Regulating Matching Markets with Distributional Constraints

KEI IKEGAMI
University of Tokyo
and
ATSUSHI IWASAKI
University of Electro-Communications
and
AKIRA MATSUSHITA
Kyoto University
and
KYOHEI OKUMURA
University of Wisconsin–Madison

Distributional constraints arise naturally in many matching markets, requiring the number of matches of specific types to satisfy predetermined bounds. This article reviews recent developments in the design and analysis of matching markets under such constraints. We discuss existing theoretical and empirical approaches. We then describe the results of [Ikegami et al. 2025], which develops a new framework for matching markets with distributional constraints and applies it to the Japan Residency Matching Program. The analysis illustrates how data can be used to evaluate regulatory instruments and to construct subsidy schemes that implement constrained-efficient outcomes.

Categories and Subject Descriptors: J.4 [**Computer Applications**]: Social and Behavioral Sciences—*Economics*

General Terms: Economics, Design, Human factors

Additional Key Words and Phrases: Matching with Distributional Constraints, Discrete Choice, Taxation, Subsidy

1. INTRODUCTION

Many matching markets are subject to distributional requirements that limit how matches can be allocated across groups or locations. These requirements are commonly implemented through rules that restrict the number of matches of particular types. Examples include affirmative action policies in college admissions, gender quotas in electoral systems, and regional caps in the Japan Residency Matching Program, which limits placements in urban hospitals to maintain adequate staffing in rural areas.

Cap-based policies are widely used as they are straightforward to implement. However, a cap is a blunt instrument that may prevent high-surplus matches. Monetary interventions, such as taxes and subsidies, offer a natural yet comparatively

Authors' addresses: ikegamikei@e.u-tokyo.ac.jp, atsushi.iwasaki@uec.ac.jp, amatsushita@i.kyoto-u.ac.jp, kyohei.okumura@wisc.edu

underexplored alternative in this context. Their potential advantage lies in their ability to account for the intensity of preferences: a well-calibrated targeted subsidy could, in principle, influence marginal participants without generating large welfare losses. Whether this theoretical potential translates into meaningful efficiency gains in practice, and under what conditions, remains a central open question for both theory and policy.

This article first reviews recent theoretical and empirical work on matching markets with distributional constraints. It then describes the framework developed in [Ikegami et al. 2025], which provides a unified approach to analyzing cap-based and monetary interventions in matching markets with distributional constraints. Applied to the Japan Residency Matching Program, the framework uses aggregate match data to quantify the effects of alternative policies and to construct subsidy schemes that implement constrained-efficient outcomes, defined as outcomes that maximize social surplus subject to the constraints.

2. EXISTING WORK ON MATCHING WITH DISTRIBUTIONAL CONSTRAINTS

2.1 Theoretical Work

Distributional objectives, particularly floor constraints, arise in many matching markets beyond the Japanese medical residency system. Rural doctor shortages have been documented in the United States [Fogarty et al. 2025], India [Alcoba 2009], Australia [Nambiar and Bavas 2010], and Korea [Chae 2025]. Similar distributional concerns also appear in other settings, such as teacher assignments to public schools in France with minimum staffing regulations [Terrier 2014] and the assignment of newly graduated cadets to U.S. military branches subject to minimum staffing requirements [Fragiadakis and Troyan 2017].

A large literature in matching theory studies mechanisms for addressing distributional imbalances, primarily in non-transferable utility (NTU) environments where agents cannot endogenously adjust transfers as part of the matching process. Early contributions typically model distributional concerns using capacity or upper-bound constraints. In school choice, [Abdulkadiroğlu and Sönmez 2003] introduce type-specific reserves that protect access for particular student types by reserving seats ex ante, without imposing ex post minimum assignment requirements. In a related vein, [Kamada and Kojima 2015; 2018] analyze the Japanese residency market by imposing regional upper bounds on matches with urban hospitals and proposing a flexible deferred acceptance algorithm that allocates limited urban capacity in response to residents' demand. Their approach encourages rural matches indirectly through urban ceilings, but it does not explicitly model lower bounds and therefore does not guarantee that minimum staffing constraints are satisfied ex post.

