

非参数估计匹配效率和弹性在一个婚姻中介平台上的应用：

2014 – 2025

Suguru Otani*

第一版：2025 年 3 月 31 日

当前版本：May 4, 2025

Abstract

本文使用日本最大的结构化匹配平台 IBJ 的新数据，考察了日本婚姻市场每月的匹配效率。与行政或约会应用程序的数据不同，IBJ 提供了基于经过验证的个人资料和确认订婚的完整搜索、约会和匹配日志。用户通过昂贵的筛选被高度筛选为认真寻找结婚对象。该数据每年涵盖全国 3.3% 的婚姻，提供了罕见的行为细节。采用非参数方法，我估计了随时间变化的匹配函数和弹性。效率随着时间的推移提高了四倍，这说明了数字中介如何在现代婚姻市场中改变伴侣搜索的方式。

关键词：匹配函数、匹配效率、匹配弹性、婚姻代理平台、约会

JEL 分类代码： J12, J10, D83

1 介绍

理解个体在现代婚姻市场中如何寻找伴侣是家庭经济学研究的核心 (Chiappori and Salanié 2023)。随着结婚率的下降和匹配技术的普及，围绕搜索效率、伴侣可获得性和市场设计的问题变得越来越突出。虽然双边搜索的理论模型层出不穷，但实证研究一直受到数据限制的影响：传统的行政记录或调查资料仅能捕捉到实际婚姻，而忽略了之前的整个搜索和匹配过程，以及那些不认

*suguru.otani@e.u-tokyo.ac.jp, Market Design Center, Department of Economics, The University of Tokyo
I thank Fuhito Kojima, Chiappori, Pierre-André, Kadachi Ye, Shangwen Li, Yusuke Ishihata, Chiro Inoue for their valuable advice. I acknowledge Shuto Fukuda and Hirokazu Tsuchiya for sharing the data and their technical and institutional knowledge in the IBJ platform. This work was supported by JST ERATO Grant Number JPMJER2301, Japan.

Declarations of interest: none

真寻找伴侣的单身人士的存在。因此，对于伴侣搜索的动力学、现代平台如何塑造匹配结果，或者现代市场运作效率的认识知之甚少。本研究通过使用来自 IBJ（日本最大的结构化婚姻匹配平台）的新数据来解决这些限制，在 2024 年该平台占据了日本约 3.3% 的结婚案例。数据集每年包含超过 10,000 个确认订婚，并允许对整个匹配过程进行行为分析，包括搜索、求婚和正式订婚。在此说明中，我们专注于记录平台上的宏观趋势作为更详细分析的第一步，将用户行为和潜在机制的深入研究留待未来的工作。

与 dating apps(Hitsch *et al.* 2010, Ong and Wang 2015, Bapna *et al.* 2016, Egebark *et al.* 2021, Rios *et al.* 2023)、speed date events(Fisman *et al.* 2006, Belot and Francesconi 2013) 和政府记录 (Chiappori *et al.* 2017) 相比，IBJ 数据具有四个独特且有利的特点。首先，数据包含丰富且经过验证的用户人口统计数据，并链接到官方记录，如扣税文件和认证健康检查结果。其次，由于有实质性的筛选和财务成本，用户被高度选择进入严肃的伙伴关系搜索。第三，该数据集记录了与确认结果相关的时间戳完整的搜索和互动行为日志，并包括关于用户在每个时间点是否在其婚姻搜索中活跃的信息。第四，数据是在个人用户行为日志层面观察到的，这使得高频率分析成为可能。重要的是，IBJ 数据的丰富性和粒度结合其一对一匹配结构，使其能够直接类比于劳动力市场匹配函数估计 (Petrongolo and Pissarides 2001)，使我们能够评估平台级别的效率和响应性。

使用此数据集，我根据 Lange and Papageorgiou (2020) 的非参数方法估计了一个匹配函数来评估平台的表现。¹ 我发现匹配效率在 2014 年至 2025 年间增加了四倍，自 2017 年起大幅增长，这与主要的平台整合相吻合。互动关系的形成也变得更加响应用户组成的变化，如不断上升的弹性系数所示。这些结果强调了数字中介在塑造现代家庭形成中的作用。与通过亲属网络或地方习俗调解的传统安排不同，像 IBJ 这样的平台提供了结构化、数据丰富的环境，在这里摩擦最小化，结果得到验证，并且行为动态可观察，从而能够获得新的实证洞察力。

