2009.

Notes: The figure reports the ratio of each prefecture's share of matched residents in 2023 relative to its share in 2009. Prefectures near Tokyo or Kanagawa are outlined in green. The 15 rural prefectures are defined by the taxonomy above, and the rural-floor counterfactuals in Section 7 impose one floor for each of them.

additional loss attributable to using caps, rather than prices, to implement it.

2 Data and Institutional Background

Throughout the empirical application, we partition Japan's 47 prefectures into three mutually exclusive groups. The *urban prefectures* are the six prefectures subject to JRMP's urban-cap policy: Tokyo, Kanagawa, Aichi, Osaka, Kyoto, and Fukuoka. We refer to the 15 prefectures with the lowest average match-to-capacity ratios over 2017–2019 as the *rural prefectures*, and the remaining 26 as *intermediate*. We use *non-urban* for the union of the rural and intermediate prefectures. These labels refer to our empirical classification rather than an official administrative designation. Medical schools and hospitals inherit the classification of the prefecture in which they are located.

2.1 Institutional Background of the JRMP

Since 2004, Japan has required medical graduates to complete a two-year residency training program. Applicants are matched to hospitals through the JRMP, a centralized clearinghouse in which applicants and hospitals submit rank-order lists and matches are determined by the deferred acceptance algorithm. Before submitting their lists, applicants gather information about hospitals, including location, size, program features, salary, and workload, and may also attend job fairs or visit hospitals. The JRMP does not regulate resident salaries; hospitals independently determine compensation. Applicants are required to take examinations or attend interviews at the hospitals they are considering. In 2023, the JRMP involved 10,202 applicants from 78 medical schools and 10,895 positions at 1,209 hospitals. Overall, 87.9% of applicants were matched; of the matched applicants, 64.3% were assigned to their first-choice hospital, 16.3% to their second choice, and the remaining 19.4% to their third or a lower-ranked choice.[3]

A central policy concern in the JRMP is the geographic distribution of residents. Previous studies and Japanese media have argued that regional disparities in physician allocation worsened under the JRMP (Iizuka and Watanabe, 2016; Sakai et al., 2013; Endo, 2019). To address these imbalances, the JRMP relies on cap-based regulation of urban residency positions. These caps are set in two steps. First, the total number of positions nationwide is determined as a multiple of the number of graduating medical students. This multiplier was approximately 1.22 in 2015 and was scheduled to decline gradually to 1.05 by 2025. Second, the nationwide total is then allocated across prefectures according to a formula that incorporates population, medical school enrollment capacity, the number of practicing physicians, physician density, geographic conditions, and the population living on isolated or remote islands. The formula is designed to limit residency positions more tightly in urban prefectures while preserving relatively more positions in underserved areas. The rationale is that reducing positions in urban areas will redirect applicants toward residency positions in underserved areas. Because physicians are more likely to practice near where they train, policymakers expect training placements in underserved areas to improve the long-run geographic distribution of physicians.

The data, however, point to important limitations of this cap-based policy. As the caps tightened, the share of applicants left unmatched in the centralized JRMP round increased substantially, from 4.7 percent to 9.5 percent. At the same time, the geographic redistribution of matches was uneven. Figure 2 shows that gains in match shares were concentrated in intermediate prefectures near major urban areas, such as those near Tokyo and Kanagawa, whereas rural prefectures experienced little comparable increase. These patterns suggest that the JRMP's cap-based policy is a blunt instrument for correcting geographic imbalances, motivating our quantitative comparison with more targeted price-based alternatives.

[3]Unmatched applicants may individually contact hospitals with vacancies, reapply through the matching process in the subsequent year, or forgo pursuing a career as a physician. Appendix A.1 describes the typical timeline of the JRMP matching process and provides the definition of "unmatched applicants."

2.2 Data Sources

We construct school–hospital match patterns from individual records in the *Physician Registration Report*. Specifically, for each medical school–hospital pair, we count the number of individuals who graduated from the medical school and trained at the hospital. This administrative data source contains information on holders of medical licenses, including residents, and allows us to identify each individual's graduating medical school and training hospital. We then aggregate the individual-level records to obtain annual pair-level match counts between medical schools and hospitals.[4] We observe matches and hospital characteristics from 2016 through 2019. Because the specification includes lagged school–hospital matches, the estimation sample covers 2017–2019, with 2016 used to construct the first lag.

For residency program characteristics, our first data source is the JRMP website. The JRMP provides information on hospital names, operating authorities, residency program offerings, the number of positions in each program, and the number of students graduating from each medical school. We collected the JRMP data for the years 2004 to 2024. In particular, the number of residency positions and the number of graduating students are used in our counterfactual simulations to study the historical effects of caps on residency positions.

Our second data source for residency program characteristics is the set of official residency-related disclosure forms published by the Ministry of Health, Labour and Welfare. These forms provide information on resident salaries and the number of beds at each hospital offering a residency program. We use, in particular, Form 4, "Notification of Addition/Revision to the Training Program," and Form 7, "Annual Report."[5] We collected these documents for the years 2016 to 2019. Because some salary information was not disclosed by the government, we supplement the missing observations using the most recent information available on hospital webpages.[6]

[4]The underlying records are based on the Ministry of Health, Labour and Welfare's physician reporting system (*iryo jujisha todokede system*); see the official description page: `https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryou/iryou/iryojujisha-todokede-sys.html`. The microdata used in this study were obtained through a formal information-disclosure request to the relevant authority.

[5]See the MHLW list of submission forms: `https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000115595.html`. We obtained the documents through formal information-disclosure requests and manually digitized the salary information. The requests were submitted to all eight Regional Bureaus of Health and Welfare, which are the competent authorities for these documents.

[6]Resident salaries in Japan appear to be highly stable over time. Moreover, hospitals are required to report any changes in resident salaries to the JRMP authorities, and such changes

Table 1. Summary Statistics of Medical-School and Hospital Covariates

	Category	Count	Mean	Std	Min	Max
Panel A. Medical Schools						
T-score	Total	78	65.89	2.58	62	74
	Non-urban	46	65.21	1.86	62	69
	Urban	32	66.88	3.14	63	74
	Public	51	66.38	2.66	63	74
	Private	27	64.96	2.17	62	72
Panel B. Hospitals						
Number of beds	Total	1032	437.89	179.19	36	1379
	Non-urban	684	423.41	154.09	36	1195
	Urban	348	466.36	217.81	38	1379
	Non-university	911	412.80	146.35	36	1097
	University	121	626.82	269.59	295	1379
Monthly Salary ($\times$100,000 JPY)	Total	1032	3.763	0.950	1.800	8.550
	Non-urban	684	3.933	0.945	1.800	8.550
	Urban	348	3.431	0.869	1.900	6.847
	Non-university	911	3.876	0.934	2.225	8.550
	University	121	2.915	0.563	1.800	4.720

* All summary statistics are computed from the cross-section in 2019.

For additional medical-school characteristics, we use the national medical licensing examination pass rate and each school's entrance-exam T-score. T-scores are widely used in Japan as an indicator of university entrance-exam difficulty, with higher scores indicating more selective universities. We use the most recent T-scores available for our estimation.[7]

2.3 Descriptive Statistics

Table 1 reports descriptive statistics for medical-school and hospital characteristics. Panel A summarizes medical-school variables. Schools are classified as public or private, with 51 public and 27 private medical schools in the sample. Higher T-scores indicate greater medical-school entrance-exam selectivity. Public medical schools have higher average T-scores than private medical schools. In our sample, there are 46 non-urban schools and 32 urban schools. Medical schools are more densely concentrated in urban areas, and T-scores are also higher there.

Panel B summarizes hospital-side characteristics. Hospitals are classified as

are subject to additional review. These institutional features suggest that using the most recent webpage information is unlikely to introduce substantial measurement error for the target years.

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

either university-affiliated or non-university-affiliated, and 88.3% are non-university-affiliated. University-affiliated hospitals tend to have more beds on average, consistent with their role as regional flagship institutions. In addition, 33.7% of hospitals are located in urban areas. As with medical schools, hospitals are more densely concentrated in urban areas, and urban hospitals also tend to be larger.

We examine salary variation in the JRMP market. The final rows of Panel B indicate substantial dispersion in resident salaries across Japan. The average monthly salary is 376.3 thousand JPY, with a standard deviation of 95.0 thousand JPY. This corresponds to an average annual salary of approximately \$29,536 and a standard deviation of \$7,456.[8] For comparison, Table 1 of Agarwal (2015) reports a mean salary of \$47,331 and a standard deviation of \$2,953 for comparable U.S. medical interns. While mean salaries are higher in the United States, salary dispersion is much larger in Japan. This variation also reflects regional and institutional pay differentials. On average, non-urban hospitals offer significantly higher monthly wages than urban hospitals, by 50,200 JPY (about \$328). By contrast, university-affiliated hospitals pay significantly less, by 96,100 JPY per month (about \$629), on average.

2.4 Empirical Matching Patterns

Figure 3 shows the empirical matching patterns in the 2019 JRMP market. The left side represents medical-school categories defined by location (urban vs. non-urban) and ownership (public vs. private). The right side represents hospital categories defined by location (urban vs. non-urban). We also include unmatched categories on both sides: *Unfilled Positions* and *Unmatched Applicants*. The width of each band between the two sides indicates the volume of realized matches between category pairs.

A few quantitative patterns stand out. First, matching is strongly segmented by geography. Among applicants from non-urban public schools, 67.7% of matches are with non-urban hospitals; among applicants from urban public and urban private schools, 62.9% and 56.3% of matches, respectively, are with urban hospitals. Second, the unmatched-applicant rates differ substantially across school categories: non-urban/private has the highest unmatched-applicant rate (22.7%), followed by urban/private (15.9%), whereas non-urban/public and urban/public are lower at 10.6% and 10.3%. Third, on the hospital side, positions at urban hospitals are filled mainly by applicants from urban schools (about 51.9% in total

[8] Dollar conversions use 1 USD = 152.887 JPY (February 12, 2026).

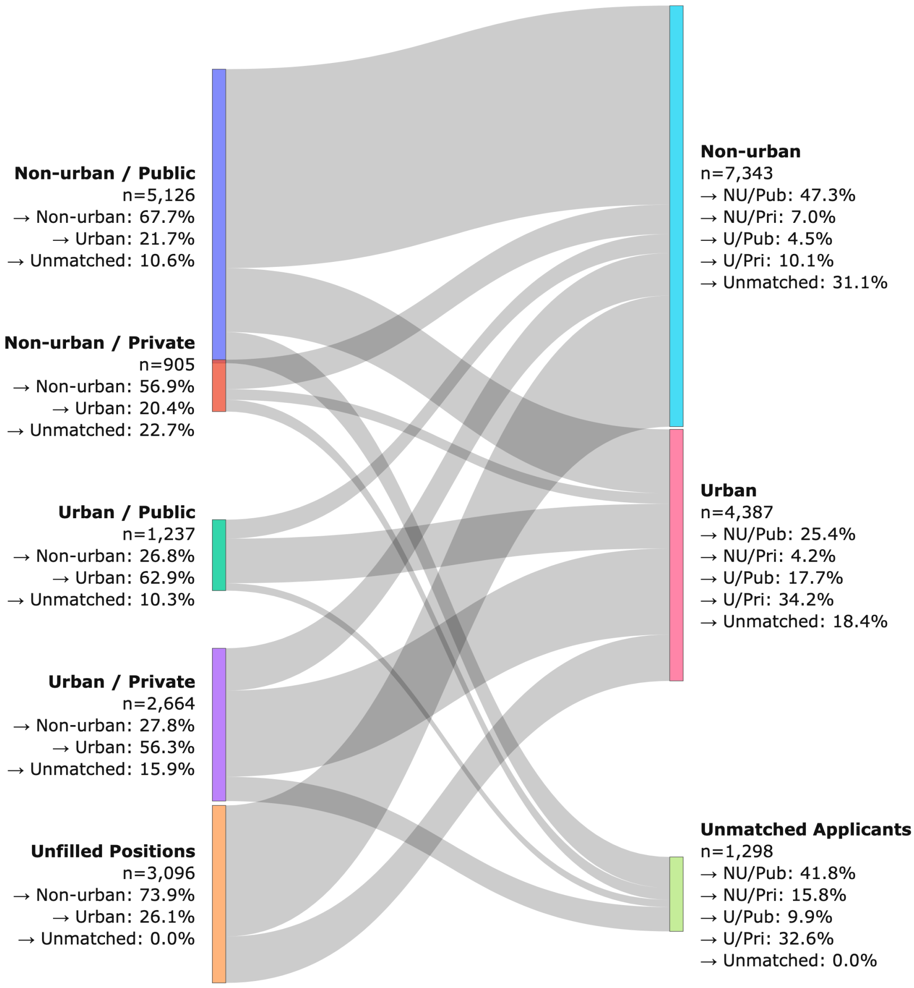


Figure 3. School–hospital matching flows with unfilled positions and unmatched applicants.

Notes: Medical schools are grouped by location and ownership, and hospitals by location. Band widths represent realized match counts. Side-label percentages are within-category shares, including unmatched outcomes.

from urban/public and urban/private), while positions at non-urban hospitals are filled mainly by applicants from non-urban public schools (47.3%). Finally, capacity imbalance is nontrivial: the number of unfilled positions (3,096) is much larger than the number of unmatched applicants (1,298), so unfilled positions outnumber unmatched applicants in this market.[9]

We run a series of regressions to study empirical matching patterns, following Agarwal (2015). Because individual-level matching data are unavailable, we first compute weighted averages of matched partners' characteristics for each medical

[9]This JRMP market is stable over time. Table O.3.1 in Online Appendix O.3.1 reports key summary statistics for other years and shows no meaningful time-series differences. In particular, entry and exit are limited: only three hospitals entered or exited between 2018 and 2019.

school and hospital. Specifically, for each hospital, we compute the match-count-weighted average characteristics of the graduating schools of its matched residents; analogously, for each medical school, we compute the match-count-weighted average characteristics of the hospitals offering positions filled by applicants from that school. We then regress these aggregated partner characteristics on the characteristics of the hospitals and medical schools. We also regress the unmatched-applicant rate by graduating medical school and the unfilled-position rate of each hospital on their respective characteristics to examine which school and hospital characteristics predict unmatched applicants and unfilled positions in this market. Standard errors are clustered at the unit level: at the hospital level for hospital-side regressions and at the graduating-medical-school level for school-side regressions.

From the hospital perspective, consistent with Figure 3, urban hospitals are more likely to match with applicants from private and urban schools. They also exhibit lower unfilled-position rates, suggesting stronger demand for urban hospitals in this market. Higher salaries are associated with matches from higher-T-score schools and with a higher share of public-university matches, whereas the coefficient on the share of urban-school matches is negative and statistically insignificant. Taken together, these results suggest that salary premia mainly help hospitals attract applicants from more selective, non-urban public schools. The positive association between salaries and unfilled-position rates is also consistent with salaries reflecting strategic behavior in hospitals' recruiting policies. Table O.3.3 reports the corresponding regression results.

From the school perspective, although public schools have higher average T-scores, applicants from those schools are no more likely to match with urban hospitals; however, they receive higher wages on average. This pattern again suggests that higher salaries may be used by non-urban hospitals to attract applicants from more selective public schools. Applicants from urban schools, by contrast, appear to have a stronger preference for matching in urban areas. Table O.3.4 reports the corresponding regression results.

Overall, applicants from public schools are more likely to secure a match, which is consistent with the common view that public schools are more selective and that applicants from public schools are preferred in the matching process. In contrast, the T-score is not statistically significant for most outcomes, except distance. This pattern highlights an important nuance of the T-score: although it captures entrance-exam selectivity, it is not necessarily a direct measure of applicant quality.

3 Model

Matching market with regional constraints We consider a two-sided matching market. Let I denote the set of applicants and J the set of positions owned by hospitals. Each applicant $i \in I$ can be matched with at most one position $j \in J$, and each position can accommodate at most one applicant. Unmatched applicants and positions are assigned to outside options j_0 and i_0, respectively. 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 applicant i is matched with position j. A matching d is *feasible* if each applicant is matched with at most one position and each position is filled by at most one applicant.

There are $L \in \mathbb{Z}_+$ regions, denoted by $Z = \{z_1, z_2, \ldots, z_L\}$, with each position 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 position j belongs. Each region z has a floor $\underline{o}_z$ and a cap $\bar{o}_z$, with $0 \leq \underline{o}_z < \bar{o}_z$. We say a feasible matching d satisfies *regional constraints* if it satisfies the regional floors and caps: $\sum_{i \in I} \sum_{j \in z} d_{ij} \in [\underline{o}_z, \bar{o}_z]$ for each $z \in Z$, where, with a slight abuse of notation, $j \in z$ means $z(j) = z$. We assume that $\underline{o}_z < |\{j \colon j \in z\}|$ for each $z \in Z$ and $\sum_{z \in Z} \underline{o}_z < |I|$.[10]

Stable outcome with policy intervention Agents form a stable outcome à la Shapley and Shubik (1971). Without policy intervention, the realized matching may not meet the regional constraints. For example, a floor for an underserved region may reflect the social value of physician access there, which the matched pair does not fully internalize.[11] We consider two classes of policies that the policymaker can use to satisfy the regional constraints: a cap-based policy makes some positions unavailable for matching; a tax-and-subsidy policy adjusts the payoff available to matched participants through taxes or subsidies.

We first define equilibrium under a tax-and-subsidy policy. When applicant i and position j are matched, they generate *(individual-level) baseline joint surplus*

[10] This assumption suffices to guarantee the existence of at least one feasible matching satisfying the regional constraints. It also ensures that the refined Slater condition (see Section 5.2.3 of Boyd and Vandenberghe (2004)) holds for the optimization problem analyzed in Section 4, thereby guaranteeing strong duality.

[11] In Japan, prices for medical services are set by a uniform national fee schedule with no regional adjustment, and providers cannot charge above the scheduled prices. Hospitals in underserved regions therefore cannot raise prices to reflect local physician scarcity. Moreover, benefits such as access to nearby emergency care accrue to the region's population without passing through hospital revenue.

$\Phi_{ij} \in \mathbb{R}$. 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. Under the tax-and-subsidy policy $w = (w_z)_{z \in Z}$, each matched pair divides the *policy-adjusted joint surplus* $\Phi_{ij} - w_{z(j)}$ instead of the baseline joint surplus.

The stable outcome under a tax-and-subsidy policy is defined as follows:

Definition 1 (Stable outcome). *Given the matching market* (I, J, Z, z, Φ)*, a profile* $(d, (u, v))$ *of feasible matching* d *and equilibrium payoffs* (u, v) *forms a* stable outcome *under the tax-and-subsidy 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* $d_{ij} = 1$*.*

When $w_z = 0$ for all z, stable outcomes in Definition 1 coincide with those in the standard definition. This equilibrium concept is often used to study decentralized matching markets such as marriage and labor markets.[12] It is also natural in our application when the residency matching process is viewed as a whole.[13]

Instead of tax-and-subsidy policies, the policymaker may employ *cap-based policies*, which restrict the set of available positions. Formally, a *cap-based policy* is specified by a subset $J' \subseteq J$, representing the positions that remain available after the policy is imposed. Given such a policy, let $\Phi' := (\Phi_{ij})_{i \in I, j \in J'}$ denote the restriction of the surplus matrix to applicants and the available positions. Under a cap-based policy J', agents form a stable outcome in the induced matching market $(I, J', Z, z|_{J'}, \Phi')$. Such quantity-based instruments are motivated by the objective of redirecting applicants toward understaffed regions by limiting capacity in high-demand areas. In this paper, we assume that tax-and-subsidy policies are not jointly applied with a cap-based policies.

The *social surplus* under matching d is defined as the sum of baseline joint surpluses from matched pairs and outside-option payoffs from unmatched applicants and positions:

$$S(\Phi, d) := \sum_{(i,j) \in I \times J} \Phi_{ij} d_{ij} + \sum_{i \in I} \left(1 - \sum_{j \in J} d_{ij}\right) \Phi_{i,j_0} + \sum_{j \in J} \left(1 - \sum_{i \in I} d_{ij}\right) \Phi_{i_0,j}.$$

[12] See Chiappori (2017) for a textbook reference.

[13] Online Appendix O.1 provides a simple formalization of this interpretation.

In the current quasi-linear environment, the social surplus coincides with the sum of agents' payoffs and government's net revenue. When the surplus matrix Φ is clear from context, we write $S(d)$ in place of $S(\Phi, d)$. When evaluating the social surplus of a cap-based policy, we treat the resulting matching in the restricted market as matching in the original market.[14]

Observed categories Let X denote the finite set of observable characteristics, or *categories*, of applicants. Each applicant $i \in I$ has a category $x(i) \in X$. Similarly, let Y denote the finite set of observable characteristics of positions, with each position $j \in J$ having a category $y(j) \in Y$. Although agents in the same category are indistinguishable to the policymaker, there can be *unobservable heterogeneity*: applicants in the same category x or positions in the same category 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 categories representing the outside options i_0 and j_0, respectively, and let $X_0 = X \cup \{x_0\}$ and $Y_0 = Y \cup \{y_0\}$ include these outside options. The set of all category pairs is denoted by $C = X_0 \times Y_0 \setminus \{(x_0, y_0)\}$. We assume each position category $y \in Y$ belongs to a unique region, denoted by $z(y) \in Z$. Let n_x be the number of applicants in category x, and m_y be the number of positions in category y. We assume that $n_x > 0$ for each $x \in X$ and $m_y > 0$ for each $y \in Y$.

Let μ_{xy} denote the number of matches between applicants in category x and positions in category y, defined as $\mu_{xy} = \sum_{i \in x} \sum_{j \in y} d_{ij}$. Let $\mu_{x,y_0} := n_x - \sum_{y \in Y} \mu_{xy}$ and $\mu_{x_0,y} := m_y - \sum_{x \in X} \mu_{xy}$. An *aggregate-level matching* $\mu = (\mu_{xy})_{(x,y) \in C}$ is said to be *feasible* if it satisfies the population constraints $\sum_{y \in Y_0} \mu_{xy} = n_x$ for each $x \in X$ and $\sum_{x \in X_0} \mu_{xy} = m_y$ for each $y \in Y$. 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 $z \in Z$.