A distinct strand of the literature instead studies floor constraints, which impose minimum matching requirements ex post. With hard floor constraints, a matching must satisfy prescribed minimums even when demand is insufficient. [Ehlers et al. 2014] first formalize this setting and show an incompatibility between feasibility—simultaneously satisfying upper and lower bounds—and stability, defined by fairness and non-wastefulness. Motivated by this impossibility, they introduce soft floor constraints that may be violated when necessary. Building on this insight, [Fragiadakis et al. 2015] design mechanisms that offer alternative tradeoffs between fairness and non-wastefulness, while [Tomoeda 2018] provides sufficient conditions on hospital preferences under which feasible and stable matchings with floor constraints exist.

Other work focuses on mechanism design under hard floor constraints. [Goto et al. 2016] propose a strategy-proof mechanism that is non-wasteful and weakly Pareto efficient, though not necessarily stable. [Fragiadakis and Troyan 2017] allow for wasteful matchings and design a strategy-proof mechanism that endogenously adjusts ceiling constraints to achieve fairness while satisfying floor requirements, showing that doctors unanimously prefer this mechanism to one with fixed ceilings. [Akin 2021] instead weaken fairness to guarantee existence: their algorithm first runs deferred acceptance without floors and then applies a serial dictatorship to satisfy remaining minimums, yielding a strategy-proof outcome.

The literature also differs in the structure of constraints it considers. Many papers impose institution-level constraints, specifying lower and upper bounds independently for each school. By contrast, other work, including [Kamada and Kojima 2015; 2018] and [Ikegami et al. 2025], studies *regional constraints* that span multiple institutions and restrict the total number of matches across a group of schools.

Compared to the NTU model, relatively few papers study constraints and policy interventions in transferable utility (TU) matching models. [Kojima et al. 2020] and [Jalota et al. 2025] analyze the existence of equilibria under various constraints. Regarding the design of optimal taxes and subsidies, [Yokote 2020] studies a many-to-one TU matching framework [Kelso and Crawford 1982] with *interval constraints*, which impose lower and upper bounds on the number of matches at each hospital.[1] [Ikegami et al. 2025] instead studies the design of taxes and subsidies under regional constraints and extends the model to incorporate unobserved heterogeneity and to enable empirical analysis using aggregate-level matching data.

2.2 Empirical Work

A central empirical challenge is to assess how distributional constraints perform in practice, including whether they achieve their intended objectives and what welfare consequences they generate. When multiple policy instruments can be used to satisfy the same constraints, an additional question is how they compare in terms of social welfare. More generally, even in the absence of a theoretically dominant policy, it is natural to ask whether data can be used to design interventions that perform well in a given market. Addressing these issues requires empirical analysis beyond purely theoretical considerations. Yet, relative to the theoretical literature, empirical work on matching markets with constraints remains limited. We briefly review the related empirical work.

[Agarwal 2015] provides a set of policy analyses of the U.S. residency matching market. The paper applies an NTU matching framework to construct an empirical model of the medical match, which is estimated using actual matching outcomes and the rank-order lists submitted to the DA-like algorithm used in practice.

Building on the preference estimates from [Agarwal 2015], [Agarwal 2017] studies

[1] [Yokote 2020] also develops a general discrete-optimization result based on hierarchical affine constraints. While the paper applies this result only to interval constraints, it may potentially be useful for studying taxes and subsidies under other types of constraints, including those considered in [Ikegami et al. 2025].

the effects of price- and quantity-based regulations aimed at addressing geographic imbalances in the supply of medical residents. The analysis suggests that neither higher wages (price regulation) nor capacity adjustments across regions (quantity regulation) substantially increase the number of residents placed in rural locations. However, both policies affect the composition of matches by altering the distribution of resident quality across regions and improving the quality of residents assigned to rural areas. The author argues that price regulation screens residents who are relatively more willing to work in rural locations, generating welfare gains for residents that may outweigh the associated fiscal costs. The paper also analyzes a restricted class of counterfactual policies and evaluates their effects, taking policy interventions as exogenously specified. The counterfactual policies considered there are not derived from an explicit theoretical benchmark.