2 数据

我使用了来自 IBJ 的机密数据，涵盖了 2014 年至 2025 年期间的数据。在 2024 年，IBJ 占日本所有婚姻的 3.3%，每年有超过 10,000 个订婚案例。虽然这一份额小于行政结婚登记处的记录，但 IBJ 提供了对婚前过程的独特可见性。该平台收集了经过验证的用户数据，包括年龄、收入、教育和婚姻状况，并追踪用户的完整行为序列：搜索、提议、消息传递和约会。所有记录都带有时间戳并链接到观察到的订婚结果。IBJ 用户高度被选入严肃的长期伙伴关系搜寻中。进入需要筛选以及

¹他们的方法广泛应用于劳动市场环境。参见 Otani (2024a,b) 和 Kanayama and Otani (2024)。

显著的前期费、月费和退出费，这有效地排除了非正式用户。与依赖自我报告资料和开放式互动的约会应用不同，IBJ 作为封闭式的咨询中介双面匹配平台运行，在这里双方同意将导致确认订婚。这种结构化的环境使得数据非常适合分析随时间变化的匹配效率和响应性。

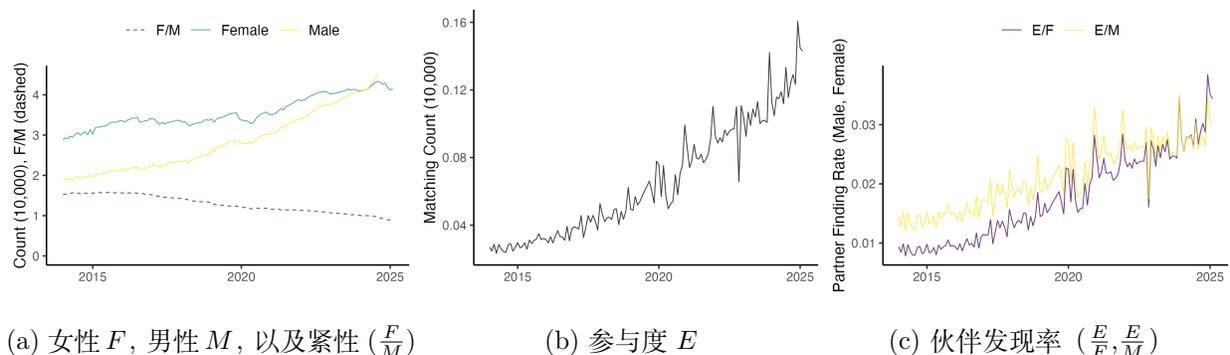


Figure 1: 关键变量的趋势 2014-2025

图 1 中的 (a) 和 (b) 展示了从 2014 年到 2025 年间 IBJ 平台上活跃用户数量及互动次数的趋势。女性参与者 F 和男性参与者 M 的数量稳步增长，但男性的增长率超过了女性，特别是在 2020 年后，这导致了市场紧俏度的逐渐下降，即性别比例平衡度 $\frac{F}{M}$ 降低。除了这一结构变化外，互动次数 E 在 2020 年前相对稳定，之后则急剧增加。这种加速可能反映了平台性能的改善，例如更好的算法匹配、用户活动的增加或数字匹配服务更广泛的社会采纳。

面板 (c) 展示了每月女性和男性用户的配对率，分别记为 $\frac{E}{F}$ 和 $\frac{E}{M}$ 。两个比率均随时间增加，女性用户始终实现更高的匹配率。这表明该平台上的匹配效率有所提高，特别是对于女性用户而言，尽管男性的供应量也在增长。男性的比率显示了更大的月度波动性，反映出其对市场条件变化的敏感程度更高。

3 经验框架

本文将婚姻市场概念化为一个双面搜索环境，其中女性用户寻找伴侣，男性用户是潜在的伴侣。² 当双方同意交往时，即视为匹配成功。借鉴劳动经济学中的匹配函数框架，我估计了约会数量作为效率输入和由性别比例平衡所捕捉的市场紧缩程度的函数。

令 E_t 表示时间 t 的参与次数， F_t 和 M_t 分别表示活跃的女性和男性用户数量， A_t 是一个随时

²虽然文献中一个重要流派将婚姻市场建模为稳定的、无摩擦的匹配环境，并使用均衡假设估计基于偏好的匹配模式 (例如, Chiappori and Salanié 2023), 本研究采取了不同的实证方法。与其在稳定匹配假设下恢复结构偏好, 我关注的是平台层面的匹配效率和弹性随时间变化的实证测量。我们的兴趣在于理解 IBJ 平台如何将搜索输入转化为实际匹配。