The category and region partitions admit several interpretations; z need only define a coarser grouping of the categories y. In our application, residency matching, a position category y corresponds to a hospital and a region z corresponds to a geographic unit such as a prefecture. In occupational licensing, y could correspond to a specific occupation, such as registered nurse, and z to a broader occupational category, such as healthcare. In school choice, y could correspond to an individual school and z to a school district.

[14] Formally, given equilibrium matching $d' = (d'_{ij})_{i \in I, j \in J'}$ under J', with a slight abuse of notation, we also write d' for its extension to $I \times J$, defined by setting $d'_{ij} = 0$ for all $j \in J \setminus J'$. The social surplus generated by the cap-based policy is $S(\Phi, d')$.

4 Theoretical Results

This section derives the main theoretical results. We begin with a complete-information benchmark in which the policymaker observes all individual-level baseline joint surpluses $(\Phi_{ij})_{ij}$. In this benchmark, the policymaker can compute an optimal tax-and-subsidy policy that implements the surplus-maximizing matching subject to the regional constraints. We then introduce unobserved heterogeneity. In this setting, individual-level baseline joint surpluses are not observed, so the complete-information policy is not directly computable. We show how the policymaker can instead use aggregate-level market primitives, learned from aggregate-level match data, to compute an aggregate-optimal tax-and-subsidy policy. Finally, we show that the aggregate-optimal policy and its induced outcomes are asymptotically equivalent to their complete-information counterparts in large markets. All proofs are in either the Appendix or the Online Appendix.

4.1 Complete-Information Benchmark: Optimal Tax-and-Subsidy Policy

We begin with a complete-information benchmark in which the policymaker observes the individual-level baseline joint surpluses $(\Phi_{ij})_{ij}$ for every possible applicant-position pair. We show that, if this information were available, the policymaker could compute the tax-and-subsidy policy that implements the surplus-maximizing matching subject to the regional constraints.

Recall the social surplus function S defined in Section 3. Consider the following *individual-level planner's problem*, in which the policymaker maximizes social surplus subject to regional constraints:

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

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

The following theorem shows that there exists a tax-and-subsidy policy w^*

that implements d^* as a stable outcome.[15] We call w^* the *optimal tax-and-subsidy policy*. Appendix B.1 further shows that w^* is characterized by a set of Lagrange multipliers for the regional constraints, and the policymaker can compute them by solving the dual problem of (P_0).

Theorem 1. *Fix any market (I, J, Z, z, Φ) and regional constraints $(\underline{o}_z, \bar{o}_z)_z$. There exists a tax-and-subsidy policy w^* and a payoff profile (u, v) such that $(d^*, (u, v))$ is a stable outcome under w^*.*

One implication is that, within our framework, the optimal tax-and-subsidy policy weakly dominates any cap-based policy in terms of social surplus: no cap-based policy that satisfies the same regional constraints can generate greater social surplus.

Corollary 1. *Fix any market and regional constraints. Suppose that the stable matching d' induced by a cap-based policy $J' \subseteq J$ satisfies the regional constraints. Then the optimal tax-and-subsidy policy weakly dominates that cap-based policy in social surplus, i.e.,*

$$S(d^*) \geq S(d').$$

Policy-loss decomposition The preceding result also provides a useful benchmark for decomposing the total policy loss from a cap-based policy. Suppose that the policymaker uses a cap-based policy to satisfy the regional constraints, with induced equilibrium matching d'.

Now consider the *no-cap-tightening benchmark*, under which all positions in J are available and agents form a stable outcome with no policy intervention. It is known in the literature that the equilibrium matching under this benchmark, denoted by d^0, maximizes the social surplus among all feasible matchings when the regional constraints are absent.[16] Since d^* maximizes social surplus among matchings that satisfy the regional constraints, and since d' is one such matching, we have

$$S(d^0) \geq S(d^*) \geq S(d').$$

[15] Theorem 1 yields weak implementation of d^*: multiple stable outcomes may arise under w^* and may deliver different social surpluses. In Online Appendix O.2, we show that, under regularity conditions, all normalized stable matchings under w^* converge to the same limit matching. Hence, any surplus discrepancy across stable outcomes is asymptotically negligible in per-capita terms.

[16] See Roth and Sotomayor (1990) for a textbook reference.

Thus, the total policy loss under the cap-based policy relative to the no-cap-tightening benchmark can be decomposed as

$$\underbrace{S(d^0) - S(d')}_{\text{total policy loss}} = \underbrace{\left(S(d^0) - S(d^*)\right)}_{\text{distributional-target cost}} + \underbrace{\left(S(d^*) - S(d')\right)}_{\text{implementation cost}}. \tag{1}$$

The first term is the distributional-target cost: the surplus loss from imposing the regional constraints themselves. The second term is the implementation cost: the additional loss from satisfying those constraints through the cap-based policy rather than through the surplus-maximizing price-based policy. In theory, both components on the RHS of (1) can take any nonnegative values; the size of the total policy loss and its components is an empirical question.

Remark 1 (Pair-specific vs. region-specific tax-and-subsidy policies)**.** When defining tax-and-subsidy policies in Section 3, we restrict attention to policies that are uniform within each region. This restriction is without loss for surplus maximization. Even if the policymaker were allowed to set pair-specific taxes and subsidies, the proof of Theorem 1 constructs an optimal policy that applies a single tax or subsidy w_z^* uniformly to all matches in each region z; this simple structure facilitates implementation. This uniformity reflects the structure of the planner's problem: because the constraints depend only on the total number of matches in each region, their Lagrange multipliers yield taxes and subsidies that vary by region rather than by pair.

4.2 Aggregate-Level Market with Unobserved Heterogeneity

With unobserved heterogeneity, the policymaker no longer observes the individual-level baseline joint surpluses $(\Phi_{ij})_{ij}$. The optimal tax-and-subsidy policy w^* is therefore not directly computable. This subsection shows how, under certain structural assumptions, the policymaker can approximate the optimal tax-and-subsidy policy using aggregate-level matching data. We first adapt the framework of Galichon and Salanié (2022), which maps the individual-level objects from Section 3, namely baseline joint surpluses $(\Phi_{ij})_{ij}$, match indicators $(d_{ij})_{ij}$, and payoffs $(u_i)_i$ and $(v_j)_j$, to aggregate-level counterparts. We then define a primal-dual pair of optimization problems whose solution yields the *aggregate-optimal tax-and-subsidy policy* w^{AE}. These problems take as input only aggregate-level market primitives, which can be identified from aggregate-level matching data. The large-market result in Section 4.3 justifies this construction: as markets grow large, w^{AE} is asymptotically equivalent to the complete-information policy w^*.

4.2.1 Discrete-choice representation and large-market payoff functions

Discrete choice representation We first connect the individual-level objects in our matching model to aggregate-level objects. The key structural assumption is that each individual-level baseline joint surplus can be decomposed into a category-level component and two idiosyncratic components. This decomposition allows the stable matching problem to be represented as a bilateral discrete-choice problem.

For any pair (i, j) with applicant i in category x and position j in category y, we assume that the individual-level baseline joint surplus has the following structure:

Assumption 1 (Additive Separability)**.** *There exists a matrix $(\Phi_{xy})_{(x,y)\in C}$ such that:*

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

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

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

$$\Phi_{x,y_0} = \Phi_{x_0,y} = 0, \quad \Phi_{i,j_0} = \varepsilon_{i,y_0}, \quad \Phi_{i_0,j} = \eta_{x_0,j}.$$

Conditional on observable categories, the error terms $(\varepsilon_i)_i$ and $(\eta_j)_j$ are independent across all applicants and positions.

Here, Φ_{xy} is the *aggregate-level baseline joint surplus* and ε_{iy} and η_{xj} capture unobserved heterogeneity, with finite first moments.[17] For each category x and every applicant $i \in x$, the vector $\varepsilon_i = (\varepsilon_{iy})_{y\in Y_0}$ is drawn from a distribution $P_x \in \Delta(\mathbb{R}^{|Y|+1})$. Similarly, for each category y and every position $j \in y$, the vector $\eta_j = (\eta_{xj})_{x\in X_0}$ is drawn from a distribution $Q_y \in \Delta(\mathbb{R}^{|X|+1})$.

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

$$U_{xy} := \min_{i\,:\,x(i)=x} \{u_i - \varepsilon_{iy}\}, \quad V_{xy} := \min_{j\,:\,y(j)=y} \{v_j - \eta_{xj}\}. \tag{2}$$

[17] This can be slightly relaxed: it suffices to assume $\max_{y\in Y_0} \mathbb{E}[|\varepsilon_{iy}| \mid i \in x] < \infty$ for each $x \in X$ and $\max_{x\in X_0} \mathbb{E}[|\eta_{xj}| \mid j \in y] < \infty$ for each $y \in Y$. The assumption is used to guarantee that $G(U)$ and $H(V)$ defined below are finite for all U and V.

Note that we have $U_{x,y_0} = V_{x_0,y} := 0$ by construction. The following lemma shows that a stable outcome admits a bilateral discrete-choice representation: each agent's individual-level equilibrium payoff equals the value of a discrete-choice problem whose systematic component is given by the aggregate-level utilities U_{xy} and V_{xy}.

Lemma 1 (Galichon and Salanié (2022))**.** *Let (u, v) be a payoff profile corresponding to a stable outcome. Under Assumption 1, for any applicant $i \in I$ and any position $j \in J$, we have*

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

Large-market payoffs By Lemma 1, the applicant-side equilibrium payoff 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 applicants in category 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, $\mathbb{E}_{\varepsilon_i \sim P_x} [\max_{y \in Y_0} \{U_{xy} + \varepsilon_{iy}\}]$.[18] This motivates the following approximation to the applicant-side equilibrium payoff:

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

Similarly, if the number of positions in category y, denoted by m_y, is large, the position-side equilibrium payoff is approximated by

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

Under the assumption that the cumulative distribution functions of the error terms are smooth (see Assumption 2 below), the Williams-Daly-Zachary theorem (McFadden, 1981) implies that G is continuously differentiable and satisfies, for each

[18]This argument is formalized in Online Appendix O.2.

$x \in X$, $y \in Y_0$, and $i \in x$,

$$\frac{\partial G}{\partial U_{xy}}(U) = n_x \Pr\left(\{y\} = \arg\max_{y' \in Y_0} \{U_{xy'} + \varepsilon_{iy'}\} \right). \tag{3}$$

For sufficiently large n_x, the RHS of (3) provides a good approximation of the number of category-wise matches μ_{xy}. Analogous results for H also hold.

Assumption 2 (Smooth Distributions)**.** *For each x and y, the distributions P_x and Q_y are absolutely continuous with respect to Lebesgue measure.*

Lastly, we introduce an additional technical assumption on the error term distributions that guarantees the strict monotonicity and strict convexity of G and H.

Assumption 3 (Full Support)**.** *For each x and y,* $\mathrm{supp}(P_x) = \mathbb{R}^{|Y_0|}$ *and* $\mathrm{supp}(Q_y) = \mathbb{R}^{|X_0|}$.

4.2.2 Aggregate-Level Planner's Problem

The *aggregate-level market primitives* are the tuple

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

where $n \coloneqq (n_x)_x$, $m \coloneqq (m_y)_y$, $\Phi \coloneqq (\Phi_{xy})_{xy}$, $P \coloneqq (P_x)_x$, and $Q \coloneqq (Q_y)_y$. We consider a policymaker who knows $\mathcal{M}$. In applications, (X, Y, n, m, Z, z) is observed directly. Section 5 describes how, given P and Q, we estimate Φ from aggregate-level matching data.

Definition 2 (Aggregate-Level Planner's Problem)**.** *Given aggregate-level market primitives $\mathcal{M}$ and regional constraints $(\underline{o}_z, \bar{o}_z)_{z \in Z}$, the* aggregate-level planner's problem *is the following primal-dual pair.*

The primal problem is

$$(P) \quad \left| \begin{array}{lll} \max\limits_{\mu \in \mathbb{R}^C_+} & \sum\limits_{(x,y) \in C} \mu_{xy} \Phi_{xy} + \mathcal{E}(\mu) & \\ \text{s.t.} & \sum\limits_{y \in Y_0} \mu_{xy} = n_x & \forall x \in X, \\ & \sum\limits_{x \in X_0} \mu_{xy} = m_y & \forall y \in Y, \\ & \underline{o}_z \leq \sum\limits_{y \in z} \sum\limits_{x \in X} \mu_{xy} \leq \bar{o}_z & \forall z \in Z, \end{array} \right.$$

where $\mathcal{E}(\mu) := -G^*(\mu) - H^*(\mu)$*, and* G^* *and* H^* *denote the Legendre-Fenchel transforms of* G *and* H*, respectively.*

The dual problem is

$$(D)\quad \left|\begin{array}{lll} \min\limits_{U,V,\bar{w},\underline{w}} & G(U) + H(V) + \sum\limits_{z\in Z} \bar{o}_z \bar{w}_z - \sum\limits_{z\in Z} \underline{o}_z \underline{w}_z & \\ \text{s.t.} & U_{xy} + V_{xy} \geq \Phi_{xy} - \bar{w}_{z(y)} + \underline{w}_{z(y)} & \forall x \in X, \forall y \in Y, \\ & U_{x,y_0} = 0 & \forall x \in X, \\ & V_{x_0,y} = 0 & \forall y \in Y, \\ & \bar{w}_z \geq 0,\ \underline{w}_z \geq 0 & \forall z \in Z. \end{array}\right.$$

A solution to the aggregate-level planner's problem is a tuple $(\mu, U, V, \bar{w}, \underline{w})$ *such that* μ *solves* (P) *and* $(U, V, \bar{w}, \underline{w})$ *solves* (D)*.*

The primal problem is the aggregate-level counterpart of the social-surplus maximization problem (P_0). The term $\mathcal{E}(\mu)$ captures the contribution of unobserved heterogeneity to expected social surplus. Intuitively, it is larger for aggregate-level matching patterns that align more closely with agents' idiosyncratic match values, and smaller for patterns that are harder to implement given those idiosyncratic preferences.

The primal problem admits an optimal solution because its objective function is continuous and its feasible set is compact and nonempty. By strong duality,[19] the dual problem (D) attains the same optimal value. Moreover, both (P) and (D) possess unique solutions, as we show below.

Theorem 2. *Assume that Assumptions 1–3 hold. Fix aggregate-level market primitives* $\mathcal{M}$ *and regional constraints* $(\underline{o}_z, \bar{o}_z)_{z\in Z}$ *for which the feasible set of* (P) *is nonempty. Then,* (P) *and* (D) *admit unique solutions, denoted by* μ^{AE} *and* $(U^{AE}, V^{AE}, \bar{w}^{AE}, \underline{w}^{AE})$*.*

Given the solution $\bar{w}^{AE}$ and $\underline{w}^{AE}$, we define the corresponding tax-and-subsidy policy by

$$w_z^{AE} := \bar{w}_z^{AE} - \underline{w}_z^{AE} \quad (z \in Z).$$

We call this w^{AE} the *aggregate-optimal tax-and-subsidy policy* given $\mathcal{M}$ and $(\underline{o}_z, \bar{o}_z)_{z\in Z}$. Note that, by the complementary slackness condition, at most one of $\bar{w}_z^{AE}$ and $\underline{w}_z^{AE}$ can be nonzero for each z.

[19] The constraints of the primal problem satisfy the refined Slater condition (see Section 5.2.3 of Boyd and Vandenberghe (2004)).

Aggregate equilibrium under a fixed policy The same aggregate-level representation can also be used to simulate outcomes under alternative primitives and policies. For example, in the counterfactual analysis in Section 7, cap-based policies are represented by the corresponding changes in aggregate-level market primitives with $w_z := 0$ for all z. Given aggregate-level market primitives $\mathcal{M}$ and a tax-and-subsidy policy w, the corresponding aggregate-level outcome $(\mu(w), U(w), V(w))$ can be obtained by solving the following pair of optimization problems.

$$
(P_w) \quad \left[\begin{array}{ll} \max\limits_{\mu \in \mathbb{R}_+^C} & \sum\limits_{x \in X} \sum\limits_{y \in Y} \mu_{xy} \big(\Phi_{xy} - w_{z(y)}\big) + \mathcal{E}(\mu) \\ \text{s.t.} & \sum\limits_{y \in Y_0} \mu_{xy} = n_x, \quad \forall x \in X, \\ & \sum\limits_{x \in X_0} \mu_{xy} = m_y, \quad \forall y \in Y. \end{array}\right.
\qquad
(D_w) \quad \left[\begin{array}{ll} \min\limits_{U,V} & G(U) + H(V) \\ \text{s.t.} & U_{xy} + V_{xy} \geq \Phi_{xy} - w_{z(y)}, \\ & \forall x \in X,\ y \in Y, \\ & U_{x,y_0} = 0, \quad \forall x \in X, \\ & V_{x_0,y} = 0, \quad \forall y \in Y. \end{array}\right.
$$

The tuple $(\mu(w), U(w), V(w))$ is referred to as an *aggregate equilibrium (AE)* under w. When w is aggregate-optimal, that is, when $w = w^{AE}$, the equilibrium is termed the *efficient aggregate equilibrium (EAE)*.

4.3 Large-Market Approximation

The complete-information benchmark in Section 4.1 computes the optimal tax-and-subsidy policy w^* from the realized individual-level surpluses. The aggregate-level construction in Section 4.2 instead computes w^{AE} using aggregate-level primitives. In large markets, these two policies become asymptotically equivalent: the aggregate-optimal tax-and-subsidy policy w^{AE} converges to the complete-information optimal tax-and-subsidy policy w^* as the market size grows, and the associated match fractions across category pairs and per-capita social surplus converge as well. Solving the aggregate-level planner's problem therefore provides a tractable and asymptotically valid method for computing the optimal tax-and-subsidy policy and evaluating surpluses across policies. Online Appendix O.2 states and proves the formal large-market result.

5 Empirical Strategy

The empirical procedure has two steps. First, aggregate medical-school–hospital match counts identify applicant- and hospital-side systematic utilities in model-numeraire units. Second, salary variation maps those utilities into money-equivalent units. The resulting joint-surplus estimates are used to simulate the counterfactual policies in Section 7.

We begin by mapping the primitives and equilibrium objects in the model to their empirical counterparts. An applicant is denoted by i, and a hospital position by j. Applicants are indexed by their graduating medical schools, and positions by the hospitals that offer them: $s(i)$ denotes applicant i's graduating medical school, and $h(j)$ denotes the hospital offering position j. Let Z denote the set of regions and $z(h)$ denote the region to which hospital h belongs.

For each annual market, let Φ_{sh} denote the aggregate-level baseline joint surplus. The individual-level baseline 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 available data consist of the aggregate-level matching $(\mu_{sh})_{s,h}$, which is the number of matches between medical school s and hospital h.

As described in Section 4.2.2, knowledge of aggregate-level baseline joint surplus Φ_{sh}, which is the sum of the aggregate-level utilities U_{sh} and V_{sh}, is sufficient to simulate counterfactual equilibria and the optimal tax-and-subsidy policy under any regional constraints. Moreover, following Galichon and Salanié (2022), we can identify the aggregate-level utilities from the observed aggregate-level matching given the distributions of the idiosyncratic error terms.

However, to evaluate social surplus under alternative policies in monetary terms, and the magnitude of the taxes and subsidies required to implement the optimal policy, it is necessary to express the aggregate-level utilities in monetary units. To this end, Section 5.1 imposes additional structure on the individual-level baseline joint surplus, which allows us to define individual-level transfers and link them to aggregate-level transfers. Section 5.2 then parameterizes the primitives introduced in Section 5.1 and derives moment conditions to identify the parameters using monthly salaries. Finally, Section 5.3 details the estimation procedure based on these moment conditions.

5.1 Transfers and Base Utilities

We begin with the standard transfer representation of a matching model with transferable utility. The *base utility* of applicant i when matched with position j, denoted by U_{ij}^{base}, is the utility that i obtains net of the transfer. Similarly, V_{ij}^{base} denotes the base utility of position j when matched with i. The baseline joint surplus generated by the pair is therefore

$$\Phi_{ij} = U_{ij}^{\text{base}} + V_{ij}^{\text{base}}.$$

Fix a stable outcome in the observed market, where taxes and subsidies are zero. For each potential pair (i, j), stability allows us to choose 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, matched-pair equality implies that the two bounds coincide. The equilibrium transfer is then uniquely determined, and the pair's payoffs can be written as

$$u_i = U_{ij}^{\text{base}} + \tau_{ij}, \qquad v_j = V_{ij}^{\text{base}} - \tau_{ij}.$$

The transfer determines how the pair divides its baseline joint surplus without changing its level. When i and j are not matched, the transfer need not be unique; choosing one admissible value for each unmatched pair yields a profile of equilibrium transfers associated with the stable outcome.

The standard TU formulation permits transfers to depend on the identities of both members of a pair, so residents matched to different positions within the same residency program may receive different transfers in the model. This flexibility accords with how compensation for JRMP residents is determined. Although a program's base salary is common, realized compensation may differ: night-duty allowances depend on post-match shift assignments, while housing and commuting allowances depend on individual circumstances. Hospitals also retain post-match discretion over terms such as rotation assignments or access to senior physicians; favorable treatment of a resident imposes an offsetting cost on the hospital. We interpret these discretionary monetary and nonmonetary terms as pair-specific transfers.

To connect these individual-level objects to the aggregate data, we impose additional structure on the two base-utility components. In accordance with additive separability (Assumption 1), we interpret the error terms as unobserved taste shocks and introduce U_{sh}^{base} and V_{sh}^{base} as *aggregate-level base utilities.* These are model primitives to be parameterized below.