Many of empirical market design studies examine the effect of altering the allocation algorithm. Examples include the design of waiting lists in kidney exchange and public housing allocation [Agarwal et al. 2021; Waldinger 2021], as well as the comparison between the deferred acceptance and Boston mechanisms in school choice [Abdulkadiroglu et al. 2011]. Closer to the literature on matching with constraints, [Combe et al. 2022] study the French teacher assignment market, with a particular focus on assignments to rural schools. Across these studies, policy interventions act through changes to the matching algorithm rather than through monetary instruments.

3. TU MODEL OF MATCHING WITH REGIONAL CONSTRAINTS

In [Ikegami et al. 2025], we develop a framework to design and analyze policies in a matching market with regional constraints. Our framework accommodates both *taxation policies*, which levy taxes or subsidies on specific matches, and *cap-based policies*, which impose quantity constraints on the number of available positions. In the proposed framework, there exists a taxation policy that induces the constrained-efficient matching.

3.1 Baseline Model

We consider a one-to-one, two-sided matching market with doctors $i \in I$ on one side and job slots $j \in J$ on the other, and there are finitely many regions. The job slots are owned by hospitals, and each hospital belongs to one region. A policymaker faces *regional constraints*. The regional constraints specify lower and upper bounds on the number of matches realized in each region.

Agents form a stable outcome à la [Shapley and Shubik 1971]. Without policy intervention, the realized matching may not meet the regional constraints. One class of interventions available to the policymaker is *taxation policies*, which alter the distribution of the joint surplus among agents to satisfy the regional constraints. When a doctor i and a slot j are matched, they generate an *(individual-level) gross joint surplus* $\Phi_{ij} \in \mathbb{R}$. The tax $w_z \in \mathbb{R}$ is imposed on each match (i, j) in region z, with negative taxes being interpreted as subsidies. With taxation policy $w = (w_z)_{z \in Z}$, each matched pair divides the *net joint surplus* $\Phi_{ij} - w_{z(j)}$ instead of the gross joint surplus. A matching market is characterized by a tuple (I, J, Φ, w), and the stable outcome under a taxation policy is defined as follows:

Definition 3.1 *Stable outcome.* Given (I, J, Φ, w), a profile $(d, (u, v))$ of feasible matching $d = (d_{ij})_{i,j}$ and equilibrium payoff profiles $u = (u_i)_i$ and $v = (v_j)_j$ forms a *stable outcome* if it satisfies:[2]

(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 $d_{ij} = 1$.

The policymaker may also employ *cap-based policies*, which restrict the set of positions offered by hospitals to induce a desired allocation of applicants. 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 doctors and the remaining slots. Under a cap-based policy J', agents form a stable outcome in the induced matching market $(I, J', \Phi', 0)$. Such quantity-based interventions are motivated by the objective of redirecting applicants toward understaffed regions by limiting capacity in high-demand areas. We assume that no taxation is applied when a cap-based policy is in place, that is, $w \equiv 0$.[3]

3.2 Results for the Baseline Model

A matching d is *constrained-efficient* if it maximizes total surplus $\sum_{i,j} d_{ij}\Phi_{ij}$ subject to the regional constraints. In the baseline model, we can show that the policymaker can compute an *optimal taxation policy* w^* that implements the constrained-efficient matching as an equilibrium outcome *if she knows the joint surplus generated by each pair.*

The optimal taxation policy, characterized by Lagrange multipliers associated with the regional constraints, can be efficiently computed via linear programming. Moreover, even if the policymaker can impose pair-specific taxes, the constrained-efficient allocation can be implemented with a uniform tax within each region. This structure substantially simplifies policy design and execution.

The constrained-efficient allocation serves as a welfare benchmark. In particular, the optimal taxation policy yields a weakly higher surplus than any cap-based policy satisfying the same regional constraints, providing a quantitative basis for evaluating the welfare losses of alternative policies.

4. IMPLEMENTING OPTIMAL TAXATION POLICY

To implement the optimal taxation policy derived in the previous section, the policymaker needs to know the preferences of market participants. We can show that the taxation policy is implementable using past match data under certain structural assumptions. We illustrate how to apply the proposed method using newly collected data on the Japan Residency Matching Program.

[2] $d_{ij} = 1$ if i and j are matched; $d_{ij} = 0$, otherwise. i_0 and j_0 denote the outside options.