间变化的匹配效率参数。令 (A, F, M) 表示由时间 t 下标对应的实际随机变量。假设在给定 F 的条件下， M 和 A 是独立的，也就是说 $M \perp A|F$ 。我假设一个形式为 $E_t = m(A_t F_t, M_t)$ 的匹配函数，具有不变规模报酬。给定这些假设，通过应用非参数识别结果 [Matzkin \(2003\)](#) 来避免功能形式假设，命题 [Lange and Papageorgiou \(2020\)](#) 中的命题 1 表明，参与度、女性用户和男性用户的观测联合分布，记为 $G(E, F, M)$ ，非参数地确定了匹配效率和女性参与度的联合分布 $H(A, F)$ ，以及匹配函数 $m(AF, M) : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$ ，直到在支持 (A, F, M) 的给定点 A_0 处对效率参数 A 进行归一化。

4 结果

使用 [Lange and Papageorgiou \(2020\)](#) 提出的非参数方法，我估计了随时间变化的匹配效率 A_t ，以及相对于有效女性搜索输入 $A_t F_t$ 和男性用户 M_t 的弹性。³

4.1 匹配效率和弹性在婚姻平台中的应用

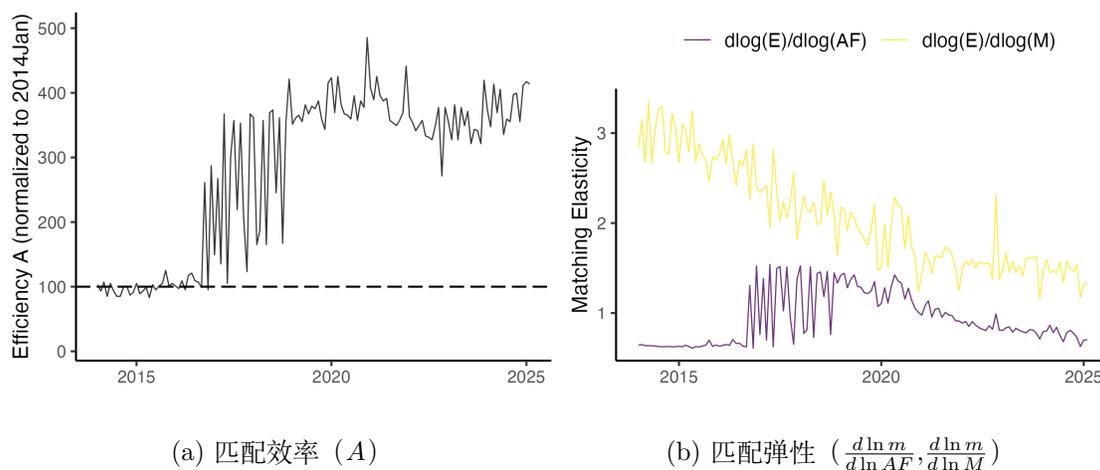


Figure 2: IBJ 平台 2014-2025

注意：我将 2014 年 1 月的匹配效率标准化为 100。

在图 2 中，匹配效率从 2014 年的基线水平增加了四倍以上，特别是自 2017 年起加速增长，并延续到 2020 年后。值得注意的是，2017 年至 2019 年间，效率几乎翻了三倍。这种显著的改进可能部分归因于平台内的结构变化，包括其他主要婚姻代理机构合并入 IBJ 网络。例如，在此期间将另一家全国性婚介咨询服务 ZWEI 整合到 IBJ 集团中，很可能扩大了有效用户基础，并改善了跨平台协调。这些发展可能增强了匹配机会的广度和推荐机制的复杂性，从而为配对形成效率带来了显著