Assumption 4. *There exist $(U_{sh}^{\text{base}})_{s,h}$ and $(V_{sh}^{\text{base}})_{s,h}$ such that, for each $i \in s$ and*

$j \in h$, U_{ij}^{base} and V_{ij}^{base} are generated as follows:

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

We also introduce aggregate counterparts of τ_{ij}. For a hospital with at least one realized match, define the *aggregate-level transfer* for hospital h, denoted by ι_h, as the average equilibrium transfer from positions at h to the matched residents occupying those positions:

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

where M_h is the number of matched pairs at hospital h. Note that the observed salary does not fully capture ι_h as mentioned above. This object also aggregates other monetary and nonmonetary terms, including unrecorded allowances and discretionary program terms. Thus, we do not treat ι_h as a perfectly observable object.

5.2 Parametrization

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. 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: $\theta \coloneqq (\beta_U, \beta_V)$.

Suppose for now that ι_h is directly observed for every hospital with $M_h > 0$. The next proposition gives the identity that links the average transfer at a hospital to the aggregate-level utilities. Galichon and Salanié (2022) 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 estimating equations for recovering θ.

Proposition 1. *Suppose Assumptions 1 and 4 hold. Fix a stable outcome in the observed market with zero taxes and subsidies, and let U and V be the aggregate utilities associated with this stable outcome. For any hospital h with $M_h > 0$, we have*

$$\iota_h = \sum_s \omega_{sh}\left(U_{sh} - U_{sh}^{\text{base}}\right) = \sum_s \omega_{sh}\left(V_{sh}^{\text{base}} - V_{sh}\right),$$

where $\omega_{sh} \coloneqq \mu_{sh}/M_h$.

Proof. See Appendix B.4. □

We introduce a measurement model: the monthly salary paid by a hospital to its residents provides a measure of this aggregate-level transfer. The salary paid by a hospital h is denoted by S_h. We assume that the school- and hospital-side aggregate-level transfers have a linear structure with respect to monetary transfers, S_h, i.e.,

$$\begin{aligned} \iota_h &= \gamma_{0,U} + \gamma_{1,U} S_h + \psi_h^U \\ -\iota_h &= \gamma_{0,V} + \gamma_{1,V} S_h + \psi_h^V. \end{aligned} \tag{4}$$

Note that $\gamma_{1,V}$ is expected to be negative because the salary is the amount of money paid to the matched resident by the hospital. We use γ to denote the set of parameters in the measurement model. The terms ψ_h^U and ψ_h^V capture the unobserved components of the transfer in the applicant- and hospital-side measurement equations, respectively. Because these components may be correlated with salary, we use instrumental variables to estimate the measurement model, as described below.

5.3 Estimation Procedure

Our estimation consists of the following two steps: (i) estimate the aggregate-level utilities U_{sh} and V_{sh} for every year, and then (ii) estimate θ and γ using the estimated aggregate-level utilities and the observed salaries.

5.3.1 Estimating Aggregate Utilities

Despite the nonparametric identification results in Galichon and Salanié (2022), we estimate a parametric specification for the aggregate-level utilities. The reason is practical: in the data, some school–hospital pairs have zero realized matches, which makes fully nonparametric estimation unstable in finite samples.

In the first step, we estimate U_{sh} and V_{sh} for each year using the moment-matching estimator proposed in Galichon and Salanié (2022). We implement this step under an i.i.d. type-I extreme-value specification for the idiosyncratic taste shocks. Under this specification, the estimator is exactly equivalent to a Poisson regression with fixed effects.[20]

[20]For completeness, Appendix B.5 derives the equivalence between the moment-matching estimator and a two-way fixed-effects Poisson regression, complementing the discussion in Galichon and Salanié (2023). Note that our estimation target is U_{sh} and V_{sh}. Relative to Galichon and Salanié (2022), we require fixed effects on only one side of the market.

As regressors in the Poisson regressions, we use second- and third-degree polynomial expansions based on X_{sh}^U and X_{sh}^V. Let X_{sh}^{poly} denote the resulting vector of polynomial terms. We thus model the aggregate-level utilities as

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

Let $\hat{\beta}_U^{\text{poly}}$ and $\hat{\beta}_V^{\text{poly}}$ denote the corresponding coefficient estimates, and define the estimated aggregate-level utilities by

$$\hat{U}_{sh} \coloneqq X_{sh}^{\text{poly}\prime} \hat{\beta}_U^{\text{poly}}, \qquad \hat{V}_{sh} \coloneqq X_{sh}^{\text{poly}\prime} \hat{\beta}_V^{\text{poly}}.$$

The covariates X_{sh}^U and X_{sh}^V are constructed from three blocks: school characteristics, hospital characteristics, and pair-specific characteristics. The same set of variables is used in the following second step.

School characteristics. On the school side, we include indicators for whether the medical school is public and whether it is located in an urban prefecture, as well as the number of applicants from the school. In addition, we include two school-level covariates, the T-score and the national examination pass rate.[21] Finally, we include prefecture fixed effects for schools, implemented as a full set of prefecture dummy variables.

Hospital characteristics. On the hospital side, we include indicators for university-affiliated hospitals and government hospitals, as well as an indicator for whether the hospital is located in an urban prefecture. We also include residency-program capacity and the logarithm of the number of beds. We also include prefecture fixed effects for hospitals, implemented as prefecture dummy variables.

Pair characteristics. We further control for several dyadic characteristics of a school–hospital pair (s, h). These include the geographic distance between s and h, an indicator for whether hospital h is affiliated with school s, the number of matches between the same pair in the previous year, and an indicator for whether the school and hospital are located in the same prefecture.

[21] We normalize them by dividing each variable by its sample maximum.

5.3.2 Money-Metric Normalization

The first-step estimates, $\hat{U}_{sh}$ and $\hat{V}_{sh}$, are measured in latent utility units. These estimates are already sufficient for comparing matching outcomes and welfare across policies in model units. The role of the second step is to estimate how these latent units translate into a monetary unit, such as Japanese yen, using observed resident salaries.

Define the weighted average variables as follows:

$$\begin{aligned} Y_h^U &:= \sum_s \omega_{sh} U_{sh}, \quad Y_h^V := \sum_s \omega_{sh} V_{sh}, \\ W_h^U &:= \sum_s \omega_{sh} X_{sh}^{U,\text{base}}, \quad W_h^V := \sum_s \omega_{sh} X_{sh}^{V,\text{base}}. \end{aligned}$$

We use $\hat{Y}_h^U$ and $\hat{Y}_h^V$ to denote the estimates of Y_h^U and Y_h^V made from the estimates of the aggregate-level utilities. Then, by combining with the identity in Proposition 1 and the measurement model (4), we have a linear model to estimate:

$$\begin{aligned} \hat{Y}_h^U &= \gamma_{0,U} + \gamma_{1,U} S_h + W_h^{U\prime} \beta_U + \psi_h^U \\ \hat{Y}_h^V &= \gamma_{0,V} + \gamma_{1,V} S_h + W_h^{V\prime} \beta_V + \psi_h^V. \end{aligned} \tag{5}$$

In practice, this linear model is constructed for each observed year. We pool all available years and estimate the two equations separately for the school side and the hospital side. The constant terms $\gamma_{0,U}$ and $\gamma_{0,V}$ are allowed to vary by prefecture-year pair while the coefficients are fixed.

We estimate (5) by instrumenting salary S_h with functions of nearby hospitals' characteristics following Berry et al. (1995). Note that we use the IV estimates only to determine the salary-to-utility conversion, for which we adopt a deliberately conservative baseline. We consider a large set of candidate IV specifications and retain those with the expected salary sign—positive on the school side and negative on the hospital side—and sufficient first-stage relevance. Among these admissible specifications, we choose the one that implies the smallest monetary magnitude for a given latent-utility quantity. Equivalently, defining the positive conversion denominators as $a_U = \hat{\gamma}_{1,U}, a_V = -\hat{\gamma}_{1,V}$, we choose the admissible specification with the largest a_k separately for each side $k \in \{U, V\}$. Online Appendix O.3.2 describes the candidate instrument set, the selection rule, and the sensitivity of the counterfactual monetary comparisons to alternative admissible normalizations.

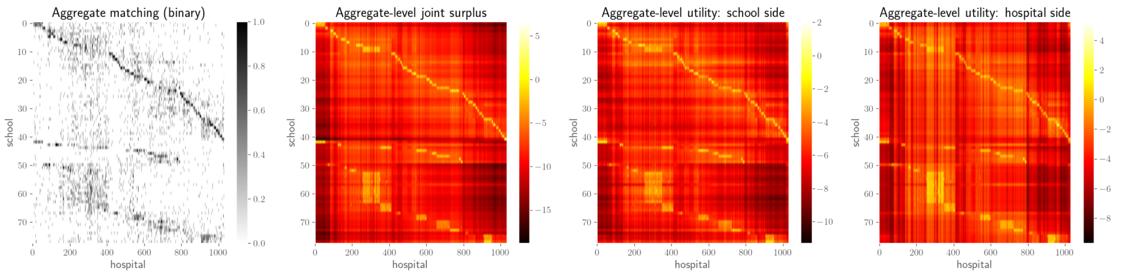


Figure 4. Observed Aggregate Matching and Estimated Aggregate Objects for 2019

Notes: This figure reports the 2019 first-step estimates using the degree-2 polynomial specification. Rows index medical schools, and columns index hospitals; the first panel shows observed matching, with dark cells indicating at least one match. The remaining panels show the estimated aggregate-level joint surplus, school-side utility, and hospital-side utility; lighter colors indicate larger estimated values. Results for other years and polynomial degrees are reported in Online Appendix O.3.5.

6 Empirical Results

This section summarizes the estimation results. Section 6.1 shows the estimation results for the first step. Section 6.2 exhibits the estimation results for the second step: the preference parameters of both sides of the market.

6.1 Model Fit

Figure 4 reports, for 2019, the observed matching matrix and three estimated aggregate objects: joint surplus, school-side utility, and hospital-side utility. We use a second-degree polynomial specification. Each panel is a heatmap with medical schools on the vertical axis and hospitals on the horizontal axis. Schools are ordered by sector, public above private, and then by location, prefecture code and north-to-south position within prefecture; hospitals are ordered similarly. In the three estimated-object panels, lighter colors indicate larger values. Results for the other years and the alternative polynomial degree are reported in Online Appendix O.3.5 and are qualitatively similar.

The estimated aggregate-level utilities and joint surplus capture the observed matching patterns well. For example, both heatmaps indicate that public medical schools are more likely to match with nearby hospitals. For each year and each polynomial degree (2 or 3), we simulate aggregate matching using the estimated aggregate-level joint surplus and report the R^2 from a regression of predicted match counts on observed match counts, without a constant. Based on this criterion, we select the degree-3 specification for 2017 and 2018 and the degree-2 specification for 2019. In the analysis that follows, we therefore use the polynomial degree selected for each year. Figure O.3.3 in the Online Appendix summarizes these model-fit comparisons.

6.2 Money-Metric Preference Estimates

The estimation results of (5) are summarized in Table 2. The first column reports the estimated parameters for school-side utility, while the second column reports the estimated parameters for hospital-side utility. The first row reports the 2SLS estimates of the salary coefficients, $\gamma_{1,U}$ and $\gamma_{1,V}$, in (5).[22] The estimated salary coefficient is positive on the school side and negative on the hospital side, as expected: applicants value salary, whereas salary is costly to hospitals.

All remaining coefficients are reported in money-metric form. Specifically, each coefficient is normalized by the salary coefficient: we divide by $\hat{\gamma}_{1,U}$ in the U column and by $-\hat{\gamma}_{1,V}$ in the V column. Under this normalization, each entry can be interpreted as the wage-equivalent value, measured in million Japanese yen (JPY) per month, of a one-unit increase in the corresponding covariate.

Focusing first on the estimated school-side utility, the coefficient on the logarithm of the number of beds is positive and statistically significant. A 10% increase in the number of beds raises school-side utility by about 8,100 JPY in wage-equivalent terms.[23] The coefficient on the logarithm of distance is negative and statistically significant: a 10% increase in distance lowers school-side utility by about 5,000 JPY. The same-prefecture indicator is positive, with a point estimate of 253,000 JPY, although it is not statistically significant. Overall, the school-side estimates suggest that applicants prefer larger hospitals and geographically closer training locations, with the strongest evidence coming from hospital size and distance.

On the hospital side, the point estimates are broadly consistent with preferences for higher-T-score schools, public schools, and geographically closer schools, but they are not statistically significant. The coefficient on the school's T-score is positive. Since the T-score is normalized by its sample maximum, 74, a one-point increase in the raw T-score corresponds to an increase of 1/74 in the regressor. The point estimate of 2.448 therefore implies a wage-equivalent change

[22] Table O.3.5 in Online Appendix O.3.5 reports the corresponding first-stage results for the IV regressions in Table 2. The first-stage F-statistic is 14.8 for the school side and 97.7 for the hospital side. Details of the selected excluded instruments are provided in Table O.3.5.

[23] For a log regressor, a 10% increase corresponds to $\Delta \ln x = \ln(1.1) \approx 0.095$. Thus, the bed effect is computed as $0.095 \times 0.085 \times 1{,}000{,}000 \approx 8{,}100$ JPY, and the distance effect as $0.095 \times 0.052 \times 1{,}000{,}000 \approx 5{,}000$ JPY.

Table 2. Salary-to-Utility Conversion and Money-Equivalent Preference Estimates

	Applicants (U)	Hospitals (V)
Panel A. Salary-to-Utility Conversion Coefficients		
Salary coefficient, per million JPY	4.802*	-1.316
	(2.730)	(0.928)
Panel B. Money-Equivalent Preference Coefficients (million JPY)		
Hospital covariates		
ln (# Beds)	0.085**	
	(0.038)	
School covariates		
T-score		2.448
		(1.818)
Public		0.277
		(0.220)
Pair covariates		
ln (distance)	-0.052**	-0.157
	(0.024)	(0.122)
Same Prefecture	0.253	1.000
	(0.166)	(0.673)
N	2847	2847

Notes: Panel A reports the 2SLS salary coefficients in latent utility units per million JPY. Panel B reports money-equivalent preference coefficients, obtained by dividing the corresponding latent-utility coefficients by $\hat{\gamma}_{1,U}$ for applicants and by $-\hat{\gamma}_{1,V}$ for hospitals; entries in Panel B are therefore measured in million JPY. All columns include year×prefecture fixed effects. IV specifications are selected from the candidate set described in Section 5.3.2. Standard errors for Panel B use the delta method. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

of $2.448/74 = 0.0331$ million JPY, or about 33,000 JPY, for a one-point increase in the raw T-score.

The coefficient on the public-university indicator is also positive, with a point estimate of 277,000 JPY. The hospital-side pair variables point in the same direction: the coefficient on the logarithm of distance is negative, so a 10% increase in distance is associated with a decrease of about 15,000 JPY in wage-equivalent hospital-side utility, while the same-prefecture indicator is positive and large, with a point estimate of 1.000 million JPY.

Taken together, the second-step estimates provide clearer evidence of geographic preferences on the school side than on the hospital side. Applicants appear to prefer larger and closer hospitals. For hospitals, the estimates are suggestive but not precise enough to support strong conclusions. Therefore, the empirical pattern of frequent matches between local schools and hospitals is most robustly supported by the school-side distance estimates, while the hospital-side local preference should be interpreted more cautiously.

7 Counterfactual Simulations

We use the estimated matching model to compare the status quo cap-based policy with a no-cap-tightening benchmark and with alternative policies evaluated against a common distributional target, defined below. Building on the policy-loss decomposition in Section 4.1, we first measure the total policy loss from the status quo policy relative to a no-cap-tightening benchmark. We then separate this loss into the distributional-target cost and the implementation cost. We further split the implementation cost into the cap-instrument loss and the cap-calibration loss to distinguish how much of the implementation cost is due to the choice of a cap-based instrument, as opposed to miscalibration of the current caps. Finally, we ask whether a tax-and-subsidy policy can recover most of the loss without requiring a large government deficit.

We simulate the 2019 market because it is the latest year in our estimation sample. We compare five policies. First, *Actual Caps* (AC) corresponds to the status quo cap-based policy enforced by the JRMP: for each hospital, we use the actual number of positions reported in 2019 and impose no monetary intervention. Second, *No Cap Tightening* (NC) is an empirical counterpart of the no-cap-tightening benchmark introduced in Section 4, which undoes the status quo tightening of urban capacity. Let c_{ht}^{AC} denote the actual number of positions at hospital h in year t. For hospitals in the six urban prefectures subject to the cap-based policy, we set the number of positions under NC to

$$c_{ht}^{\mathsf{NC}} = \max_{\tau \leq t} c_{h\tau}^{\mathsf{AC}}.$$

For all other hospitals, $c_{ht}^{\mathsf{NC}} = c_{ht}^{\mathsf{AC}}$. Thus, NC restores past urban capacity when the observed policy subsequently reduced it.

The remaining three policies require some lower bound on the number of matchings as inputs Under the geographic taxonomy defined in Section 2, there is one such lower bound for each rural prefecture. We refer to these 15 lower bounds collectively as the *rural floors*. Each prefecture's floor equals the simulated number of matches it receives under AC in 2019. The JRMP does not publish prefecture-level placement targets, so rather than infer the policymaker's distributional goal, we define the rural floors using the match counts delivered by the status quo caps and ask whether the same floors can be satisfied at lower surplus cost through a different instrument. Because the floors are defined from the AC allocation, AC satisfies them by construction, whereas NC need not.

Table 3. Counterfactual Policies in 2019

	AC	TC	NC	OS	BTS
Rural floors satisfied	Yes	Yes	No	Yes	Yes
Price-based instrument	No	No	No	Subsidy	Tax/subsidy
Social surplus	76,930	77,147	80,160	80,009	80,148
Policy loss relative to NC	3,230	3,013	0	152	12
Government revenue	0	0	0	-336	-1
Tax revenue	0	0	0	0	305
Subsidy spending	0	0	0	336	307
# floor violations	0	0	15	0	0

Notes: Money entries are measured in millions of JPY per month and rounded to the nearest million; totals may not add exactly because of rounding. Social surplus equals converted school-side and hospital-side surplus plus government revenue. School-side and hospital-side surplus are converted using their respective estimated salary coefficients; fiscal transfers, including government revenue, tax revenue, and subsidy spending, are converted using the estimated hospital-side salary coefficient. Negative government revenue indicates a deficit.

The third policy, *Tuned Caps* (TC), is the best cap-based policy found by our search among policies that satisfy the rural floors. The fourth policy, *Optimal Subsidy* (OS), is the surplus-maximizing subsidy policy given the rural floors. The last policy, *Balanced Tax-and-Subsidy* (BTS), combines taxes and subsidies to satisfy the same rural floors while keeping the government budget balanced.

To report social surplus and government revenue in monetary units, we convert model-numeraire quantities into JPY using the estimated salary coefficients. School-side and hospital-side participant surplus are converted using their respective salary coefficients. For fiscal transfers, the main tables adopt the conservative convention that all taxes and subsidies are paid or collected through hospitals. Under this convention, one unit of numeraire is worth about 0.76 million JPY, corresponding to the hospital-side salary coefficient of about 1.32. This convention maximizes the reported subsidy cost of OS and minimizes its reported social surplus. Section 7.4 discusses this accounting choice further.

Table 3 summarizes the main results. Relative to NC, AC generates a total policy loss of 3,230 million JPY per month. We decompose this loss as

$$\underbrace{S^{\mathsf{NC}} - S^{\mathsf{AC}}}_{\text{total policy loss}} = \underbrace{\left(S^{\mathsf{NC}} - S^{\mathsf{OS}}\right)}_{\text{distributional-target cost}} + \underbrace{\underbrace{\left(S^{\mathsf{OS}} - S^{\mathsf{TC}}\right)}_{\text{cap-instrument loss}} + \underbrace{\left(S^{\mathsf{TC}} - S^{\mathsf{AC}}\right)}_{\text{cap-calibration loss}}}_{\text{implementation cost}},$$

where S^{P} denotes social surplus under policy P. The decomposition first separates total policy loss into distributional-target cost and implementation cost. The distributional-target cost, $S^{\mathsf{NC}} - S^{\mathsf{OS}}$, is the cost of satisfying the rural floors; it is 152 million JPY per month. The implementation cost, $S^{\mathsf{OS}} - S^{\mathsf{AC}}$, is the additional cost of using the actual cap policy rather than the optimal subsidy policy to satisfy those floors; it is 3,079 million JPY per month, about 95% of total policy loss.

We further decompose this implementation cost into two components. The cap-instrument loss, $S^{\mathsf{OS}} - S^{\mathsf{TC}}$, is the surplus gap between the optimal subsidy policy and the tuned cap policy, both of which satisfy the same rural floors; it is 2,862 million JPY per month. The AC cap-calibration loss, $S^{\mathsf{TC}} - S^{\mathsf{AC}}$, is the surplus gap between the tuned and actual cap policies; it is 217 million JPY per month.

Taken together, these results suggest that cap-based policies are inefficient instruments for satisfying the rural floors in this application: most of the total policy loss reflects the use of the cap instrument rather than the rural floors themselves or the particular way the actual caps are calibrated.

Moreover, we find that a price-based alternative need not require a large government deficit: BTS satisfies the rural floors, delivers social surplus comparable to OS, and leaves a deficit of only about 1 million JPY per month.[24]

In the following subsections, we discuss these five policies in detail. Section 7.1 compares the status quo caps, the no-cap-tightening benchmark, and tuned caps to assess the limits of cap-based policies. Section 7.2 evaluates the subsidy-only and balanced tax-and-subsidy policies, and shows the effectiveness of price-based instruments. Section 7.3 explains why price-based instruments can preserve more surplus than caps. Finally, Section 7.4 discusses how model-numeraire transfers are converted into monetary surplus and fiscal costs.