[3] In our subsequent analysis incorporating unobserved heterogeneity, we assume that positions within a given region are removed uniformly at random under a cap-based policy.

4.1 Model with Unobserved Heterogeneity

Let X represent the finite set of observable characteristics, or *types*, of doctors. Each doctor $i \in I$ has a type $x(i) \in X$. Similarly, let Y represent the finite set of observable characteristics of job slots, with each slot $j \in J$ having a type $y(j) \in Y$. Agents with the same type are indistinguishable to the policymaker, but there can be *unobservable heterogeneity*: doctors of the same type x or job slots of the same type y may generate different joint surpluses when matched. We assume each job slot type $y \in Y$ belongs to a unique region. Let μ_{xy} denote the number of matches between type-x doctors and type-y job slots. We call $\mu = (\mu_{xy})_{x \in X, y \in Y}$ an *aggregate-level matching*.

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

4.2 Implementing Optimal Taxation using Data

Suppose that the policymaker observes historical aggregate match outcomes $(\mu_{xy})_{xy}$. Under the identification results of [Galichon and Salanié 2021], the *aggregate joint surplus* Φ_{xy}—the analogue of individual-level surplus Φ_{ij} in the presence of unobserved heterogeneity—can be recovered under certain structural assumptions. Embedding our model in this framework, we show that the optimal taxation policy can be computed from the identified joint surplus by solving a convex optimization problem, the aggregate counterpart of the linear program in the baseline model.

To apply the results to the data and compute the optimal taxes and subsidies, we estimate the joint surplus Φ_{xy} parametrically, decomposing it into doctors' and hospitals' systematic utilities. The specification uses observable attributes of medical schools x and hospitals y, including wages, quality measures, geographic distance, and past match frequencies. Wage information is especially important, as it allows us to express surplus—and, in counterfactual exercises, taxes and subsidies—in monetary units. Because [Galichon and Salanié 2021] does not address the use of transfer data in empirical analysis, we refer to [Ikegami et al. 2025] for a detailed description of the empirical model and the estimation procedure.

We use the estimates to conduct counterfactual simulations that compare the current regulatory regime with two alternative policies. The first scenario, *Artificial Caps* (AC), replicates the allocation under the cap-based policy currently implemented in the Japan Residency Matching Program. The second, *No Caps* (NC), provides an unconstrained welfare benchmark by removing all caps and reinstating all residency positions eliminated between 2017 and 2019. The third, *Optimal Subsidy* (OS), constructs the constrained-optimal benchmark: caps are removed, and subsidies are chosen to maximize total surplus subject to distributional constraints requiring designated rural regions to receive at least the number of residents assigned under AC.

Comparing welfare across these three scenarios isolates the sources of efficiency loss. By construction, NC attains the highest surplus and AC the lowest. The difference between NC and OS reflects the welfare cost of the distributional constraints themselves, while the difference between OS and AC measures the loss due to the choice of policy instrument. The simulations indicate that the former loss is small (approximately 6 million JPY per month), whereas the latter is large (exceeding 2,600 million JPY per month).[4] Thus, the cap-based policy has limited effectiveness in redirecting residents during the 2017–2019 period, and comparable distributional goals can be achieved at substantially lower welfare cost using a targeted subsidy scheme.

5. CONCLUSION

We conclude by highlighting several directions for future research. First, the framework of [Ikegami et al. 2025] abstracts from the possibility that the policymaker can create new positions. Extending the analysis to optimal capacity planning problems that allow for both expansions and contractions would be valuable. Second, practical policy design is often constrained by budget balance, which motivates the study of optimal policies under explicit budget constraints. Finally, the implementability result in [Ikegami et al. 2025]—the computability of optimal taxation policies using aggregate match data—relies on structural assumptions introduced in [Galichon and Salanié 2021]. Evaluating the restrictiveness of these assumptions, examining how they may be relaxed with richer data, and studying optimal policy design under ambiguity about agents' preferences are important avenues for future research.

REFERENCES

Abdulkadiroglu, A., Che, Y.-K., and Yasuda, Y. 2011. Resolving conflicting preferences in school choice: The "boston mechanism" reconsidered. *American Economic Review 101,* 1, 399–410.