³细节见附录。有限样本性能通过蒙特卡罗模拟在 [Otani \(2024b\)](#) 中展示。

提升。女性输入的弹性范围从 0.5 到 1.2，而男性输入的弹性通常超过 1.0，这表明该平台已经变得更加响应市场双方的变化。

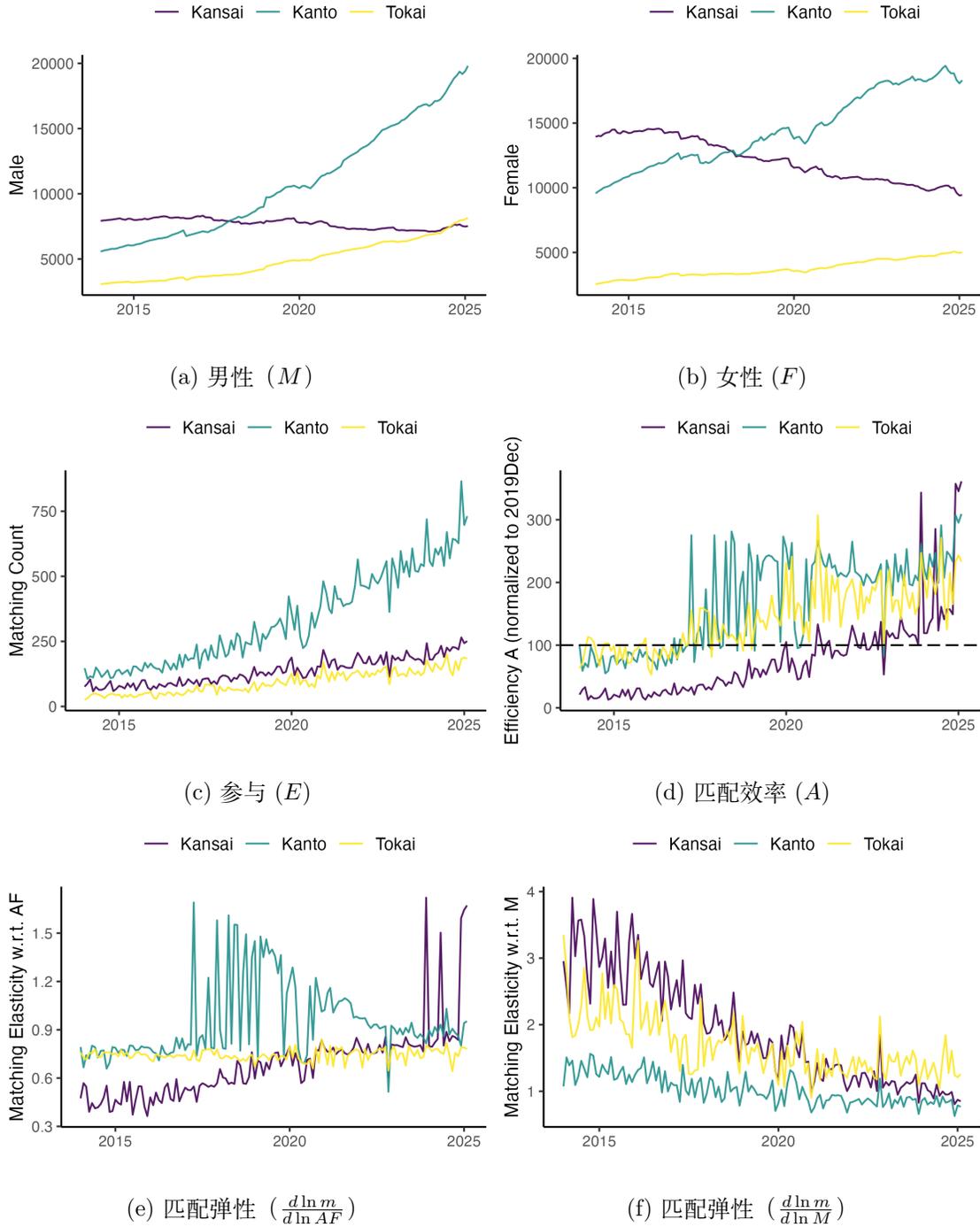


Figure 3: IBJ 平台 2014-2025

注意：我将 2014 年 1 月东京的匹配效率标准化为 100。

这些趋势共同表明，IBJ 平台不仅在产生互动方面变得更加高效，而且对用户基础双方的变化也更加敏感。效率和弹性的提高暗示了潜在匹配过程的功能得到了改进，这可能是由更好的推荐算法、更大的用户基数或更精准的用户参与策略驱动的。

4.2 地区匹配效率和弹性差异

我也探讨了日本三大地区——关东（东京）、近畿（大阪）和东海（爱知）在图 3 中的区域差异。所有地区的匹配效率与 2014 年 1 月相比都有所提高，关东地区一直领先。各地区的弹性大致相似，表明全国范围内的平台功能稳定。总体而言，IBJ 平台在主要城市地区的匹配效率不断提高，且弹性模式保持稳定，这反映了平台层面的改进和区域特定用户增长。