7.1 Limits of Cap-Based Policies

Relative to NC, AC increases matches in the 15 rural prefectures from 1,129.1 to 1,204.3 expected matches, an increase of 75.2, or 6.7%. The increase is positive in every rural prefecture, ranging from 3.3 expected matches in Yamagata to 8.3 in Okayama. However, relative to NC, AC comes at a large surplus loss. How much of this loss is due to the choice of a cap-based instrument, as opposed to miscalibration of the current caps?

[24]Online Appendix O.3.2 reports a sensitivity analysis of these welfare comparisons to the choice of IV specification, focusing on the comparison among AC, NC, and OS.

Table 4. Urban-Prefecture Caps under the Tuned Caps Policy in 2019

Prefecture	$0.7 \times$ AC	AC	TC	NC	TC/AC	TC−AC
Tokyo	1,005	1,436	1,800	1,800	1.25	364
Kanagawa	486	695	493	859	0.71	-202
Aichi	402	575	402	756	0.70	-173
Kyoto	187	267	187	394	0.70	-80
Osaka	415	593	870	876	1.47	277
Fukuoka	295	421	417	680	0.99	-4

Notes: The table reports prefecture-level total caps for the six urban prefectures. The lower bound is the integer-rounded value of $0.7 \times$ AC used in the grid search, and the upper bound is NC. Positive values in the last column indicate relaxed caps relative to the actual cap; negative values indicate additional tightening.

Ideally, we would answer this question by searching over all possible hospital-level capacity vectors in the six urban prefectures, imposing no taxes or subsidies, and retaining only policies that satisfy the rural floors. Such an exhaustive search is computationally infeasible: the capacity vector is high-dimensional, and each candidate requires solving the aggregate equilibrium. We therefore approximate this ideal benchmark with a heuristic search. The search chooses, for each of the six urban prefectures, a total capacity between 70% of that prefecture's AC total and its NC total, maps each prefecture-level total into hospital-level capacities, and iteratively expands the candidate set around promising feasible and near-feasible vectors. Online Appendix O.3.3 describes the search procedure. Among all candidates evaluated, TC is the cap-based policy that achieves the highest social surplus among those that satisfy the rural floors.

Table 3 shows that TC improves on AC, but only modestly. The tuned cap-based policy reduces the total policy loss relative to NC from 3,230 to 3,013 million JPY per month, recovering about 217 million JPY per month, or 6.7% of the total policy loss under AC. Table 4 shows why the improvement is limited. The tuned capacity vector restores Tokyo's capacity to its NC level and sets Osaka close to its NC level, but offsets these increases by setting Kanagawa, Aichi, and Kyoto close to the lower endpoint. The search therefore reallocates scarce urban capacity across the six urban prefectures, but it still relies on suppressing urban matches in order to satisfy the rural floors. The status quo capacity vector is not fully optimized, but the AC cap-calibration loss is small relative to the cap-instrument loss.

7.2 Effectiveness of Price-Based Instruments

We next evaluate price-based instruments. The subsidy-only policy OS restores urban capacity to the no-cap-tightening benchmark and chooses region-level subsidies by solving the aggregate-level planner's problem subject to the rural floors. As shown in Table 3, the subsidy-only policy eliminates the implementation cost incurred under AC. This gain requires 336 million JPY per month in subsidy spending.

We then ask whether these surplus gains can be preserved while keeping the government budget nearly balanced. The balanced tax-and-subsidy policy BTS also restores urban capacity to the no-cap-tightening benchmark and satisfies the same rural floors. To compute urban taxes that help finance the subsidies required to satisfy the rural floors, we introduce auxiliary upper bounds on matches in the six urban prefectures. These upper bounds are not treated as additional policy objectives; rather, they are computational devices whose Lagrange multipliers generate taxes on high-demand urban regions. We select BTS from a grid search over 4,096 candidate vectors of auxiliary urban upper bounds, choosing the candidate that nearly balances government revenue under the conservative accounting convention. Online Appendix O.3.4 describes the grid search procedure.

Under BTS, the government deficit is 1 million JPY per month, almost eliminating the 336 million JPY per month deficit under OS. The resulting social surplus is 80,148 million JPY per month, only 12 million JPY below NC and 3,218 million JPY above AC. The implied tax and subsidy rates are also moderate: the largest prefecture-level tax is 0.306 million JPY per matched resident per month in Aichi, and the largest subsidy in a rural prefecture is 0.292 million JPY per matched resident per month in Tokushima, compared with 0.321 million JPY under OS. These results show that price-based instruments can satisfy the rural floors with large surplus gains and only a small government deficit.

7.3 Mechanisms

Two features of price-based instruments explain why they preserve more surplus than caps. First, taxes and subsidies can be calibrated to influence the decisions of agents who are closest to being indifferent between urban and rural regions, without eliminating urban positions that generate high joint surplus. A subsidy need only compensate applicants whose surplus loss from a rural match is relatively small; a tax need only impose costs on applicants whose surplus gain from an urban match is small. By contrast, a cap is less flexible: by reducing urban

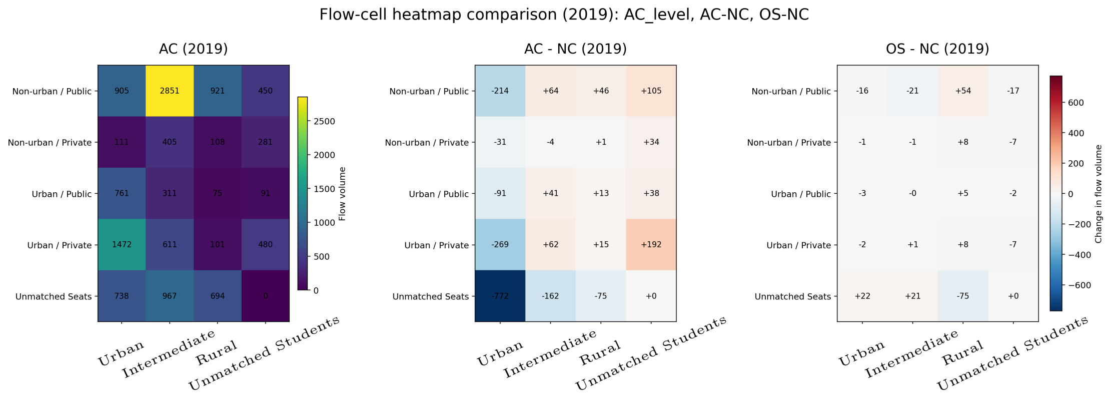


Figure 5. Aggregate Matching Flows and Counterfactual Changes in 2019

Notes: Rows group medical schools using the binary urban/non-urban classification and ownership, whereas columns group hospitals using the three-way classification into urban, intermediate, and rural prefectures; unmatched outcomes are recorded separately. The left panel reports the level of the flow matrix under AC. The middle panel reports AC−NC, and the right panel reports OS−NC. Positive entries in the middle panel are flows that are larger under the status quo caps than under no cap tightening. Positive entries in the right panel are flows that are larger under the optimal subsidy policy than under no cap tightening. Entries are expected aggregate-equilibrium counts, rounded to the nearest integer in the figure.

capacity, it rules out potential matches, including high-surplus matches that a price instrument would preserve.[25]

Second, price-based instruments can be applied directly to the target regions. The rural floors are set in the 15 target prefectures, whereas the cap-based policy restricts capacity in six urban prefectures. Caps therefore affect the rural prefectures only indirectly, by changing the set of urban positions available to applicants. This indirect channel is harder to control: it produces broader reallocations, redirecting some matches to intermediate prefectures and leaving some applicants unmatched.

Figure 5 illustrates these mechanisms in the 2019 counterfactuals. The hospital columns use the urban, intermediate, and rural categories. The middle panel, AC−NC, shows that the status quo caps reduce urban-hospital matches and induce broad reallocations, including toward intermediate prefectures and unmatched outcomes. The right panel, OS−NC, shows a more targeted adjustment: the positive changes are concentrated in hospitals in the rural prefectures, and their magnitudes are small relative to the reallocations induced by caps. The subsidy policy therefore preserves most of the efficiency gain from restoring urban

[25] In our counterfactuals, we assume that the unobserved components of joint surplus are drawn from the same distributions regardless of the number of available positions. This means that, when the policymaker reduces a hospital's capacity, the policy does not selectively remove positions with low realized joint surplus.

capacity while redirecting the marginal matches needed to satisfy the rural floors.

7.4 Model Numeraire and Monetary Implementation

Our theoretical framework assumes that utility is perfectly transferable in a model numeraire. That is, there is a common unit in which payoff can be shifted one-for-one between an applicant and a hospital position: giving one unit to the applicant lowers the position's payoff by exactly one unit, and vice versa. As discussed in Section 5.1, the individual-level equilibrium transfer τ_{ij} combines monetary compensation with nonmonetary program terms, and it is through this composite that the pair divides its baseline joint surplus. The one-for-one property applies to the composite transfer measured in model-numeraire units; it does not extend to money on its own, since a unit of money need not convert to the same number of model-numeraire units for applicants and hospitals.

This asymmetry matters when the government implements taxes and subsidies using money. If the government paid through the numeraire, the induced matching would not depend on which side receives the subsidy. When paid through money, however, it does, because the two sides can have different marginal utilities of money. Our estimates imply an applicant-side salary coefficient of $|\hat{\gamma}_{1,U}| \approx 4.80$ and a hospital-side coefficient of $|\hat{\gamma}_{1,V}| \approx 1.32$. Treating these coefficients as money-to-numeraire conversion rates, delivering a subsidy of q model-numeraire units costs approximately $q/|\hat{\gamma}_{1,U}| = 0.21q$ million JPY when it is paid to applicants and $q/|\hat{\gamma}_{1,V}| = 0.76q$ million JPY when it is channeled through hospitals. Applicant-side implementation therefore requires less money to deliver the same model-numeraire subsidy.

The main analysis assumes that all taxes and subsidies are paid through hospitals. Because $|\hat{\gamma}_{1,V}| < |\hat{\gamma}_{1,U}|$, this convention assigns the larger monetary cost to subsidy spending and hence yields a conservative money-equivalent surplus for OS. Applicant and hospital payoffs are converted at $|\hat{\gamma}_{1,U}|$ and $|\hat{\gamma}_{1,V}|$, respectively, while fiscal transfers are converted at $|\hat{\gamma}_{1,V}|$. Money-equivalent social surplus is therefore not model-numeraire social surplus scaled by a common factor, and the two measures need not deliver the same policy ranking. This is why OS attains higher surplus than BTS in the model numeraire, while BTS ranks slightly higher under the reported money-equivalent accounting.

Our main takeaway remains valid despite this ranking reversal between (OS) and (BTS). Even under the conservative hospital-side convention, price-based instruments meet the rural floors far more efficiently than quantity-based instruments do. They would achieve an even higher money-equivalent surplus if some taxes and subsidies were instead channeled through applicants, since this would let the government deliver the same model-numeraire amount at a lower fiscal cost. The choice between price- and quantity-based instruments is therefore robust to how taxes and subsidies are implemented.

8 Conclusion

This paper develops a framework for evaluating quantity-based and price-based instruments in matching markets with distributional disparities. We build a transferable-utility matching model with regional constraints, such as caps and floors on the number of matches, and show that region-level taxes and subsidies can implement the welfare-maximizing allocation among matchings satisfying a given distributional objective. By embedding this model in the aggregate matching framework of Galichon and Salanié (2022), we provide an empirical procedure to construct such policies using aggregate match data and an observable measure of transfers, such as salary.

Applying the framework to the Japan Residency Matching Program, we find that the status quo policy that caps positions in urban areas generates substantial losses in social surplus relative to the no-intervention benchmark, while redirecting some applicants toward positions in underserved areas. We show that most of the surplus loss is due to the implementation cost, not the cost of the distributional target itself. A targeted subsidy policy can achieve the same matches in rural prefectures as in the status quo policy, while recovering about 95 percent of the surplus loss. Our results also suggest that this is not simply a matter of recalibrating the caps: even the most efficient cap policy we identify recovers only a small fraction of the loss, indicating that the loss is largely intrinsic to cap-based regulation.

These findings suggest that price-based instruments can be a more efficient tool than quantity-based regulations in matching markets where policymakers pursue regional or distributional objectives. More broadly, the framework applies to settings in which transfers are present but individual-level preference data are unavailable. Future work could extend the approach to richer distributional constraints or to dynamic outcomes, such as physicians' subsequent location choices and long-run settlement patterns.

References

Abdulkadiroğlu, A. and T. Sönmez (2003): “School Choice: A Mechanism Design Approach,” *American Economic Review*, 93, 729–747.

Agarwal, N. (2015): “An Empirical Model of the Medical Match,” *American Economic Review*, 105, 1939–1978.

——— (2017): “Policy Analysis in Matching Markets,” *American Economic Review*, 107, 246–250.

Berry, S., J. Levinsohn, and A. Pakes (1995): “Automobile Prices in Market Equilibrium,” *Econometrica*, 63, 841–890.

Bobba, M., T. Ederer, G. León-Ciliotta, C. Neilson, and M. G. Nieddu (2024): “Teacher Compensation and Structural Inequality: Evidence from Centralized Teacher School Choice in Perú,” Working Paper 29068, National Bureau of Economic Research.

Boyd, S. and L. Vandenberghe (2004): *Convex Optimization*, Cambridge: Cambridge University Press.

Chiappori, P.-A. (2017): *Matching with Transfers: The Economics of Love and Marriage*, Princeton, NJ: Princeton University Press.

Clarke, F. (2013): *Functional Analysis, Calculus of Variations and Optimal Control*, vol. 264 of *Graduate Texts in Mathematics*, London: Springer London.

Combe, J., U. Dur, O. Tercieux, C. Terrier, and M. U. Ünver (2025): “Market Design for Distributional Objectives in (Re)assignment: An Application to Improve the Distribution of Teachers in Schools,” CEPR Discussion Paper 19893, Centre for Economic Policy Research.

Dworczak, P., S. D. Kominers, and M. Akbarpour (2021): “Redistribution Through Markets,” *Econometrica*, 89, 1665–1698.

Ehlers, L., I. E. Hafalir, M. B. Yenmez, and M. A. Yildirim (2014): “School Choice with Controlled Choice Constraints: Hard Bounds versus Soft Bounds,” *Journal of Economic Theory*, 153, 648–683.

Endo, H. (2019): "Ishi no Sotsugo Kenshu no Genjo to Kadai: Ishi no Chiiki Henzai Mondai o Chushin ni [The Current Status and Challenges of Postgraduate Medical Training for Physicians: With a Focus on Regional Maldistribution of Physicians]," *Journal of Social Security Research*, 3, 476–491, in Japanese.

Fragiadakis, D. and P. Troyan (2017): "Improving Matching under Hard Distributional Constraints," *Theoretical Economics*, 12, 863–908.

Galichon, A. (2016): *Optimal Transport Methods in Economics*, Princeton, NJ: Princeton University Press.

Galichon, A. and B. Salanié (2022): "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," *Review of Economic Studies*, 89, 2600–2629.

——— (2023): "Structural Estimation of Matching Markets with Transferable Utility," in *Online and Matching-Based Market Design*, ed. by F. Echenique, N. Immorlica, and V. V. Vazirani, Cambridge: Cambridge University Press, 552–572.

Guesnerie, R. and K. Roberts (1984): "Effective Policy Tools and Quantity Controls," *Econometrica*, 52, 59–86.

Hafalir, I. E., M. B. Yenmez, and M. A. Yildirim (2013): "Effective Affirmative Action in School Choice," *Theoretical Economics*, 8, 325–363.

Iizuka, T. and Y. Watanabe (2016): "The Impact of Physician Supply on the Healthcare System: Evidence from Japan's New Residency Program," *Health Economics*, 25, 1433–1447.

Jalota, D., M. Ostrovsky, and M. Pavone (2025): "Matching with Transfers under Distributional Constraints," *Games and Economic Behavior*, 152, 313–332.

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

Kojima, F. (2012): "School Choice: Impossibilities for Affirmative Action," *Games and Economic Behavior*, 75, 685–693.

Kojima, F., N. Sun, and N. N. Yu (2020): "Job Matching under Constraints," *American Economic Review*, 110, 2935–2947.

McFadden, D. (1981): "Econometric Models of Probabilistic Choice," in *Structural Analysis of Discrete Data with Econometric Applications*, ed. by C. F. Manski and D. McFadden, Cambridge, MA: MIT Press, 198–272.

Roth, A. E. and M. Sotomayor (1990): *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*, vol. 18 of *Econometric Society Monographs*, Cambridge: Cambridge University Press.

Sakai, R., W. Wang, N. Yamaguchi, H. Tamura, R. Goto, and I. Kawachi (2013): "The Impact of Japan's 2004 Postgraduate Training Program on Intra-Prefectural Distribution of Pediatricians in Japan," *PLOS ONE*, 8, e77045.

Shapley, L. S. and M. Shubik (1971): "The Assignment Game I: The Core," *International Journal of Game Theory*, 1, 111–130.

Weitzman, M. L. (1974): "Prices vs. Quantities," *Review of Economic Studies*, 41, 477–491.

——— (1977): "Is the Price System or Rationing More Effective in Getting a Commodity to Those Who Need It Most?" *Bell Journal of Economics*, 8, 517–524.

Williamson, D. P. (2019): *Network Flow Algorithms*, Cambridge: Cambridge University Press.

Yokote, K. (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, 37–77.

A Additional Institutional Details

A.1 Definitions of Unmatched Applicants

The definitions of "unmatched" in the statistics reported by the JRMP ("JRMP unmatched") and in our dataset ("final unmatched") differ. This appendix clarifies the precise definitions and provides relevant institutional details.

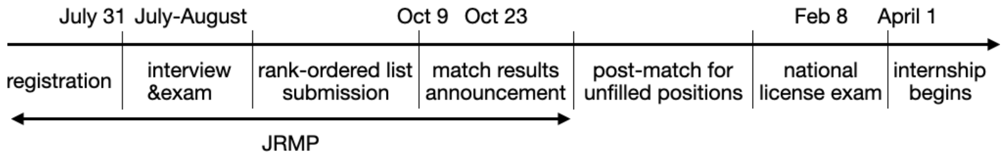


Figure A.1. JRMP Timeline

Figure A.1 illustrates the timeline of the matching process, using the 2025 JRMP schedule as an example. The centralized matching procedure (JRMP) concludes in late October, when the matching results are announced. "JRMP unmatched" refers to applicants who fail to obtain a residency position at this stage. After the centralized match, unmatched applicants may continue to search for positions through a decentralized "post-match" process before residency training begins on April 1.

Following these two stages—the centralized and decentralized matches—all applicants take the national licensing examination in February. Those who fail the examination are not permitted to start their residency training and must reenter the matching process in the subsequent year, even if they had already secured a position. The positions originally assigned to these applicants are then released back to the market.[26] In our dataset, these applicants are also classified as "unmatched." Thus, the "final unmatched" category in our data includes (i) applicants who did not secure a position in either the centralized or decentralized matching stages, and (ii) applicants who failed the national examination. All hospitals and applicants participating in the post-match market must be registered with the JRMP and are therefore subject to the regional caps. In our analysis, the "matching outcome" refers to the overall result.

[26] A few applicants continue to search for positions after the examination to fill the positions released by those who failed the national exam.

B Omitted Proofs

B.1 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 $j \in J$.[27] The goal of this subsection is to prove the following result (I repeat the statement of Theorem 1):

Theorem B.1. *Fix any market (I, J, Z, z, Φ) and regional constraints $(\underline{o}_z, \bar{o}_z)_z$. There exists a tax-and-subsidy policy w^* and a payoff profile (u, v) such that $(d^*, (u, v))$ is a stable outcome under w^*.*

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:

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

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 B.1. *Fix any matching market (I, J, Z, z, Φ) and integral 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 B.1

The proof can be done in two steps: first, we reduce (P_0') to a minimum-cost circulation problem with integral lower and upper bounds. We then apply the following standard result.[28]

Lemma B.1. *Consider a feasible minimum-cost circulation problem on a directed netweork in which all edges have integral lower bounds and integral upper bounds. Then, there exists an optimal circulation whose flow on each edge is integral.*

[27] 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.

[28] See, for example, Property 5.5 of Williamson (2019).

Below, we show that we can reduce the social surplus maximization problem (P_0') to a minimum-cost flow problem. Proposition B.1 follows from Lemma B.1.

Lemma B.2. (P_0') *can be reduced to a minimum-cost flow problem.*

Proof. Construct a directed network $G = (V, E)$ as follows (see also Figure B.1.) The node set is

$$V = \{s, t\} \cup I \cup J \cup Z.$$

The edge set E is

$$E = \{(s, i) : i \in I\} \cup (I \times J) \cup \{(j, z(j)) : j \in J\} \cup \{(z, t) : z \in Z\} \cup \{(t, s)\}.$$

For each edge $e \in E$, define lower bounds ℓ_e, upper bounds u_e, and costs c_e by

$$\begin{aligned}
&\ell_{s,i} = 0, u_{s,i} = 1, c_{s,i} = 0 && (i \in I),\\
&\ell_{i,j} = 0, u_{i,j} = 1, c_{i,j} = -\Phi_{ij} && (I \times J),\\
&\ell_{j,z(j)} = 0, u_{j,z(j)} = 1, c_{j,z(j)} = 0 && (j \in J),\\
&\ell_{z,t} = \underline{o}_z, u_{z,t} = \overline{o}_z, c_{z,t} = 0 && (z \in Z),\\
&\ell_{t,s} = 0, u_{t,s} = |I| \wedge |J|, c_{t.s} = 0.
\end{aligned}$$

There is a natural bijection between feasible solutions in (P_0') and feasible flows in the network: any flow f induces a solution in (P_0') via $d_{ij} = f_{ij}$ for each $(i, j) \in I \times J$; any matching d induces a flow such that

$$f_{ij} = d_{ij},\ f_{si} = \sum_{j \in J} d_{ij},\ f_{j,z(j)} = \sum_{i \in I} d_{ij},\ f_{zt} = \sum_{j \in z} \sum_{i \in I} d_{ij},\ f_{t,s} = \sum_{z \in Z} f_{zt}.$$

Integral solutions in (P_0') correspond to integral flows.