Abdulkadiroğlu, A. and Sönmez, T. 2003. School choice: A mechanism design approach. *American Economic Review 93,* 3, 729–747.

Agarwal, N. 2015. An Empirical Model of the Medical Match. *American Economic Review 105,* 7, 1939–1978.

Agarwal, N. 2017. Policy analysis in matching markets. *American Economic Review 107,* 5, 246–50.

Agarwal, N., Ashlagi, I., Rees, M. A., Somaini, P., and Waldinger, D. 2021. Equilibrium allocations under alternative waitlist designs: Evidence from deceased donor kidneys. *Econometrica 89,* 1, 37–76.

Akin, S. 2021. Matching with floor constraints. *Theoretical Economics 16,* 3, 911–942.

Alcoba, N. 2009. Hitting the road to alleviate india's rural doctor shortage. *Canadian Medical Association Journal 180,* 10, E34–E36.

Chae, H.-S. 2025. Rural hospitals are short of doctors and not even high pay will attract them. Accessed: 2025-12-21.

Combe, J., Tercieux, O., and Terrier, C. 2022. The design of teacher assignment: Theory and evidence. *Review of Economic Studies 89,* 6, 3154–3222.

Ehlers, L., Hafalir, I. E., Yenmez, M. B., and Yildirim, M. A. 2014. School choice with controlled choice constraints: Hard bounds versus soft bounds. *Journal of Economic Theory 153*, 648–683.

Fogarty, C. T., Byun, H., and Huffstetler, A. N. 2025. Family physician workforce trends: The toll on rural communities. *The Annals of Family Medicine 23,* 6, 535–538.

[4] The estimated total surplus under NC is approximately 65,000 million JPY per month.

Fragiadakis, D., Iwasaki, A., Troyan, P., Ueda, S., and Yokoo, M. 2015. Strategyproof matching with minimum quotas. *ACM Transactions on Economics and Computation 4*.

Fragiadakis, D. and Troyan, P. 2017. Improving matching under hard distributional constraints. *Theoretical Economics 12,* 2, 863–908.

Galichon, A. and Salanié, B. 2021. Cupid's Invisible Hand: Social Surplus and Identification in Matching Models. *The Review of Economic Studies 89,* 5, 2600–2629.

Goto, M., Iwasaki, A., Kawasaki, Y., Kurata, R., Yasuda, Y., and Yokoo, M. 2016. Strategyproof matching with regional minimum and maximum quotas. *Artificial Intelligence 235*, 40–57.

Ikegami, K., Iwasaki, A., Matsushita, A., and Okumura, K. 2025. Evaluating the efficiency of regulation in matching markets with distributional disparities. In *Proceedings of the 26th ACM Conference on Economics and Computation (EC'25)*.

Jalota, D., Ostrovsky, M., and Pavone, M. 2025. Matching with transfers under distributional constraints. *Games and Economic Behavior 152*, 313–332.

Kamada, Y. and Kojima, F. 2015. Efficient matching under distributional constraints: Theory and applications. *American Economic Review 105,* 1, 67–99.

Kamada, Y. and Kojima, F. 2018. Stability and strategy-proofness for matching with constraints: A necessary and sufficient condition. *Theoretical Economics 13,* 2, 761–793.

Kelso, A. and Crawford, V. 1982. Job matching, coalition formation, and gross substitutes. *Econometrica 50,* 6, 1483–1504.

Kojima, F., Sun, N., and Yu, N. N. 2020. Job matching under constraints. *American Economic Review 110,* 9, 2935–47.

Nambiar, M. and Bavas, J. 2010. Rudd plan won't fix rural doctor shortage: Rda. Accessed: 2025-12-21.

Shapley, L. S. and Shubik, M. 1971. The assignment game I: The core. *International Journal of Game Theory 1,* 1, 111–130.

Terrier, C. 2014. Matching practices for secondary public school teachers – france. Matching in Practice Country Profile No. 20. Accessed: 2025-12-21.

Tomoeda, K. 2018. Finding a stable matching under type-specific minimum quotas. *Journal of Economic Theory 176*, 81–117.

Waldinger, D. 2021. Targeting in-kind transfers through market design: A revealed preference analysis of public housing allocation. *American Economic Review 111,* 8, 2660–96.

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,* 1, 37–77.