5 结论

本文记录了在日本一个大型婚姻平台上匹配效率的重大改进。利用经过验证的用户数据和明确界定的参与结果，我采用非参数方法来估计随时间变化的匹配函数。效率和弹性估算表明，该平台变得更加有效，并且对用户动态更加敏感。区域分析证实主要城市地区的功能一致。这些发现强调了数字匹配平台通过结构化中介和算法改进重塑传统伴侣搜索市场的潜力。

References

- BAPNA, R., RAMAPRASAD, J., SHMUELI, G. and UMYAROV, A. (2016). One-way mirrors in online dating: A randomized field experiment. *Management Science*, **62** (11), 3100–3122.
- BELOT, M. and FRANCESCONI, M. (2013). Dating preferences and meeting opportunities in mate choice decisions. *Journal of Human Resources*, **48** (2), 474–508.
- CHIAPPORI, P.-A. and SALANIÉ, B. (2023). Mating markets. In *Handbook of the Economics of the Family*, vol. 1, Elsevier, pp. 49–109.
- , SALANIÉ, B. and WEISS, Y. (2017). Partner choice, investment in children, and the marital college premium. *American Economic Review*, **107** (8), 2109–2167.
- EGEBARK, J., EKSTRÖM, M., PLUG, E. and VAN PRAAG, M. (2021). Brains or beauty? causal evidence on the returns to education and attractiveness in the online dating market. *Journal of Public Economics*, **196**, 104372.
- FISMAN, R., IYENGAR, S. S., KAMENICA, E. and SIMONSON, I. (2006). Gender differences in mate selection: Evidence from a speed dating experiment. *The Quarterly Journal of Economics*, **121** (2), 673–697.
- HITSCH, G. J., HORTAÇSU, A. and ARIELY, D. (2010). Matching and sorting in online dating. *American Economic Review*, **100** (1), 130–163.
- KANAYAMA, H. and OTANI, S. (2024). Nonparametric estimation of matching efficiency and elasticity in a spot gig work platform: 2019-2023. *arXiv preprint arXiv:2412.19024*.
- LANGE, F. and PAPAGEORGIOU, T. (2020). *Beyond Cobb-Douglas: flexibly estimating matching functions with unobserved matching efficiency*. Tech. rep., National Bureau of Economic Research.
- MATZKIN, R. L. (2003). Nonparametric estimation of nonadditive random functions. *Econometrica*, **71** (5), 1339–1375.
- ONG, D. and WANG, J. (2015). Income attraction: An online dating field experiment. *Journal of Economic Behavior & Organization*, **111**, 13–22.

- OTANI, S. (2024a). Nonparametric estimation of matching efficiency and elasticity on a private on-the-job search platform: Evidence from japan, 2014-2024. *arXiv preprint arXiv:2410.17011*.
- (2024b). Nonparametric estimation of matching efficiency and mismatch in labor markets via public employment security offices in japan, 1972-2024. *arXiv preprint arXiv:2407.20931*.
- PETRONGOLO, B. and PISSARIDES, C. A. (2001). Looking into the black box: A survey of the matching function. *Journal of Economic literature*, **39** (2), 390–431.
- RIOS, I., SABAN, D. and ZHENG, F. (2023). Improving match rates in dating markets through assortment optimization. *Manufacturing & Service Operations Management*, **25** (4), 1304–1323.

A 在线附录（不用于出版）

A.1 估计详情

我首先估计分布函数 $F(A_0 | F)$ ，遵循 [Lange and Papageorgiou \(2020\)](#) 中的识别策略。逻辑依赖于给定女性和男性用户数量 (F, M) 的参与条件分布 E 。形式上：

$$F(A_0 | \psi F_0) = G_{E|F,M}(\psi E_0 | \psi F_0, \psi M_0),$$

$$F(\psi A_0 | \lambda F_0) = G_{E|F,M}(\psi E_0 | \lambda F_0, \psi M_0),$$

其中 ψ 是一个任意标量，而 λ 是缩放因子。通过改变 (ψ, λ) ，我在 (A, F) 的支持上追踪出 $F(A | F)$ 。

在实践中，使用有限的的数据，我构造了一个非参数估计器 $G_{E|F,M}$ 。对于任何评估点 (E_τ, F_τ, M_τ) ，我计算位于 (F_τ, M_τ) 附近且参与度少于 E_τ 的观测值的比例，在 (F, M) 空间中。核权重会降低远处点的影响。该估计器表示为：

$$F(\psi A_0 | \lambda F_0) = G_{E|F,M}(\psi E_0 | \lambda F_0, \psi M_0),$$

$$\hat{F}(\psi A_0 | \lambda F_0) = \sum 1(E_t < \psi E_0) \cdot \kappa(F_t, M_t; \lambda F_0, \psi M_0),$$