The cost of a flow, $-\sum_{ij} f_{ij} \Phi_{ij}$, is the negative of the social surplus generated by the corresponding matching. Thus, minimizing the cost in this network corresponds to maximizing the objective function in (P_0').

□

Proof of Theorem B.1

Consider the following dual problem of (P_0'):

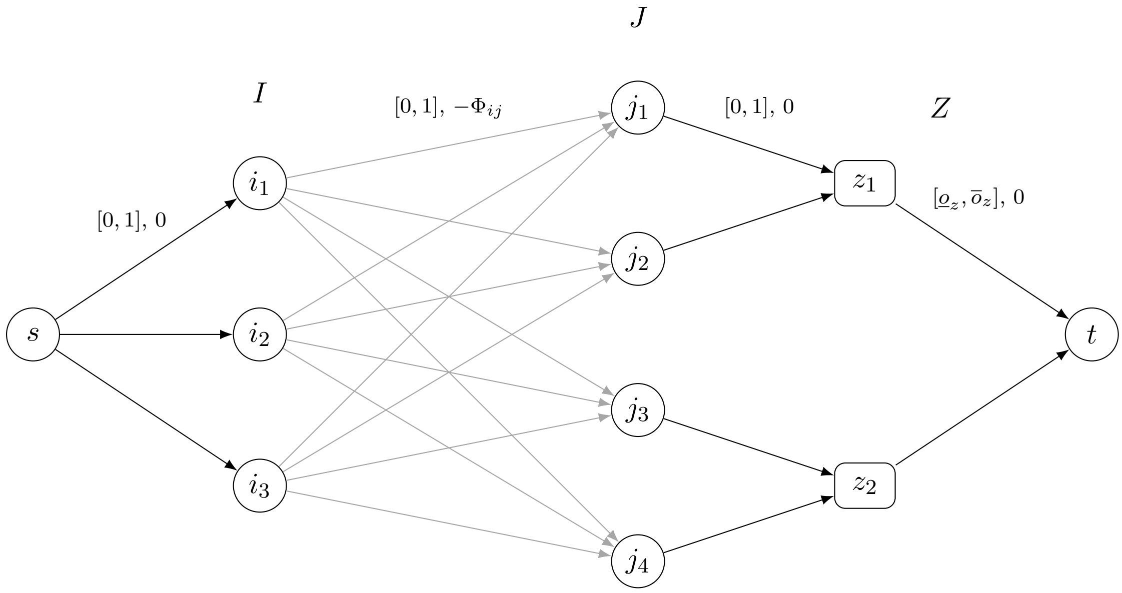


Figure B.1. Network for the case $|I| = 3$, $|J| = 4$, and $Z = \{z_1, z_2\}$ with $z_1 = \{j_1, j_2\}$, $z_2 = \{j_3, j_4\}$. The edge (t, s) is omitted.

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

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'). We can consider a pair of LPs under w^* that characterizes the stable outcomes given w^*:

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

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

By construction, $(d^*, (u, v))$ is a solution to $(\mathrm{P}_{\mathrm{w}^*})$ and $(\mathrm{D}_{\mathrm{w}^*})$: the optimality conditions of (P_0') and (D_0') imply that d^* and (u, v) are feasible for $(\mathrm{P}_{\mathrm{w}^*})$ and $(\mathrm{D}_{\mathrm{w}^*})$, respectively, and the complementary slackness conditions of $(\mathrm{P}_{\mathrm{w}^*})$ and $(\mathrm{D}_{\mathrm{w}^*})$ are satisfied since they are included in the optimality conditions of (P'_0) and (D'_0). Therefore, $(d^*, (u, v))$ is a stable outcome given Φ and w^*.

Corollary B.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 tax-and-subsidy 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 $j \in J \setminus J'$. □

B.2 Proof of Lemma 1

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

Lemma B.3. *Suppose that U and V are aggregate-level utilities given (u, v) (i.e., (2) holds), and 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 $x \in X$ and $y \in Y$.*

Proof. Fix any $i \in x$ and $j \in y$. By stability, we have

$$u_i + v_j \geq \Phi_{ij} - w_{z(y)} = \Phi_{xy} - w_{z(y)} + \varepsilon_{iy} + \eta_{xj}.$$

Thus, we have

$$u_i - \varepsilon_{iy} \geq \Phi_{xy} - w_{z(y)} - (v_j - \eta_{xj}).$$

Since this holds for any $i \in x$ and $j \in y$, by taking the infimum with respect to $i \in x$ and $j \in y$ separately in both sides, we have

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

and thus $U_{xy} + V_{xy} \geq \Phi_{xy} - w_{z(y)}$ holds.

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 $U_{xy} + V_{xy} > \Phi_{xy} - w_{z(y)}$. By (2), 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.

□

[29] The following proof of Lemma B.3 is almost identical to the proof of Proposition 1 of Galichon and Salanié (2022).

Proof of Lemma 1. Fix any category $x \in X$ and applicant $i \in x$. 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 category $y \in Y$ and position $j \in J$ in category 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 category $x \in X$ and applicant $i \in x$ such that

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

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

$$\begin{aligned} \Phi_{ij} - w_{z(y)} &= u_i + v_j \\ &> \left(\max_{y' \in Y_0} \{U_{xy'} + \varepsilon_{iy'}\} \right) + \left(\max_{x' \in X_0} \{V_{x'y} + \eta_{x'j}\} \right) \\ &\geq U_{xy(j)} + \varepsilon_{iy(j)} + V_{xy(j)} + \eta_{xj} \\ &\geq \Phi_{xy} - w_{z(y)} + \varepsilon_{iy(j)} + \eta_{xj} \quad (\because \text{ Lemma B.3}) \\ &= \Phi_{ij} - w_{z(y)}. \end{aligned}$$

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

$$u_i = \Phi_{i,j_0} = \varepsilon_{i,y_0} > \max_{y \in Y_0} \{U_{xy} + \varepsilon_{iy}\} \geq \varepsilon_{i,y_0}.$$

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. □

B.3 Proof of Theorem 2

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

Lemma B.4. *Under Assumptions 1, 2, and 3, G and H are strictly increasing and strictly convex.*

Proof. Recall that we normalize $U_{x,y_0} = V_{x_0,y} = 0$. Below, U and V are defined on this normalized domain.

G is strictly increasing.

Take any U^1, U^2 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 and U^2, and let $s \in [0, 1]$. Since

$$\begin{aligned}
&sG(U^1) + (1-s)G(U^2) \\
&= \sum_x n_x \mathbb{E}\left[\left(\max_y\ s(U^1_{xy} + \varepsilon_{iy})\right) + \left(\max_y\ (1-s)(U^2_{xy} + \varepsilon_{iy})\right)\right] &(6)\\
&\geq \sum_x n_x \mathbb{E}\left[\max_y\ \left\{sU^1_{xy} + (1-s)U^2_{xy}\right\} + \varepsilon_{iy}\right] &(7)\\
&= G\left(sU^1 + (1-s)U^2\right)
\end{aligned}$$

holds, G is a convex function.

Now suppose $U^1 \neq U^2$.

For some $x \in X$, the vectors U^1_x and U^2_x differ. Let $\Delta_y \coloneqq U^1_{xy} - U^2_{xy}$ for

$y \in Y$. Since $U_x^1 \neq U_x^2$, there exists $y \in Y$ such that $\Delta_y \neq 0$. By normalization, $\Delta_{y_0} = 0$. Hence Δ is not constant on Y_0, so there exist $y, y' \in Y_0$ with $\Delta_y > \Delta_{y'}$. Equivalently,

$$U_{xy'}^1 - U_{xy}^1 < U_{xy'}^2 - U_{xy}^2.$$

Choose $\varepsilon_{iy} - \varepsilon_{iy'}$ strictly between these two numbers and choose the remaining shock coordinates sufficiently low. This describes a nonempty open set on which y is the unique maximizer under U_x^1 and y' is the unique maximizer under U_x^2. Full support of P_x gives this set positive probability. On that set, for every $s \in (0,1)$,

$$\Big(\max_y \; s(U_{xy}^1 + \varepsilon_{iy})\Big) + \Big(\max_y \; (1-s)(U_{xy}^2 + \varepsilon_{iy})\Big) > \max_y \left\{sU_{xy}^1 + (1-s)U_{xy}^2 + \varepsilon_{iy}\right\}$$

occurs with strictly positive probability, and thus (6) > (7) holds. Therefore, for any $s \in (0,1)$, we have

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

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

Using this lemma, we can prove the theorem:

Proof of Theorem 2. For (P), since G and H are differentiable (WDZ theorem), G^* and H^* are strictly convex (Proposition D.14 of Galichon (2016)). 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 $(\bar{w}, \underline{w}) \neq (\bar{w}', \underline{w}')$. At either optimum, the KKT multiplier on the constraint indexed by (x, y) equals $\partial G / \partial U_{xy}$. The WDZ formula and full support imply that this multiplier is strictly positive. Complementary slackness therefore makes every such constraint bind:

$$U_{xy} + V_{xy} = \Phi_{xy} - \bar{w}_{z(y)} + \underline{w}_{z(y)}, \qquad U'_{xy} + V'_{xy} = \Phi_{xy} - \bar{w}'_{z(y)} + \underline{w}'_{z(y)}.$$

Since $U = U'$ and $V = V'$, this implies, for every z,

$$\bar{w}_z - \underline{w}_z = \bar{w}'_z - \underline{w}'_z. \tag{8}$$

Since $\bar{o}_z > \underline{o}_z$, we must have $\bar{w}_z \underline{w}_z = 0$; otherwise, reducing both positive multipliers by the same small amount leaves the constraints unchanged and strictly lowers the objective. Similarly, $\bar{w}'_z \underline{w}'_z = 0$. Together with (8), this yields $\bar{w} = \bar{w}'$ and $\underline{w} = \underline{w}'$, a contradiction. □

B.4 Proof of Proposition 1

Lemma B.5. *Suppose Assumption 1 holds. Fix a stable outcome in the observed market with zero taxes and subsidies, and let U and V be the aggregate utilities associated with this stable outcome. If applicant $i \in s$ is matched with position $j \in h$, then*

$$u_i = U_{sh} + \varepsilon_{ih}, \qquad v_j = V_{sh} + \eta_{sj}.$$

Proof. Fix a medical school s and a hospital h. By the definition of the aggregate utilities,

$$U_{sh} = \min_{i' \in s}\{u_{i'} - \varepsilon_{i'h}\}, \qquad V_{sh} = \min_{j' \in h}\{v_{j'} - \eta_{sj'}\}.$$

Hence, for every $i' \in s$ and $j' \in h$,

$$u_{i'} - \varepsilon_{i'h} \geq U_{sh}, \qquad v_{j'} - \eta_{sj'} \geq V_{sh}.$$

We first show that $U_{sh} + V_{sh} \geq \Phi_{sh}$. For any $i' \in s$ and $j' \in h$, stability in the observed market with zero taxes and subsidies implies

$$u_{i'} + v_{j'} \geq \Phi_{i'j'}.$$

By Assumption 1,

$$\Phi_{i'j'} = \Phi_{sh} + \varepsilon_{i'h} + \eta_{sj'}.$$

Therefore,

$$(u_{i'} - \varepsilon_{i'h}) + (v_{j'} - \eta_{sj'}) \geq \Phi_{sh}$$

for every $i' \in s$ and $j' \in h$. Taking the minimum over $i' \in s$ and $j' \in h$ yields

$$U_{sh} + V_{sh} \geq \Phi_{sh}.$$

Now suppose applicant $i \in s$ is matched with position $j \in h$. Matched-pair equality in the stable outcome gives

$$u_i + v_j = \Phi_{ij}.$$

Using Assumption 1 again,

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

Thus,

$$(u_i - \varepsilon_{ih}) + (v_j - \eta_{sj}) = \Phi_{sh}.$$

Since

$$U_{sh} \leq u_i - \varepsilon_{ih}, \qquad V_{sh} \leq v_j - \eta_{sj},$$

we also have

$$U_{sh} + V_{sh} \leq \Phi_{sh}.$$

Combining this with $U_{sh} + V_{sh} \geq \Phi_{sh}$, we obtain

$$U_{sh} + V_{sh} = \Phi_{sh}.$$

Moreover,

$$u_i - \varepsilon_{ih} \geq U_{sh}, \qquad v_j - \eta_{sj} \geq V_{sh},$$

and the sum of the two left-hand sides equals the sum of the two right-hand sides. Therefore both inequalities must bind:

$$u_i - \varepsilon_{ih} = U_{sh}, \qquad v_j - \eta_{sj} = V_{sh}.$$

Equivalently,

$$u_i = U_{sh} + \varepsilon_{ih}, \qquad v_j = V_{sh} + \eta_{sj}.$$

□

Proof of Proposition 1. Fix a hospital h with

$$M_h = \sum_{s'} \mu_{s'h} > 0, \qquad \omega_{sh} = \frac{\mu_{sh}}{M_h}.$$

For each medical school s, define the set of realized matches in cell (s, h) by

$$\mathcal{M}_{sh} = \{(i, j) : i \in s,\ j \in h,\ d_{ij} = 1\}.$$

Then $|\mathcal{M}_{sh}| = \mu_{sh}$.

Take any $(i,j) \in \mathcal{M}_{sh}$. Since the observed market has zero taxes and subsidies and (i,j) is a realized match, matched-pair equality gives

$$u_i + v_j = \Phi_{ij}.$$

By the definition of base utilities,

$$\Phi_{ij} = U_{ij}^{\text{base}} + V_{ij}^{\text{base}}.$$

Therefore,

$$u_i - U_{ij}^{\text{base}} = V_{ij}^{\text{base}} - v_j.$$

Hence, for a realized matched pair, the admissible interval defining τ_{ij} collapses to a singleton:

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

By Lemma B.5,

$$u_i = U_{sh} + \varepsilon_{ih}, \qquad v_j = V_{sh} + \eta_{sj}.$$

By Assumption 4,

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

Substituting these expressions gives

$$\tau_{ij} = u_i - U_{ij}^{\text{base}} = U_{sh} - U_{sh}^{\text{base}},$$

and also

$$\tau_{ij} = V_{ij}^{\text{base}} - v_j = V_{sh}^{\text{base}} - V_{sh}.$$

Thus, for every realized match $(i,j) \in \mathcal{M}_{sh}$,

$$\tau_{ij} = U_{sh} - U_{sh}^{\text{base}} = V_{sh}^{\text{base}} - V_{sh}.$$

Averaging this pointwise identity over all realized matches at hospital h, we obtain

$$\begin{aligned}
\iota_h &= \frac{1}{M_h}\sum_s \sum_{(i,j)\in\mathcal{M}_{sh}} \tau_{ij} \\
&= \frac{1}{M_h}\sum_s \mu_{sh}\left(U_{sh} - U_{sh}^{\text{base}}\right) \\
&= \sum_s \omega_{sh}\left(U_{sh} - U_{sh}^{\text{base}}\right).
\end{aligned}$$

The same averaging argument using $\tau_{ij} = V_{sh}^{\text{base}} - V_{sh}$ gives

$$\iota_h = \sum_s \omega_{sh}\left(V_{sh}^{\text{base}} - V_{sh}\right).$$

Combining the two identities yields

$$\iota_h = \sum_s \omega_{sh}\left(U_{sh} - U_{sh}^{\text{base}}\right) = \sum_s \omega_{sh}\left(V_{sh}^{\text{base}} - V_{sh}\right),$$

which proves the proposition. □

B.5 Derivation of two-way fixed effects Poisson regression of Galichon and Salanié (2023)

We derive the estimator based on two-way fixed effects Poisson regression, which is proposed in Galichon and Salanié (2023). 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 μ_{xy}, the number of matches between categories x and y. 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 ignoring constants is written as follows:

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

The minimization objective function is

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

As shown in Galichon and Salanié (2023), minimizing F is equivalent to the moment matching estimator.

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 (2023). Note that, while they take the form of fixed effects, they do *not* capture baseline utilities for each category—such as the value of remaining unmatched. As noted in Section 4.2.1, the average unmatched values are normalized to 0 in our model.

Online Appendix

O.1 Stable Outcomes as an Equilibrium of the JRMP game

Although deferred acceptance determines the final assignment in JRMP, the algorithm is only one of the many stages of a broader process. Earlier stages are decentralized: before rank-order lists are submitted, hospitals choose compensation and program attributes, applicants decide where to apply and interview, and both sides form rankings based on these interactions. In this section, we show that a stable outcome (Definition 1) arise as an equilibrium of a stylized game that mirrors the JRMP's matching process; this provides a justification for using stable outcomes as a reduced-form equilibrium notion for the broader matching process.

The game consists of three periods. In Period 1, each position simultaneously announces a profile of transfers $t = (t_{ij})_{i,j}$, where $t_j = (t_{ij})_i$ is the transfer profile chosen by position j. In Period 2, observing all the wage offers, applicants and positions submit their preference lists. In Period 3, the applicant-proposing deferred acceptance (DA) algorithm determines the final matching. The utility of applicant i being matched with a position j is defined as follows: the *base utility*, which is the utility i derives net of transfer, is denoted by U_{ij}. Given transfer offer t, when matched with position j, applicant i's payoff is $U_{ij} + t_{ij}$. Similarly, the base utility of position j when matched with applicant i is denoted by V_{ij}, and the payoff of j when matched with i given t 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, using some tie-breaking rule.[30] Given truthful reporting in Period 2 and a fixed tie-breaking rule, the continuation outcome after any transfer profile is uniquely determined by applicant-proposing deferred acceptance. Hence, the strategic interaction reduces to a one-shot transfer-posting game among positions in Period 1.

Note that although the base-salary component is in fact published in JRMP listings before applications, other components of transfers, such as hospital discretion, are exercised after the match (night-shift assignment, rotation placement, allocation of senior physicians' time). The posted transfer in period 1 is interpreted

[30]Since the DA algorithm is strategy-proof for the proposing side, this means that we assume the truthful report from the other side.

as the anticipated total compensation package rather than a formally posted contract: players form beliefs about this discretionary treatment through hospital visits, interviews, and program reputations.

The following proposition establishes the link between the stable outcome and the equilibrium of this game. It shows that any outcome resulting from a pure Nash equilibrium of the period-1 transfer-posting game is stable in the original TU matching model.

Proposition O.1.1. *Fix any tie-breaking rule used both on and off path. Let*

$$t^* = (t^*_{ij})_{i \in I, j \in J}$$

be a pure-strategy Nash equilibrium of the period-1 transfer-posting game in the JRMP game, and let

$$(d^*, u^*, v^*)$$

be the induced matching and payoffs. Then (d^*, u^*, v^*) *is a stable outcome of the original TU matching market* (I, J, Φ)*.*

Proof. Let μ^* be the matching produced by applicant-proposing deferred acceptance under the strict preference profile induced by t^*, where applicant i ranks each $j \in J$ by $U_{ij} + t^*_{ij}$ and the outside option j_0 by Φ_{ij_0}, and position j ranks each $i \in I$ by $V_{ij} - t^*_{ij}$ and the outside option i_0 by Φ_{i_0j}; ties are broken by the fixed deterministic rule. Let d^* be the matrix representation of μ^*. Since deferred acceptance always returns a matching, d^* is feasible.

Define the induced payoff profile (u^*, v^*) by

$$u^*_i = \begin{cases} U_{i\mu^*(i)} + t^*_{i,\mu^*(i)} & \text{if } \mu^*(i) \in J, \\ \Phi_{ij_0} & \text{if } \mu^*(i) = j_0, \end{cases} \qquad v^*_j = \begin{cases} V_{\mu^*(j),j} - t^*_{\mu^*(j),j} & \text{if } \mu^*(j) \in I, \\ \Phi_{i_0j} & \text{if } \mu^*(j) = i_0. \end{cases}$$

We show that (d^*, u^*, v^*) is a stable outcome of the original TU matching market (I, J, Φ).

First, because μ^* is stable for the induced nontransfer preference profile, it is individually rational. Hence, for every applicant $i \in I$,

$$u^*_i \geq \Phi_{ij_0},$$

with equality if $\mu^*(i) = j_0$, and for every position $j \in J$,

$$v^*_j \geq \Phi_{i_0j},$$

with equality if $\mu^*(j) = i_0$.

Second, if $\mu^*(i) = j \in J$, then

$$u_i^* = U_{ij} + t_{ij}^*, \qquad v_j^* = V_{ij} - t_{ij}^*,$$

so

$$u_i^* + v_j^* = U_{ij} + V_{ij} = \Phi_{ij}.$$

Thus, equality holds for every matched pair.

It remains to show that

$$u_i^* + v_j^* \geq \Phi_{ij} \qquad \text{for all } i \in I,\ j \in J.$$

Suppose, toward a contradiction, that there exist $i \in I$ and $j \in J$ such that

$$u_i^* + v_j^* < \Phi_{ij}.$$

Choose $\varepsilon > 0$ such that

$$u_i^* + v_j^* + 2\varepsilon < \Phi_{ij}.$$

Consider the unilateral deviation by position j given by

$$\tilde{t}_{ij} = u_i^* - U_{ij} + \varepsilon, \qquad \tilde{t}_{kj} = t_{kj}^* \quad \text{for all } k \neq i,$$

while all transfers posted by positions $h \neq j$ remain unchanged.

Under this deviation,

$$U_{ij} + \tilde{t}_{ij} = u_i^* + \varepsilon > u_i^*.$$

Hence under the deviating transfer profile, applicant i ranks position j strictly above her equilibrium assignment $\mu^*(i)$. Also,

$$V_{ij} - \tilde{t}_{ij} = U_{ij} + V_{ij} - u_i^* - \varepsilon = \Phi_{ij} - u_i^* - \varepsilon > v_j^* + \varepsilon > v_j^*.$$

Therefore, if position j ends up matched with applicant i after the deviation, then position j's payoff is strictly greater than v_j^*.