其中 $\kappa(\cdot)$ 是一个带宽设置为 0.01 的二元正态核函数。

一旦恢复了分布 $F(A | F)$ ，我使用观察到的参与情况将其反演，以检索每个时间段的隐含匹配效率 A_t 。

$$A_t = F^{-1}(G(E_t | F_t, M_t) | F_t).$$

使用恢复的 A_t 值，然后反转匹配函数：

$$m(A_t F_t, M_t) = G^{-1}(F(A_t | F_t) | F_t).$$

最后，我通过回归 E 与女性和男性用户数量以及估计的有效搜索输入 AF 的交互作用来计算

匹配函数的局部弹性。这使我们能够估计参与度对女性和男性参与变化的边际响应：

$$\text{Elasticity w.r.t. } F : \frac{d \log m(AF, M)}{d \log F} = \frac{d \log m(AF, M)}{d \log AF},$$

通过估计匹配函数关于 AF 的导数，并乘以比率 $\frac{AF}{E}$ ，使用经验规范中的回归系数获得。

A.2 区域间的流动性

图 4 记录了随着时间推移，居住在同一地理区域的用户之间形成的匹配比例（“区域内匹配”）。面板 (a) 展示了日本本州岛主要地区的趋势，而面板 (b) 则展示了岛屿地区的情况。在所有地区中，大多数互动发生在同一地区的用户之间，这表明在寻找伴侣的过程中跨区域流动性有限。像关东和近畿这样的本州地区一直保持着超过 70-80% 的区域内匹配率，而像中国和东北这样较小或人口密度较低的地区则显示出更大的波动性，尤其是在样本早期年份。

相比之下，像冲绳和北海道这样的岛屿地区显示出更高的波动性和更大比例的跨区域匹配，尤其是在早期年份，这可能反映了本地合作伙伴池更加受限。随着时间的推移，岛屿地区的区域内匹配率也倾向于趋同于本州岛水平，表明本地匹配能力增加或搜索行为更加地理筛选。这些模式表明，尽管平台具有数字化特性，但物理地理仍然在塑造匹配结果方面发挥着重要作用。

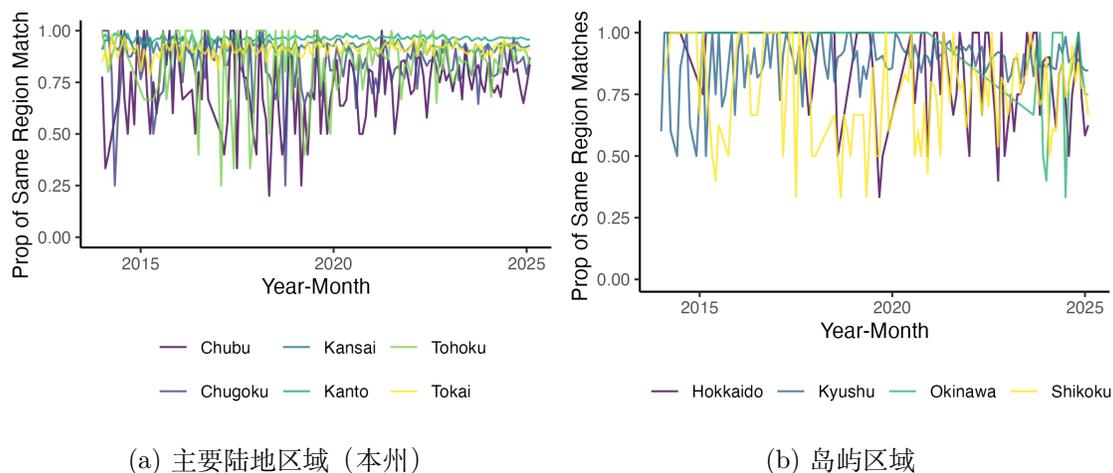


Figure 4: 区域内婚姻比例