We next show that, under the deviating transfer profile, applicant i must apply to position j during deferred acceptance.

Consider the two runs of applicant-proposing deferred acceptance, one under t^* and one under the deviating transfer profile. The only changes in preferences are the following: position j may move up in applicant i's ranking, while the relative order of all other positions for applicant i is unchanged; and position j's ranking of applicant i may change, while the relative order of all other applicants for position j is unchanged.

We claim that, until the round in which applicant i applies to position j under the deviating transfer profile, the history of deferred acceptance is identical to the history under t^*.

We prove this by induction on rounds. In round 1, every applicant $k \neq i$ makes exactly the same proposal as under t^*, because their preferences are unchanged. If applicant i proposes to position j already in round 1, then the claim is proved. Otherwise applicant i's most preferred position among those ranked above j is unchanged, so applicant i also makes the same round-1 proposal as under t^*. Hence the tentative assignments after round 1 coincide in the two runs.

Now suppose that the histories coincide through round $r-1$, and applicant i has not yet applied to position j. Then every applicant $k \neq i$ has exactly the same rejection history as under t^*, so each such applicant makes the same round-r proposal as under t^*. Applicant i has also received exactly the same rejections from all positions she ranks above j, because those positions have unchanged preferences and have seen exactly the same applications as under t^*. Therefore in round r, applicant i either proposes to the same position as under t^*, if that position is still ranked above j, or proposes to j, if every position ranked above j has already rejected her. Thus, unless applicant i applies to j in round r, the full round-r history is the same as under t^*. This completes the induction.

Since under the deviation applicant i ranks j strictly above $\mu^*(i)$, every position that applicant i ranks above j under the deviation is also ranked above $\mu^*(i)$ under t^*. Under the run at t^*, applicant i is eventually rejected by every position ranked above $\mu^*(i)$ if $\mu^*(i) \in J$, and by every acceptable position if $\mu^*(i) = j_0$. Hence the induction implies that applicant i must eventually apply to position j under the deviating transfer profile.

When applicant i applies to position j, position j either tentatively holds i, or rejects i in favor of some applicant k such that

$$V_{kj} - \tilde{t}_{kj} \geq V_{ij} - \tilde{t}_{ij} > v_j^*.$$

In the former case, position j's tentative payoff is already strictly greater than v_j^*. In the latter case, the applicant tentatively held by position j yields payoff strictly greater than v_j^*. Since in deferred acceptance a position always keeps its most preferred applicant among those seen so far, the tentative payoff of position j can only weakly increase over time. Therefore position j's final payoff under the deviation is strictly greater than v_j^*.

Thus position j has a profitable unilateral deviation from t^*, contradicting the assumption that t^* is a pure-strategy Nash equilibrium. We conclude that

$$u_i^* + v_j^* \geq \Phi_{ij} \qquad \text{for all } i \in I,\ j \in J.$$

Combining feasibility, individual rationality, equality on matched pairs, and the no-blocking-pair inequality, we conclude that (d^*, u^*, v^*) is a stable outcome of the original TU matching market (I, J, Φ). □

O.2 Large-Market Approximation

This appendix establishes the large-market approximation stated in Section 4.3. We compare the realized individual-level market with the aggregate planner's problem. Both converge to a limit aggregate problem. The proof first establishes uniform convergence under fixed pair-specific taxes and then applies this result to the optimal regional taxes and subsidies.

O.2.1 Market sequence and main result

Consider a sequence of markets indexed by N, where I^N is the set of applicants, J^N is the set of positions, and

$$N = |I^N| + |J^N|.$$

The finite category sets X, Y, and Z; the regional assignment $z : Y \to Z$; the systematic surplus matrix $\Phi = (\Phi_{xy})_{x \in X, y \in Y}$; and the shock distributions $(P_x)_{x \in X}$ and $(Q_y)_{y \in Y}$ are fixed across markets. Define the normalized category masses and regional bounds by

$$p_x^N = \frac{n_x^N}{N}, \qquad q_y^N = \frac{m_y^N}{N}, \qquad \underline{\rho}_z^N = \frac{\underline{o}_z^N}{N}, \qquad \bar{\rho}_z^N = \frac{\bar{o}_z^N}{N}.$$

We construct all markets on one probability space. For each applicant category x, let $(\varepsilon_{xk})_{k \geq 1}$ be an i.i.d. sequence with law P_x. For each position category y, let $(\eta_{yk})_{k \geq 1}$ be an i.i.d. sequence with law Q_y. These sequences are mutually independent. Market N contains the first n_x^N applicant draws in category x and the first m_y^N position draws in category y. Under Assumption 1, the realized surplus from matching applicant i with position j is

$$\Phi_{ij}^N = \Phi_{x(i),y(j)} + \varepsilon_{i,y(j)} + \eta_{x(i),j}.$$

Let $\mathcal{M}_N$ denote the set of one-to-one matchings before regional constraints are imposed. For $d \in \mathcal{M}_N$, define its normalized category-pair match shares by

$$\pi_N(d)_{xy} = \frac{1}{N} \sum_{i \in x} \sum_{j \in y} d_{ij}, \qquad x \in X,\ y \in Y.$$

For any nonnegative matrix π on $X \times Y$, its match share in region z is

$$\rho_z(\pi) = \sum_{x \in X} \sum_{y: z(y)=z} \pi_{xy}.$$

The regionally feasible individual matchings are

$$\mathcal{M}_N^{\text{con}} = \left\{ d \in \mathcal{M}_N : \underline{\rho}_z^N \le \rho_z(\pi_N(d)) \le \bar{\rho}_z^N \text{ for every } z \in Z \right\}.$$

The realized per-capita social surplus is

$$\mathcal{S}_N^{\text{I}}(d) = \frac{1}{N} S(\Phi^N, d).$$

The feasible match shares in the finite aggregate problem are

$$\Pi_N = \left\{ \pi \in \mathbb{R}_+^{X \times Y} : \sum_{y \in Y} \pi_{xy} \le p_x^N \text{ for every } x, \quad \sum_{x \in X} \pi_{xy} \le q_y^N \text{ for every } y \right\}.$$

The outside-option shares are

$$\pi_{x,y_0} = p_x^N - \sum_{y \in Y} \pi_{xy}, \qquad \pi_{x_0,y} = q_y^N - \sum_{x \in X} \pi_{xy}.$$

Let Π_∞ be the corresponding set with limiting masses p_x and q_y. Regional feasibility gives

$$\Pi_N^{\text{con}} = \left\{ \pi \in \Pi_N : \underline{\rho}_z^N \le \rho_z(\pi) \le \bar{\rho}_z^N \text{ for every } z \right\},$$

and Π_∞^{con} is defined analogously.

For category-level payoffs U and V, normalized so that $U_{x,y_0} = V_{x_0,y} = 0$, define

$$\bar{G}_N(U) = \sum_{x\in X} p_x^N \mathbb{E}\left[\max_{y\in Y_0}\{U_{xy} + (\varepsilon_x)_y\}\right],$$
$$\bar{H}_N(V) = \sum_{y\in Y} q_y^N \mathbb{E}\left[\max_{x\in X_0}\{V_{xy} + (\eta_y)_x\}\right].$$

The functions $\bar{G}_\infty$ and $\bar{H}_\infty$ replace p_x^N and q_y^N with p_x and q_y, respectively. For a convex function f on the real-cell payoff space $\mathbb{R}^{X\times Y}$, write

$$f^*(\pi) = \sup_a\{\langle \pi, a\rangle - f(a)\}, \qquad \langle \pi, a\rangle = \sum_{x\in X}\sum_{y\in Y} \pi_{xy} a_{xy}.$$

Define the aggregate heterogeneity term and per-capita social surplus by

$$\bar{\mathcal{E}}_N(\pi) = -(\bar{G}_N)^*(\pi) - (\bar{H}_N)^*(\pi), \quad \mathcal{S}_N^{\mathrm{A}}(\pi) = \langle \pi, \Phi\rangle + \bar{\mathcal{E}}_N(\pi),$$
$$\bar{\mathcal{E}}_\infty(\pi) = -(\bar{G}_\infty)^*(\pi) - (\bar{H}_\infty)^*(\pi), \quad \mathcal{S}_\infty(\pi) = \langle \pi, \Phi\rangle + \bar{\mathcal{E}}_\infty(\pi).$$

The effective domains of the two finite-market conjugates are

$$\begin{aligned}\operatorname{dom}(\bar{G}_N)^* &= \left\{\pi \geq 0 : \sum_y \pi_{xy} \leq p_x^N \text{ for every } x\right\},\\ \operatorname{dom}(\bar{H}_N)^* &= \left\{\pi \geq 0 : \sum_x \pi_{xy} \leq q_y^N \text{ for every } y\right\}.\end{aligned} \tag{9}$$

The same identities hold in the limit after replacing p_x^N, q_y^N with p_x, q_y. In particular, the intersection of the two effective domains is Π_N in the finite aggregate problem and Π_∞ in the limit problem.

Assumption 5. *For every $x \in X$ and $y \in Y$,*

$$\mathbb{E}\max_{y'\in Y_0}|(\varepsilon_x)_{y'}| < \infty, \qquad \mathbb{E}\max_{x'\in X_0}|(\eta_y)_{x'}| < \infty.$$

Moreover, $p_x^N \to p_x > 0$ and $q_y^N \to q_y > 0$ for every category.

Assumption 6. *For every $z \in Z$,*

$$\underline{\rho}_z^N \to \underline{\rho}_z, \qquad \bar{\rho}_z^N \to \bar{\rho}_z.$$

There are $\delta_0 > 0$ and $\pi^\circ \in \Pi_\infty$ such that

$$\underline{\rho}_z + \delta_0 \leq \rho_z(\pi^\circ) \leq \bar{\rho}_z - \delta_0 \qquad \textit{for every } z \in Z.$$

Assumption 6 requires the limiting regional constraints to admit an allocation with uniform slack. The allocation π° need not solve the planner's problem. The slack ensures feasibility in sufficiently large finite markets and bounds the regional shadow prices.

Let d_N^{I} be any solution to

$$\max_{d \in \mathcal{M}_N^{\mathrm{con}}} \mathcal{S}_N^{\mathrm{I}}(d),$$

and let w_N^{I} be any associated signed regional Lagrange multiplier. Define

$$\pi_N^{\mathrm{I}} = \pi_N(d_N^{\mathrm{I}}), \qquad W_N^{\mathrm{I}} = \mathcal{S}_N^{\mathrm{I}}(d_N^{\mathrm{I}}).$$

Let π_N^{A} be the unique solution to

$$\max_{\pi \in \Pi_N^{\mathrm{con}}} \mathcal{S}_N^{\mathrm{A}}(\pi),$$

let w_N^{A} be its signed regional shadow price, and define $W_N^{\mathrm{A}} = \mathcal{S}_N^{\mathrm{A}}(\pi_N^{\mathrm{A}})$. Choose any stable individual-level matching $d_N^{\mathrm{A,impl}}$ under w_N^{A}, and set

$$\pi_N^{\mathrm{A,impl}} = \pi_N(d_N^{\mathrm{A,impl}}), \qquad W_N^{\mathrm{A,impl}} = \mathcal{S}_N^{\mathrm{I}}(d_N^{\mathrm{A,impl}}).$$

Note that the computed aggregate allocation π_N^{A} need not equal the individual-level stable allocation $\pi_N^{\mathrm{A,impl}}$ in a finite market.

These normalized problems are the per-capita versions of the planner's problems in the main text. We use w_N^* for the selected signed optimal regional multiplier of the complete-information problem. Thus,

$$w_N^{\mathrm{I}} = w_N^*, \qquad w_N^{\mathrm{A}} = w_N^{\mathrm{AE}}.$$

In the limit aggregate problem, let π_∞ be the unique solution to

$$\max_{\pi \in \Pi_\infty^{\mathrm{con}}} \mathcal{S}_\infty(\pi),$$

let w_∞ be its unique signed regional shadow price, and define $W_\infty = \mathcal{S}_\infty(\pi_\infty)$.

In the rest of this section, we prove the following theorem:

Theorem O.2.1 (Large-market approximation)**.** *Suppose Assumptions 1–3, 5, and 6 hold. There is an event of probability one on which the following conclusions hold for every sequence of individual-market optima d_N^{I}, associated optimal signed* *regional multipliers* w_N^{I}*, and stable matchings* $d_N^{\mathrm{A,impl}}$ *under* w_N^{A}*:*

$$\begin{aligned} w_N^{\mathrm{I}} &\longrightarrow w_\infty, & w_N^{\mathrm{A}} &\longrightarrow w_\infty, & & \\ \pi_N^{\mathrm{I}} &\longrightarrow \pi_\infty, & \pi_N^{\mathrm{A}} &\longrightarrow \pi_\infty, & \pi_N^{\mathrm{A,impl}} &\longrightarrow \pi_\infty, \\ W_N^{\mathrm{I}} &\longrightarrow W_\infty, & W_N^{\mathrm{A}} &\longrightarrow W_\infty, & W_N^{\mathrm{A,impl}} &\longrightarrow W_\infty. \end{aligned}$$

All vector and matrix limits use the sup norm. In addition, if $\widehat{d}_N$ *is any stable matching under* w_N^{I}*, then*

$$\pi_N(\widehat{d}_N) \longrightarrow \pi_\infty, \qquad \mathcal{S}_N^{\mathrm{I}}(\widehat{d}_N) \longrightarrow W_\infty.$$

O.2.2 Uniform approximation under fixed pair-specific taxes

To identify every category-pair match share, we perturb surplus by a pair-specific tax matrix $\tau \in \mathbb{R}^{X\times Y}$. A regional policy $w \in \mathbb{R}^{|Z|}$ is the special case

$$\tau(w)_{xy} = w_{z(y)},$$

where $\tau_{xy} = \tau_{x'y'}$ holds for any (x, y) and (x', y') such that $z(y) = z(y')$.

Define the fixed-tax value functions by

$$\begin{aligned} \mathcal{V}_N^{\mathrm{I}}(\tau) &= \max_{d\in\mathcal{M}_N} \{\mathcal{S}_N^{\mathrm{I}}(d) - \langle \pi_N(d), \tau\rangle\}, \\ \mathcal{V}_N^{\mathrm{A}}(\tau) &= \max_{\pi\in\Pi_N} \{\mathcal{S}_N^{\mathrm{A}}(\pi) - \langle \pi, \tau\rangle\}, \\ \mathcal{V}_\infty(\tau) &= \max_{\pi\in\Pi_\infty} \{\mathcal{S}_\infty(\pi) - \langle \pi, \tau\rangle\}. \end{aligned}$$

Let $\widetilde{\Pi}_N^{\mathrm{I}}(\tau)$ be the set of category-pair match shares induced by all maximizers of the first problem. The aggregate problems have unique maximizers, denoted by $\widetilde{\pi}_N^{\mathrm{A}}(\tau)$ and $\widetilde{\pi}_\infty(\tau)$. For a fixed-tax maximizer with match shares π, gross social surplus equals net participant payoff plus net government revenue. Accordingly, define

$$\begin{aligned} \mathcal{B}_N^{\mathrm{I}}(\tau;\pi) &= \mathcal{V}_N^{\mathrm{I}}(\tau) + \langle \pi, \tau\rangle, \\ \mathcal{B}_N^{\mathrm{A}}(\tau) &= \mathcal{V}_N^{\mathrm{A}}(\tau) + \langle \widetilde{\pi}_N^{\mathrm{A}}(\tau), \tau\rangle, \\ \mathcal{B}_\infty(\tau) &= \mathcal{V}_\infty(\tau) + \langle \widetilde{\pi}_\infty(\tau), \tau\rangle. \end{aligned}$$

Proposition O.2.1. *Suppose Assumptions 1–3 and 5 hold. For every compact* *set* $T \subset \mathbb{R}^{X\times Y}$*,*

$$\begin{aligned}
&\sup_{\tau\in T} |\mathcal{V}_N^{\mathrm{I}}(\tau) - \mathcal{V}_N^{\mathrm{A}}(\tau)| \longrightarrow 0 \quad \textit{almost surely},\\
&\sup_{\tau\in T} |\mathcal{V}_N^{\mathrm{A}}(\tau) - \mathcal{V}_\infty(\tau)| \longrightarrow 0.
\end{aligned} \tag{10}$$

Moreover,

$$\begin{aligned}
\sup_{\tau\in T} \sup_{\pi\in\widetilde{\Pi}_N^{\mathrm{I}}(\tau)} \|\pi - \widetilde{\pi}_\infty(\tau)\|_\infty \longrightarrow 0 \quad \textit{almost surely},\\
\sup_{\tau\in T} \|\widetilde{\pi}_N^{\mathrm{A}}(\tau) - \widetilde{\pi}_\infty(\tau)\|_\infty \longrightarrow 0,
\end{aligned} \tag{11}$$

and

$$\begin{aligned}
\sup_{\tau\in T} \sup_{\pi\in\widetilde{\Pi}_N^{\mathrm{I}}(\tau)} |\mathcal{B}_N^{\mathrm{I}}(\tau;\pi) - \mathcal{B}_\infty(\tau)| \longrightarrow 0 \quad \textit{almost surely},\\
\sup_{\tau\in T} |\mathcal{B}_N^{\mathrm{A}}(\tau) - \mathcal{B}_\infty(\tau)| \longrightarrow 0.
\end{aligned} \tag{12}$$

The function $\widetilde{\pi}_\infty$ *is continuous.*

Proof. We prove the proposition in three steps: a common dual representation, uniform convergence of the value functions, and convergence of their subgradients.

For $\tau \in \mathbb{R}^{X\times Y}$, let

$$\mathcal{D}(\tau) = \{(U,V) : U_{x,y_0} = V_{x_0,y} = 0, \quad U_{xy} + V_{xy} \geq \Phi_{xy} - \tau_{xy} \text{ for every } x, y\}.$$

On this set, define

$$\begin{aligned}
\mathcal{L}_N^{\mathrm{I}}(U,V) &= \frac{1}{N}\sum_{x\in X}\sum_{k=1}^{n_x^N} \max_{y\in Y_0}\{U_{xy} + (\varepsilon_{xk})_y\} + \frac{1}{N}\sum_{y\in Y}\sum_{k=1}^{m_y^N} \max_{x\in X_0}\{V_{xy} + (\eta_{yk})_x\},\\
\mathcal{L}_N^{\mathrm{A}}(U,V) &= \bar{G}_N(U) + \bar{H}_N(V),\\
\mathcal{L}_\infty(U,V) &= \bar{G}_\infty(U) + \bar{H}_\infty(V).
\end{aligned}$$

Linear-programming duality and Lemma 1 give

$$\mathcal{V}_N^{\mathrm{I}}(\tau) = \min_{(U,V)\in\mathcal{D}(\tau)} \mathcal{L}_N^{\mathrm{I}}(U,V). \tag{13}$$

For the aggregate problems, replace Φ by $\Phi - \tau$ in the primal–dual pair in Section 4.2.2. Positive category masses imply that the effective domains in (9) have intersecting relative interiors for all sufficiently large N. Fenchel duality therefore gives the following representations:

$$\mathcal{V}_N^{\mathrm{A}}(\tau) = \min_{(U,V)\in\mathcal{D}(\tau)} \mathcal{L}_N^{\mathrm{A}}(U,V), \qquad \mathcal{V}_\infty(\tau) = \min_{(U,V)\in\mathcal{D}(\tau)} \mathcal{L}_\infty(U,V). \tag{14}$$

We next show that all minimizers in (13)–(14) lie in one compact set, uniformly over $\tau \in T$. Let

$$B_T = \|\Phi\|_\infty + \sup_{\tau\in T}\|\tau\|_\infty, \qquad M_+(U,V) = \max\{0, \max_{x,y} U_{xy}, \max_{x,y} V_{xy}\}.$$

Feasibility implies

$$\|(U,V)\|_\infty \le M_+(U,V) + B_T. \tag{15}$$

Indeed, if one coordinate of U were below $-M_+(U,V) - B_T$, the corresponding constraint would require the associated coordinate of V to exceed $M_+(U,V)$; the same argument applies after interchanging U and V.

Let

$$a = \frac{1}{2}\min\left\{\min_x p_x, \min_y q_y\right\} > 0.$$

If $M_+(U,V)$ is attained at U_{xy}, then the applicant-side maximum for every applicant in category x is at least $U_{xy} + (\varepsilon_{xk})_y$. All other agents can choose their outside options. The same argument applies when the maximum is attained at a coordinate of V. Assumption 5, convergence of the category masses, and the strong law of large numbers therefore give a deterministic constant $C_T < \infty$ such that, almost surely for all sufficiently large N,

$$\min\{\mathcal{L}_N^{\mathrm{I}}(U,V), \mathcal{L}_N^{\mathrm{A}}(U,V), \mathcal{L}_\infty(U,V)\} \ge aM_+(U,V) - C_T \tag{16}$$

for every $\tau \in T$ and $(U,V) \in \mathcal{D}(\tau)$. The feasible choice $U_{xy} = \Phi_{xy} - \tau_{xy}$, $V = 0$, also yields a common almost-sure upper bound for the three minima. Together, (15) and (16) imply that all minimizers lie in a deterministic compact set

$$K_T = \{(U,V) : \|(U,V)\|_\infty \le R_T,\ U_{x,y_0} = V_{x_0,y} = 0\}$$

for some $R_T < \infty$, almost surely for all sufficiently large N. Note that we can choose R_T so that this holds uniformly over $\tau \in T$.

The map

$$(U_x, e) \longmapsto \max_{y\in Y_0}\{U_{xy} + e_y\}$$

is 1-Lipschitz in U_x. On K_T, it has the integrable envelope $R_T + \max_y |e_y|$. Applying the strong law on finite nets of K_T and then letting the mesh converge to zero yields

$$\sup_{(U,V)\in K_T} |\mathcal{L}_N^{\mathrm{I}}(U,V) - \mathcal{L}_N^{\mathrm{A}}(U,V)| \longrightarrow 0 \qquad \text{almost surely.}$$

Convergence of the category masses gives

$$\sup_{(U,V)\in K_T} |\mathcal{L}_N^{\mathrm{A}}(U,V) - \mathcal{L}_\infty(U,V)| \longrightarrow 0.$$

For every τ, the three minimizations have the same feasible set $\mathcal{D}(\tau)$, and their minimizers lie in K_T. The inequality $|\inf_D f - \inf_D g| \le \sup_D |f-g|$ now proves (10).

It remains to prove convergence of match shares. Each fixed-tax value function is convex in τ, and its maximizing match shares satisfy

$$-\pi \in \partial\mathcal{V}_N^{\mathrm{I}}(\tau) \quad \text{for every } \pi \in \widetilde{\Pi}_N^{\mathrm{I}}(\tau), \qquad -\widetilde{\pi}_N^{\mathrm{A}}(\tau) \in \partial\mathcal{V}_N^{\mathrm{A}}(\tau).$$

By Lemma B.4, the positive limiting masses and Assumption 3 make $\bar{G}_\infty$ and $\bar{H}_\infty$ strictly convex. Assumption 2 makes both functions differentiable by the Williams–Daly–Zachary theorem (McFadden, 1981). These imply that the limit fixed-tax problem has a unique maximizing match matrix $\widetilde{\pi}_\infty(\tau)$.

The functions $(\bar{G}_\infty)^*$ and $(\bar{H}_\infty)^*$ are lower-semicontinuous convex functions that are finite on their compact polyhedral effective domains. A lower-semicontinuous convex function that is finite-valued on a finite-dimensional polytope is continuous relative to that polytope. Hence $\mathcal{S}_\infty$ is continuous on Π_∞. Danskin's theorem (Clarke, 2013, Theorem 10.22) gives

$$\nabla\mathcal{V}_\infty(\tau) = -\widetilde{\pi}_\infty(\tau) \tag{17}$$

for every τ. Because $\mathcal{V}_\infty$ is finite, convex, and differentiable on $\mathbb{R}^{X\times Y}$, its gradient is continuous (Clarke, 2013, Proposition 4.16).

Apply (10) to closed balls of integer radius and intersect the resulting probability-one events. On this event, both finite value functions $\mathcal{V}_N^{\mathrm{I}}$ and $\mathcal{V}_N^{\mathrm{A}}$ converge locally uniformly to $\mathcal{V}_\infty$. Suppose that either conclusion in (11) fails. Along a subsequence, there are $\tau_N \in T$ and maximizing shares π_N such that $\|\pi_N - \widetilde{\pi}_\infty(\tau_N)\|_\infty$ is bounded away from zero, where this π_N is either a maximizing match share in the individual-level market or a maximizing match share in the aggregate market. Compactness of T and boundedness of feasible match shares yield a further subsequence with $\tau_N \to \bar{\tau}$ and $\pi_N \to \bar{\pi}$.

Let $\mathcal{V}_N$ denote $\mathcal{V}_N^{\mathrm{I}}$ in the individual-market case and $\mathcal{V}_N^{\mathrm{A}}$ in the aggregate-market case. Fix an alternative tax matrix τ', and choose an integer m such that $T \cup \{\tau'\}$ is contained in the closed ball of radius m. Local uniform convergence and $\tau_N \to \bar{\tau}$ imply

$$\mathcal{V}_N(\tau') \longrightarrow \mathcal{V}_\infty(\tau'), \qquad \mathcal{V}_N(\tau_N) \longrightarrow \mathcal{V}_\infty(\bar{\tau}).$$

Taking limits in the subgradient inequality and using $\pi_N \to \bar{\pi}$ gives

$$\mathcal{V}_\infty(\tau') \geq \mathcal{V}_\infty(\bar{\tau}) - \langle \bar{\pi}, \tau' - \bar{\tau} \rangle.$$

Because this inequality holds for every τ', it follows that

$$-\bar{\pi} \in \partial \mathcal{V}_\infty(\bar{\tau}).$$

Equation (17) implies $\bar{\pi} = \widetilde{\pi}_\infty(\bar{\tau})$. Since $\tau_N \to \bar{\tau}$ and $\widetilde{\pi}_\infty$ is continuous,

$$\widetilde{\pi}_\infty(\tau_N) \to \widetilde{\pi}_\infty(\bar{\tau})\bar{\pi}.$$

Together with $\pi_N \to \bar{\pi}$, this implies

$$\|\pi_N - \widetilde{\pi}_\infty(\tau_N)\|_\infty \to 0,$$

contradicting the choice of the subsequence. This proves (11) for every maximizing selection in the individual-level market.

Finally, gross social surplus adds $\langle \pi, \tau \rangle$ to the fixed-tax value. Uniform value convergence, uniform match-share convergence, and boundedness of T therefore imply (12). □

O.2.3 Regional shadow prices and optimal policies

For a scalar r, write $r^+ = \max\{r, 0\}$ and $r^- = \max\{-r, 0\}$. For a signed regional Lagrange multiplier $w \in \mathbb{R}^{|Z|}$, define the contribution of the regional bounds to the dual objective by

$$\kappa_N(w) = \sum_{z \in Z} \left(\bar{\rho}_z^N w_z^+ - \underline{\rho}_z^N w_z^- \right),$$
$$\kappa_\infty(w) = \sum_{z \in Z} \left(\bar{\rho}_z w_z^+ - \underline{\rho}_z w_z^- \right).$$

After minimizing over participant payoffs, the regional dual objectives are

$$\begin{aligned}\Psi^{\mathrm{I}}_N(w) &= \mathcal{V}^{\mathrm{I}}_N(\tau(w)) + \kappa_N(w),\\ \Psi^{\mathrm{A}}_N(w) &= \mathcal{V}^{\mathrm{A}}_N(\tau(w)) + \kappa_N(w),\\ \Psi_\infty(w) &= \mathcal{V}_\infty(\tau(w)) + \kappa_\infty(w).\end{aligned}$$

Proposition O.2.2. *Under the assumptions of Theorem O.2.1, there are $\delta > 0$, deterministic constants $C_A, C_I \in \mathbb{R}$, and, for all sufficiently large N, a deterministic matrix $\pi^\circ_N \in \Pi_N$ and an individual matching $d^\circ_N \in \mathcal{M}_N$ such that*

$$\pi_N(d^\circ_N) = \pi^\circ_N, \qquad \underline{\rho}^N_z + \delta \le \rho_z(\pi^\circ_N) \le \bar{\rho}^N_z - \delta \quad \textit{for every } z,$$

and

$$\mathcal{S}^{\mathrm{A}}_N(\pi^\circ_N) \ge C_A, \qquad \mathcal{S}_\infty(\pi^\circ) \ge C_A. \tag{18}$$

Almost surely for all sufficiently large N,

$$\mathcal{S}^{\mathrm{I}}_N(d^\circ_N) \ge C_I. \tag{19}$$

The regional shadow prices minimize the reduced dual objectives:

$$w^{\mathrm{I}}_N \in \arg\min_w \Psi^{\mathrm{I}}_N(w), \qquad \{w^{\mathrm{A}}_N\} = \arg\min_w \Psi^{\mathrm{A}}_N(w), \qquad \{w_\infty\} = \arg\min_w \Psi_\infty(w). \tag{20}$$

The constrained optima and stable implementation satisfy

$$\begin{aligned}\pi^{\mathrm{I}}_N &\in \widetilde{\Pi}^{\mathrm{I}}_N(\tau(w^{\mathrm{I}}_N)), \quad \pi^{\mathrm{A}}_N = \widetilde{\pi}^{\mathrm{A}}_N(\tau(w^{\mathrm{A}}_N)),\\ \pi^{\mathrm{A,impl}}_N &\in \widetilde{\Pi}^{\mathrm{I}}_N(\tau(w^{\mathrm{A}}_N)), \quad \pi_\infty = \widetilde{\pi}_\infty(\tau(w_\infty)),\end{aligned} \tag{21}$$

and their social surpluses satisfy

$$\begin{aligned}W^{\mathrm{I}}_N &= \mathcal{B}^{\mathrm{I}}_N(\tau(w^{\mathrm{I}}_N); \pi^{\mathrm{I}}_N), \qquad W^{\mathrm{A}}_N = \mathcal{B}^{\mathrm{A}}_N(\tau(w^{\mathrm{A}}_N)),\\ W^{\mathrm{A,impl}}_N &= \mathcal{B}^{\mathrm{I}}_N(\tau(w^{\mathrm{A}}_N); \pi^{\mathrm{A,impl}}_N), \quad W_\infty = \mathcal{B}_\infty(\tau(w_\infty)).\end{aligned} \tag{22}$$

Finally, there is a deterministic compact set $\mathcal{W} \subset \mathbb{R}^{|Z|}$ containing every minimizer of Ψ^{I}_N and Ψ^{A}_N, almost surely for all sufficiently large N, and containing the minimizer of Ψ_∞.

Proof. We first construct a strictly feasible finite allocation. We then use the KKT conditions and assignment linear-programming duality to obtain the reduced dual characterization. The strictly feasible allocation finally yields a common bound on all optimal regional Lagrange multipliers.

Define

$$\lambda_N = \min\left\{1, \min_{x\in X}\frac{p_x^N}{p_x}, \min_{y\in Y}\frac{q_y^N}{q_y}\right\}, \qquad (\pi_N^\circ)_{xy} = \frac{\lfloor N\lambda_N\pi_{xy}^\circ\rfloor}{N}.$$

Assumption 5 implies $\lambda_N \to 1$. Rounding down preserves every row and column capacity because

$$\sum_y (\pi_N^\circ)_{xy} \le \lambda_N \sum_y \pi_{xy}^\circ \le \lambda_N p_x \le p_x^N,$$

and the analogous inequality holds for every column. The integer cell counts $N(\pi_N^\circ)_{xy}$ can be implemented by a one-to-one matching: partition the required applicants and positions by category pair and match them within each cell. Denote one such matching by d_N°. This construction gives $\pi_N(d_N^\circ) = \pi_N^\circ$. Moreover, $\|\pi_N^\circ - \pi^\circ\|_\infty \to 0$. Because the type and region sets are finite and $\pi_N^\circ \to \pi^\circ$, we have

$$\max_z |\rho_z(\pi_N^\circ) - \rho_z(\pi^\circ)| \to 0.$$

Therefore, taking, for example, $\delta = \delta_0/2$, for all sufficiently large N,

$$\underline{\rho}_z^N + \delta \le \rho_z(\pi_N^\circ) \le \bar{\rho}_z^N - \delta \quad \text{for every } z.$$

Hence the matching d_N° implementing π_N° satisfies every finite-market regional constraint with a slack independent of N.

Assumption 5 also gives the following surplus bounds. For every $\pi \in \Pi_N$, the conjugate definition and the outside-option shares imply

$$\begin{aligned} (\bar{G}_N)^*(\pi) &\le \sum_{x\in X} p_x^N \mathbb{E}\max_{y\in Y_0}|(\varepsilon_x)_y|, \\ (\bar{H}_N)^*(\pi) &\le \sum_{y\in Y} q_y^N \mathbb{E}\max_{x\in X_0}|(\eta_y)_x|. \end{aligned} \tag{23}$$

For example, the first inequality follows by completing each row with its outside-option share and using

$$p_x^N \max_y\{U_{xy} + (\varepsilon_x)_y\} \ge \sum_{y\in Y_0}\pi_{xy}\{U_{xy} + (\varepsilon_x)_y\}.$$

Because the total match share is at most one, $\langle \pi, \Phi\rangle \ge -\|\Phi\|_\infty$. Thus (23), mass convergence, and the corresponding limit inequalities give (18).

Every applicant and position contributes either a matched shock or an outside-option shock to realized surplus. Consequently,

$$\mathcal{S}_N^{\mathrm{I}}(d_N^\circ) \geq -\|\Phi\|_\infty - \frac{1}{N}\sum_{x\in X}\sum_{k=1}^{n_x^N}\max_{y\in Y_0}|(\varepsilon_{xk})_y| - \frac{1}{N}\sum_{y\in Y}\sum_{k=1}^{m_y^N}\max_{x\in X_0}|(\eta_{yk})_x|.$$

The strong law and mass convergence bound the right-hand side below by a deterministic constant almost surely for all sufficiently large N. This proves (19).

We next verify the constraint qualification for the aggregate problems. The effective-domain identity (9) shows that a matrix lies in the relative interiors of both conjugate domains when every real-cell share and every outside-option share is strictly positive. Positive category masses allow such a matrix in both the finite and limit markets. A sufficiently small convex combination of this matrix with π_N°, or with π° in the limit, retains strict regional slack. These convex combinations satisfy Slater's condition. Strong duality, dual attainment, and the KKT characterization therefore apply (Boyd and Vandenberghe, 2004, Sections 5.2.3 and 5.5.3). The uniqueness argument from Theorem 2 applies after normalization. It also applies to each unconstrained fixed-tax problem after replacing Φ by $\Phi - \tau$. Hence the constrained aggregate allocation and its regional shadow prices are unique for all sufficiently large N and in the limit, and the fixed-tax maximizers used above are unique.

To derive the reduced dual, consider the finite aggregate problem. If $\bar{\lambda}_z$ and $\underline{\lambda}_z$ are the nonnegative multipliers on the regional cap and floor, respectively, set $w = \bar{\lambda} - \underline{\lambda}$. The partial Lagrangian is

$$\begin{aligned} L_N(\pi, \bar{\lambda}, \underline{\lambda}) &= \mathcal{S}_N^{\mathrm{A}}(\pi) + \sum_z \bar{\lambda}_z\{\bar{\rho}_z^N - \rho_z(\pi)\} + \sum_z \underline{\lambda}_z\{\rho_z(\pi) - \underline{\rho}_z^N\} \\ &= \mathcal{S}_N^{\mathrm{A}}(\pi) - \langle \pi, \tau(w)\rangle + \sum_z \bar{\rho}_z^N\bar{\lambda}_z - \sum_z \underline{\rho}_z^N\underline{\lambda}_z. \end{aligned} \tag{24}$$

The strict regional slack implies $\bar{\rho}_z^N > \underline{\rho}_z^N$. Therefore, the cap and floor cannot both bind. Complementary slackness gives

$$\bar{\lambda}_z = w_z^+, \qquad \underline{\lambda}_z = w_z^-.$$

Taking the supremum of (24) over $\pi \in \Pi_N$ gives $\Psi_N^{\mathrm{A}}(w)$. Strong duality and uniqueness therefore yield $\{w_N^{\mathrm{A}}\} = \arg\min_w \Psi_N^{\mathrm{A}}(w)$. The same calculation gives $\{w_\infty\} = \arg\min_w \Psi_\infty(w)$.

The saddle-point property of the partial Lagrangian also implies that the constrained aggregate optimum maximizes the fixed-policy objective at its regional shadow price. This proves the aggregate and limit identities in (21). Complementary slackness gives

$$\kappa_N(w_N^{\mathrm{A}}) = \langle \pi_N^{\mathrm{A}}, \tau(w_N^{\mathrm{A}}) \rangle, \qquad \kappa_\infty(w_\infty) = \langle \pi_\infty, \tau(w_\infty) \rangle.$$

Adding net government revenue to participants' net payoff gives the aggregate and limit identities in (22).

For the realized individual-level planner, relax the binary assignment constraints to nonnegativity. Proposition B.1 implies that the relaxation has an integral optimum and is therefore exact. Linear-programming duality and (13) give

$$W_N^{\mathrm{I}} = \min_{w \in \mathbb{R}^{|Z|}} \Psi_N^{\mathrm{I}}(w).$$

For any primal–dual optimum, the saddle-point property and complementary slackness imply that the constrained matching maximizes the realized fixed-policy objective at its signed regional multiplier. Hence every selected optimal signed multiplier w_N^{I} and associated planner-optimal matching share π_N^{I} satisfy (20), (21), and the individual-planner identity in (22).

A stable outcome of a transferable-utility assignment game is a primal–dual optimum of its assignment linear program (Shapley and Shubik, 1971). Therefore, every individual-level stable matching under w_N^{A} maximizes the realized fixed-policy objective at $\tau(w_N^{\mathrm{A}})$. This proves the implementation identities in (21) and (22).

It remains to bound the optimal regional multipliers. Evaluating each fixed-policy value at the strictly feasible allocation constructed above gives, for every $w \in \mathbb{R}^{|Z|}$,

$$\begin{aligned} &\Psi_N^{\mathrm{A}}(w) \geq C_A + \delta \|w\|_1, \\ &\Psi_N^{\mathrm{I}}(w) \geq C_I + \delta \|w\|_1 \quad \text{almost surely for all sufficiently large } N, \\ &\Psi_\infty(w) \geq C_A + \delta \|w\|_1. \end{aligned} \tag{25}$$

For instance, when $w_z \geq 0$, the z-component remaining after adding $\kappa_N(w)$ is $(\bar{\rho}_z^N - \rho_z(\pi_N^\circ))w_z \geq \delta |w_z|$. When $w_z < 0$, it is $(\rho_z(\pi_N^\circ) - \underline{\rho}_z^N)|w_z| \geq \delta |w_z|$.

At $w = 0$, Proposition O.2.1 and (10) give one deterministic upper bound on $\Psi_N^{\mathrm{I}}(0)$, $\Psi_N^{\mathrm{A}}(0)$, and $\Psi_\infty(0)$, almost surely for all sufficiently large N. Comparing that bound with (25) places every minimizer in a common closed and bounded subset of $\mathbb{R}^{|Z|}$. This subset is the required deterministic compact set $\mathcal{W}$. □

O.2.4 Proof of Theorem O.2.1

Proof. Work on a probability-one event on which Propositions O.2.1 and O.2.2 hold for every compact tax set used below. Let $\mathcal{W}$ be the compact policy set in Proposition O.2.2. Its image

$$T_{\mathcal{W}} = \{\tau(w) : w \in \mathcal{W}\}$$

is compact. Convergence of the regional bounds implies

$$\sup_{w\in\mathcal{W}} |\kappa_N(w) - \kappa_\infty(w)| \longrightarrow 0.$$

Together with (10), this gives

$$\begin{aligned} &\sup_{w\in\mathcal{W}} |\Psi_N^{\mathrm{I}}(w) - \Psi_\infty(w)| \longrightarrow 0 \quad \text{almost surely,} \\ &\sup_{w\in\mathcal{W}} |\Psi_N^{\mathrm{A}}(w) - \Psi_\infty(w)| \longrightarrow 0. \end{aligned} \tag{26}$$

The function Ψ_∞ is continuous and has the unique minimizer w_∞. Uniform convergence in (26) therefore implies convergence of every sequence of minimizers. To verify this claim directly, suppose that a sequence of finite-market minimizers does not converge to w_∞. Compactness of $\mathcal{W}$ gives a subsequence converging to some $\bar{w} \neq w_\infty$. Finite-market optimality and uniform convergence imply $\Psi_\infty(\bar{w}) \leq \Psi_\infty(w_\infty)$, contradicting uniqueness. Equation (20) now gives

$$w_N^{\mathrm{I}} \longrightarrow w_\infty, \qquad w_N^{\mathrm{A}} \longrightarrow w_\infty.$$

By (21), the three finite-market match matrices are fixed-policy maximizers evaluated at either w_N^{I} or w_N^{A}. Uniform convergence in (11), applied on $T_{\mathcal{W}}$, and continuity of $\widetilde{\pi}_\infty$ give

$$\pi_N^{\mathrm{I}} \longrightarrow \pi_\infty, \qquad \pi_N^{\mathrm{A}} \longrightarrow \pi_\infty, \qquad \pi_N^{\mathrm{A,impl}} \longrightarrow \pi_\infty.$$

The welfare identities (22), uniform convergence in (12), and continuity of $\mathcal{B}_\infty$ similarly give

$$W_N^{\mathrm{I}} \longrightarrow W_\infty, \qquad W_N^{\mathrm{A}} \longrightarrow W_\infty, \qquad W_N^{\mathrm{A,impl}} \longrightarrow W_\infty.$$

Finally, let $\widehat{d}_N$ be any stable matching under w_N^{I}. The Shapley–Shubik char-

Table O.3.1. Environments and Outcomes of JRMP

	2017	**2018**	**2019**
Panel A: School side			
Number of medical schools	78	78	78
Number of applicants	9830	9916	9932
Number of matched residents	8530	8369	8634
Number of unmatched applicants	1300	1547	1298
Unmatched-applicant rate (%)	13.22	15.60	13.07
Panel B: Hospital side			
Number of hospitals	1025	1025	1022
Number of positions	11716	11468	11730
Number of filled positions	8530	8369	8634
Number of unfilled positions	3186	3099	3096
Unfilled-position rate (%)	27.19	27.02	26.39
Number of excess positions	1886	1551	1798
Excess rate (%)	16.01	13.52	15.33

acterization used above implies

$$\pi_N(\widehat{d}_N) \in \widetilde{\Pi}^{\mathrm{I}}_N(\tau(w^{\mathrm{I}}_N)), \qquad \mathcal{S}^{\mathrm{I}}_N(\widehat{d}_N) = \mathcal{B}^{\mathrm{I}}_N(\tau(w^{\mathrm{I}}_N); \pi_N(\widehat{d}_N)).$$

The same uniform match-share and welfare convergence therefore prove the final two conclusions of the theorem. □

O.3 Additional analysis

O.3.1 Empirical matching patterns for other years

Table O.3.1 summarizes the environment and outcomes of JRMP for 2017, 2018, and 2019. Since we use lagged matching information as controls, the estimation sample excludes 2016 and is restricted to these three years. Panel A and Panel B of Table O.3.1 report the key market environment and matching outcomes, respectively. The matching-market environment—characterized by the numbers of medical schools, applicants, hospitals, and available positions—is highly stable across the three years, with little entry or exit. Matching outcomes, including the unmatched-applicant rate, are also stable over this period. Based on these patterns, we assume that school- and hospital-side preference structures remain unchanged during the sample period.

O.3.2 Monetary-Normalization Sensitivity

For the monetary normalization, we construct candidate instruments from nearby hospitals' university-affiliation status, government-hospital status, log beds, and residency-program capacity. These characteristics are summarized by means, standard deviations, and sums over five neighborhoods: within 10, 20, and 30 km of hospital h, and within the 10–20 km and 20–30 km rings. This yields fifteen candidate instrument blocks. We estimate all non-empty combinations of these blocks, $2^{15} - 1 = 32{,}767$ specifications, using 2SLS versions of (5) separately for the school and hospital sides.

For each IV specification m, define the salary-to-utility conversion factors

$$a_U^{(m)} \equiv \hat{\gamma}_{1,U}^{(m)}, \qquad a_V^{(m)} \equiv -\hat{\gamma}_{1,V}^{(m)}.$$

The expected signs are $a_U^{(m)} > 0$ and $a_V^{(m)} > 0$. We call an IV specification admissible if it delivers the expected sign and satisfies the first-stage relevance criterion. The baseline monetary normalization used in the main counterfactual table selects, within the admissible set, the specification with the largest $a_k^{(m)}$ for each side $k \in \{U, V\}$. Since money-metric values are obtained by dividing latent-utility quantities by these salary coefficients, this rule yields the most conservative monetary magnitudes among admissible specifications.

Sensitivity analysis of counterfactual simulation results. Table O.3.2 reports how the main counterfactual comparisons vary when the salary-to-utility conversion is recomputed across admissible IV specifications. For each year, the $\mathsf{NC} - \mathsf{AC}$ comparison is total policy loss, the $\mathsf{NC} - \mathsf{OS}$ comparison is distributional-target cost, and the $\mathsf{OS} - \mathsf{AC}$ comparison is implementation cost. These entries are money-equivalent differences under the stated salary-to-utility conversion. The underlying counterfactual allocations and model-numeraire welfare objects are held fixed; only the conversion from model units into JPY is varied.

We define an IV specification as admissible if the estimated salary coefficient has the expected sign and the first-stage Wald statistic is at least 10. For each admissible specification, we recompute the money-metric values using the corresponding salary coefficient. The Main column uses the conservative normalization adopted in the main counterfactual table, while the p10, Median, and p90 columns summarize the distribution across admissible IV specifications. These percentiles should therefore be interpreted as a specification-sensitivity exercise for the monetary normalization, not as statistical confidence intervals.

The main conclusions are unchanged. Across admissible normalizations, total policy loss and implementation cost remain quantitatively large. Distributional-target cost is much smaller and more sensitive to the monetary normalization.

Table O.3.2. Sensitivity of the Money-Equivalent Policy-Loss Decomposition to IV Normalization

Year	Component	Main	p10	Median	p90
2017	Total policy loss ($\mathsf{NC}-\mathsf{AC}$)	2,680.9	16,231.9	53,010.0	296,346.2
	Distributional-target cost ($\mathsf{NC}-\mathsf{OS}$)	273.9	-10,408.1	1,224.1	34,538.0
	Implementation cost ($\mathsf{OS}-\mathsf{AC}$)	2,407.0	15,506.6	50,100.2	266,638.7
	OS government revenue	-475.4	-45,650.8	-4,855.4	-1,048.1
	OS subsidy spending	475.4	1,048.1	4,855.5	45,651.3
2018	Total policy loss ($\mathsf{NC}-\mathsf{AC}$)	3,058.6	15,990.8	54,427.9	320,198.6
	Distributional-target cost ($\mathsf{NC}-\mathsf{OS}$)	128.7	-6,910.0	418.7	17,708.4
	Implementation cost ($\mathsf{OS}-\mathsf{AC}$)	2,930.0	16,139.3	54,315.8	316,364.0
	OS government revenue	-279.3	-26,823.4	-2,852.9	-615.8
	OS subsidy spending	279.3	615.9	2,853.0	26,823.9
2019	Total policy loss ($\mathsf{NC}-\mathsf{AC}$)	3,230.2	17,709.6	59,652.8	349,252.6
	Distributional-target cost ($\mathsf{NC}-\mathsf{OS}$)	151.6	-7,598.0	534.2	20,473.6
	Implementation cost ($\mathsf{OS}-\mathsf{AC}$)	3,078.6	17,693.2	58,756.1	337,473.3
	OS government revenue	-335.9	-32,254.9	-3,430.6	-740.5
	OS subsidy spending	335.9	740.6	3,430.7	32,255.4

Notes: All entries are in millions of JPY per month. The Main column uses the conservative monetary normalization reported in the main counterfactual table. The p10, median, and p90 columns summarize the distribution obtained by recomputing the monetary conversion over all admissible IV specifications from the expanded candidate set. A specification is admissible if the salary coefficient has the expected sign and the first-stage Wald statistic is at least 10. The first three rows for each year are money-equivalent differences under the stated conversion and use the same hospital-side fiscal accounting convention as the main table. Because the components of model-numeraire social surplus are converted using different salary coefficients, an entry for money-equivalent distributional-target cost can be negative under some normalizations. Such entries do not contradict the nonnegativity of distributional-target cost in model-numeraire units. Government revenue is negative when the policy requires subsidy spending.

O.3.3 Tuned Caps Search

This subsection describes the search procedure used to construct TC in Section 7.1. The ideal cap-only benchmark would search over all hospital-level capacity vectors in the six urban prefectures, impose no taxes or subsidies, and retain only policies that satisfy the rural floors. Because this search space is high-dimensional and integer-valued, we instead search over prefecture-level capacity totals and then map each prefecture-level total into hospital-level capacities.

Let z index one of the six urban prefectures, and let $\mathcal{H}_z$ denote the set of hospitals in prefecture z. Let κ_z^{AC} and κ_z^{NC} denote total capacity in prefecture z under AC and NC. For each urban prefecture, the search considers candidate totals satisfying

$$0.7\kappa_z^{\mathsf{AC}} \leq \kappa_z \leq \kappa_z^{\mathsf{NC}}.$$

For each hospital $h \in \mathcal{H}_z$, let c_h^{AC} and c_h^{NC} denote its capacity under AC and NC. Given a candidate prefecture-level total κ_z, we assign hospital-level capacity according to the following AC-anchored rule:

$$c_h(\kappa_z) = \begin{cases} \dfrac{\kappa_z}{\kappa_z^{\mathsf{AC}}} c_h^{\mathsf{AC}}, & \text{if } \kappa_z < \kappa_z^{\mathsf{AC}}, \\[2ex] c_h^{\mathsf{AC}} + \dfrac{\kappa_z - \kappa_z^{\mathsf{AC}}}{\kappa_z^{\mathsf{NC}} - \kappa_z^{\mathsf{AC}}} \left(c_h^{\mathsf{NC}} - c_h^{\mathsf{AC}} \right), & \text{if } \kappa_z \geq \kappa_z^{\mathsf{AC}}. \end{cases}$$

Thus, capacity reductions below AC are proportional to hospitals' AC capacities, while capacity increases above AC restore each hospital's AC-to-NC capacity increase in the same proportion. In implementation, the resulting hospital-level capacities are rounded to integer positions while preserving the candidate prefecture-level total. This construction guarantees that choosing $\kappa_z = \kappa_z^{\mathsf{AC}}$ reproduces the status quo hospital-level capacity vector, while choosing $\kappa_z = \kappa_z^{\mathsf{NC}}$ restores the no-cap-tightening vector.

We construct the candidate set iteratively. The initial candidate set allows each urban prefecture to take one of three anchor values: the lower endpoint, AC, or NC. For each candidate vector κ, we solve the aggregate equilibrium without taxes or subsidies. Let $M_r(\kappa)$ denote the simulated number of matches in rural prefecture r, and let F_r denote its floor. We measure the candidate's rural-floor shortfall by

$$V(\kappa) = \max_r \max\{F_r - M_r(\kappa), 0\}.$$

A candidate is feasible if and only if $V(\kappa) = 0$. After evaluating the current candidate set, we select previously evaluated candidates as base points for local refinement. These base points include feasible candidates with high model-numeraire social surplus and infeasible candidates with small $V(\kappa)$. For each base point $\bar{\kappa}$ and refinement radius δ, we add neighboring vectors whose zth component is $\bar{\kappa}_z(1+\eta_z)$ with $\eta_z \in \{-\delta, 0, \delta\}$, subject to the capacity bounds above. We evaluate the newly added candidates, update the set of base points, and repeat this step for $\delta = 0.05$, 0.02, and 0.01. Among all candidates evaluated during this iterative search, TC is the feasible cap-based policy with the highest model-numeraire social surplus, subject to a minimum rural-floor slack of 0.01 matches.

O.3.4 Balanced Tax-and-Subsidy Search

This subsection describes the grid search used to construct BTS in Section 7.2. The rural floors define the distributional target. The urban upper bounds introduced below are auxiliary constraints used to compute urban taxes; they do not define additional distributional targets.

Let $\mathcal{Z}^R$ denote the 15 rural prefectures and let $\mathcal{Z}^U$ denote the six urban prefectures subject to the status quo cap-based policy. For each $z \in \mathcal{Z}^R$, let F_z denote the rural floor, defined as the simulated number of matches in prefecture z under AC. We begin from the no-cap-tightening capacity vector and impose these rural floors, as in OS. Let M_z^{OS} denote the simulated number of matches in urban prefecture $z \in \mathcal{Z}^U$ under OS.

For each urban prefecture $z \in \mathcal{Z}^U$, we consider four scale factors,

$$\alpha_z \in \{0.85, 0.90, 0.95, 1.00\}.$$

A grid candidate is a vector $\alpha = (\alpha_z)_{z \in \mathcal{Z}^U}$, so the search evaluates $4^6 = 4{,}096$ candidates. Given α, we set the auxiliary urban upper bound in prefecture z to

$$B_z(\alpha) = \alpha_z M_z^{\mathsf{OS}}.$$

We then solve the aggregate-level planner's problem with the no-cap-tightening capacity vector, the rural floors

$$\sum_{y \in z} \sum_{x \in X} \mu_{xy} \geq F_z \qquad (z \in \mathcal{Z}^R),$$

and the auxiliary urban upper bounds

$$\sum_{y \in z} \sum_{x \in X} \mu_{xy} \leq B_z(\alpha) \qquad (z \in \mathcal{Z}^U).$$

The Lagrange multipliers on the rural floors imply subsidies in rural prefectures, while the multipliers on the auxiliary urban upper bounds imply taxes in high-demand urban prefectures.

For each candidate α, let $M_z(\alpha)$ denote the simulated number of matches in prefecture z, and let

$$r(\alpha) = \sum_z w_z(\alpha) M_z(\alpha)$$

denote government revenue in model-numeraire units, where positive values indicate tax revenue and negative values indicate net subsidy spending. Following the conservative accounting convention used in Table 3, we convert this revenue into JPY using the hospital-side salary coefficient:

$$R^H(\alpha) = \frac{r(\alpha)}{|\beta_H|}.$$

We select the BTS candidate by minimizing the absolute hospital-side budget imbalance,

$$I^H(\alpha) = |R^H(\alpha)|.$$

The selected candidate is

$$(\alpha_{13}, \alpha_{14}, \alpha_{23}, \alpha_{26}, \alpha_{27}, \alpha_{40}) = (1.00, 1.00, 0.95, 0.95, 1.00, 1.00),$$

where prefecture codes 13, 14, 23, 26, 27, and 40 correspond to Tokyo, Kanagawa, Aichi, Kyoto, Osaka, and Fukuoka, respectively. For this candidate, the signed government revenue is $R^H(\alpha) = -1.197$ million JPY per month. The corresponding deficit magnitude is 1.197 million JPY per month, which rounds to the 1 million JPY deficit reported in Table 3.

O.3.5 Additional Estimation Results, Figures and Tables

O.3.5.1 Descriptive Sorting Regressions

Table O.3.3. Matching from Hospital Viewpoint

	(1) Public	(2) Urban Univ	(3) T-score	(4) Distance	(5) Unfilled-Position Rate
Urban Hospital	-0.172***	0.376***	0.727***	-6.822	-0.151***
	(0.017)	(0.014)	(0.093)	(9.021)	(0.015)
ln Wage	0.1006***	-0.0074	5.0446***	27.0238***	0.0600***
	(0.0099)	(0.0076)	(0.0575)	(5.0860)	(0.0100)
ln Beds	-0.0873***	0.0519***	0.0136	-27.9942***	-0.0825***
	(0.0215)	(0.0163)	(0.1244)	(10.6837)	(0.0215)
Observations	2847	2847	2847	2847	3072
Year dummy	✓	✓	✓	✓	✓

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

Cols. (1)-(4) exclude hospital-year observations with zero realized matches (weighted averages undefined).

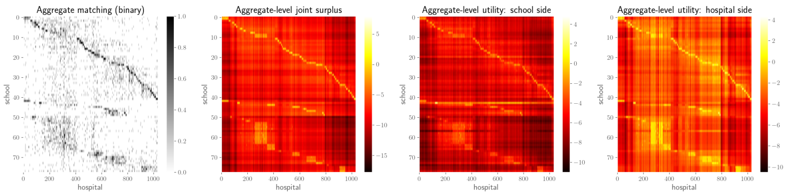
Col. (5) excludes hospital-year observations with zero capacity.

Table O.3.4. Matching from School Viewpoint

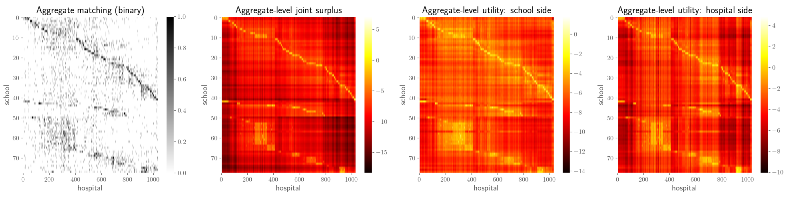
	(1) Urban Hospital	(2) ln Wage	(3) ln Beds	(4) Distance	(5) Unmatched-Applicant Rate
Public University	0.006	0.097***	-0.229***	75.221**	-0.091***
	(0.044)	(0.030)	(0.044)	(36.979)	(0.033)
Urban University	0.441***	-0.032	0.035	-6.428	-0.024
	(0.040)	(0.030)	(0.041)	(36.597)	(0.027)
T-score	0.0007	-0.0040	0.0022	-19.4853***	-0.0045
	(0.0062)	(0.0044)	(0.0063)	(5.5169)	(0.0031)
Observations	234	234	234	234	234
Year dummy	✓	✓	✓	✓	✓

Standard errors in parentheses

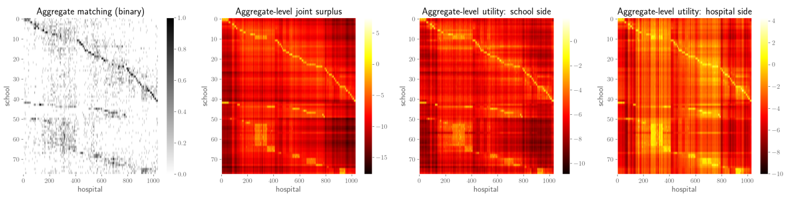
* p<0.1, ** p<0.05, *** p<0.01



(a) 2017

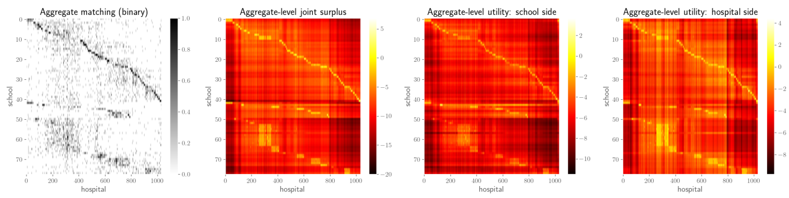


(b) 2018

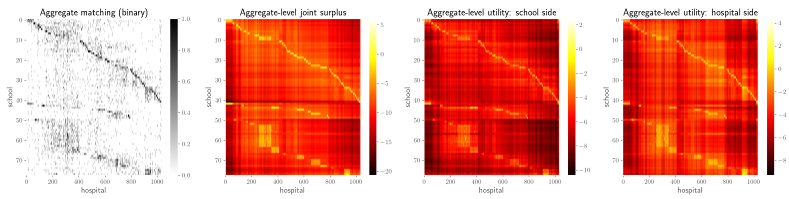


(c) 2019

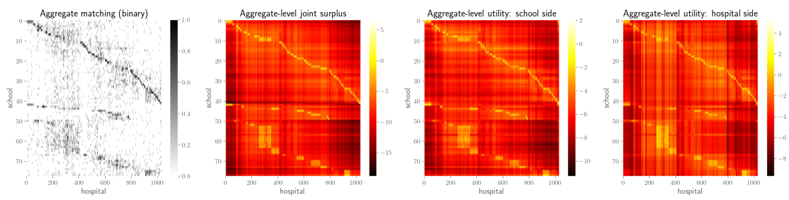
Figure O.3.1. Aggregate Matching and Estimated Aggregate-level Joint Surplus for Degree 3.



(a) 2017

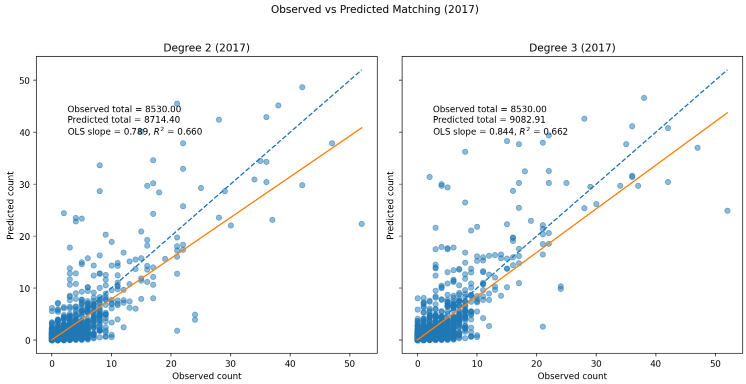


(b) 2018

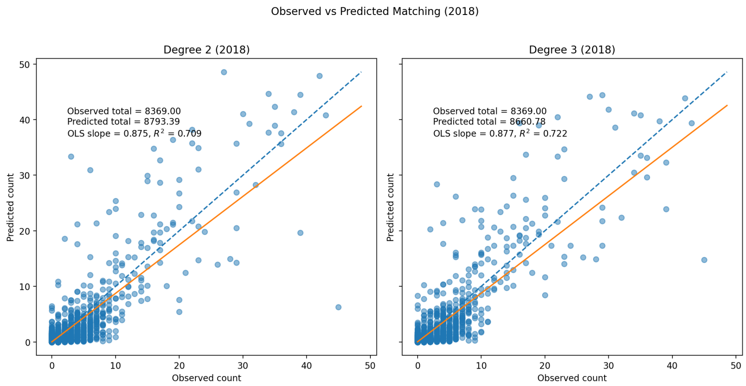


(c) 2019

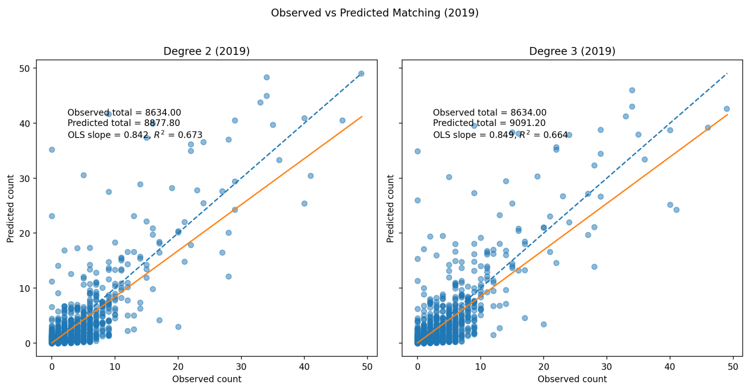
Figure O.3.2. Aggregate Matching and Estimated Aggregate-level Joint Surplus for Degree 2.



(a) 2017



(b) 2018



(c) 2019

Figure O.3.3. Observed and predicted matches by year.

Note: Each panel plots observed counts on the horizontal axis and predicted counts on the vertical axis. The dashed blue line is the 45-degree line, which represents perfect prediction. The solid orange line is the fitted line from the OLS regression of predicted counts on observed counts with a constant. “Observed total” and “Predicted total” denote the total number of matches in the observed and predicted data, respectively. The reported R^2 is the uncentered R^2 from the corresponding regression without a constant.

Table O.3.5. First-stage regressions for salary_million (selected IV sets)

	U	V
Avg. Capacity (20–30km ring)	-0.000	
	(0.000)	
Avg. Government (20–30km ring)	0.016*	
	(0.009)	
Avg. ln (# Beds) (20–30km ring)	-0.001	
	(0.001)	
Avg. University Hospital (20–30km ring)	0.049***	
	(0.015)	
Std. Capacity (0–30km)		-0.001
		(0.001)
Std. Government (0–30km)		0.044***
		(0.011)
Std. ln (# Beds) (0–30km)		0.016
		(0.015)
Std. University Hospital (0–30km)		-0.035*
		(0.020)
Sum Capacity (10–20km ring)		-0.000
		(0.000)
Sum Government (10–20km ring)		-0.002
		(0.001)
Sum ln (# Beds) (10–20km ring)		-0.000***
		(0.000)
Sum University Hospital (10–20km ring)		0.005***
		(0.001)
First-stage F	14.8	97.7
N	2847	2847
# excluded IV used	4	8

Notes: First-stage regressions correspond to each 2SLS column (same controls and year×prefecture fixed effects). The selected IV sets come from unrestricted winner selection, i.e. no requirement that a std block must be accompanied by the corresponding avg block. Rows list coefficients on the excluded instruments used after rank/zero-variance cleanup. The reported F is the partial first-stage F-statistic for the excluded instruments (linearmodels diagnostics).