Top of the Batch: Interviews and the Match

Federico Echenique[*] Ruy González[†] Alistair Wilson[‡] Leeat Yariv[§][¶]

May, 2021

Abstract

Most doctors in the NRMP match with one of their most-preferred internship programs. However, surveys indicate doctors' preferences are similar, suggesting a puzzle: how can so many doctors match with their top choices when positions are scarce? We provide one possible explanation. We show that the patterns in the NRMP data may be an artifact of the interview process that precedes the match. Our study highlights the importance of understanding market interactions occurring before and after a matching clearinghouse. It casts doubts on analyses of clearinghouses that take reported preferences at face value.

Key words: NRMP, Deferred acceptance, Interviews, First-rank matches

JEL: C78, D47, J44

---

[*] Division of the Humanities and Social Sciences, Caltech.

[†] Caltech.

[‡] Department of Economics, University of Pittsburgh.

[§] Department of Economics, Princeton University, CEPR, and NBER.

[¶] Echenique gratefully acknowledges the support of NSF SES-1558757 and CNS-1518941. Yariv gratefully acknowledges the support of NSF grant SES-1629613. Detailed comments from SangMok Lee, Juan Pereyra, the coeditor, and four anonymous reviewers were very helpful in revising and improving the paper. A suggestion from SangMok Lee allowed us to generalize our main result.

1 Introduction

The National Resident Matching Program (NRMP) has matched millions of doctors to residency programs across the United States. In 2020 alone, 45,000 active applicants attempted to match with just over 37,000 positions. Match results reported by the NRMP for 2020 suggest comforting news for doctors: 46% of freshly-minted MDs from US schools were matched to their first-ranked choice, while 71% were matched to one of their top-three choices. The 2020 figures are by no means an aberration. The fraction of applicants matched to their first-ranked choice has been at least as high over the past two decades. We suggest these figures should not be taken at face value. In particular, we show that interactions outside of the main match—the interview process that precedes it—may be at least as important as the match itself.

Why should a large fraction of doctors matching to their first-ranked residency be surprising? The algorithm governing the NRMP match implements a stable matching over reported preferences. If applicants report similar preferences, only a few applicants can get their most-preferred option. For example, suppose 100 prospective residents are matched to 100 positions. Common preferences on both sides (an assortative market) yield an outcome in which only *1%* are matched to their first-ranked program. As we show, even a small common component in doctors' preferences implies relatively few matches to top-ranked hospitals.

One explanation for the NRMP outcomes is that applicants' preferences are diametrically opposed, with only a handful ranking each position as their top outcome. This stands in the face of survey data and estimation results suggesting the importance of preference commonalities (Rees-Jones, 2018; Agarwal, 2015). Another explanation might be that preferences are similar across participants, but that each doctor and hospital consider only their top-k partners acceptable, as in Immorlica and Mahdian (2015). Matched participants would then have to receive a top-k outcome. As we demonstrate, this explanation too has shortcomings. First, it does not explain the relative prevalence of matches with first-ranked partners. Second, for k small enough to generate an effect, many applicants would be unmatched (see Arnosti, 2015; Beyhaghi and Tardos, 2018; Lee, 2016).

We propose another story. In the months preceding the centralized match, applicants submit their biographical and academic records, personal statements, and letters of recommendation to hospitals.[1] From these applications, the hospitals select a group to interview. The process used to determine who interviews with whom is decentralized with two important features. First, applications and interviews are costly—both market sides have limited capacity. Second, hospitals and doctors submit rankings to the NRMP only for those *they interviewed with*.[2]

We assume that hospitals' and applicants' preferences are decomposable into common and idiosyncratic components. For hospitals, the common component may reflect doctors' academic performance, test scores, and the strength of their letters (Agarwal, 2015). For doctors, it may reflect hospital rankings, quality of life in the local area, etc. In contrast, the idiosyncratic component reflects match-specific values. Using this model of preferences we consider the pre-match interview-selection process. Each candidate has a limit on the number of interviews they can attend, k, while each hospital has a maximum number of interview slots they can offer, k'. The decentralized interview process is then modeled as a stable many-to-many matching under the (k,k') capacity constraints. In the centralized matching stage, participants report rankings over interview partners only—their "interview-truncated" preference.

Truncation induced by interviewing necessarily narrows agents' original preferences. However, unlike the truncation to the top-k, it is endogenous. Participants' preferences are linked through the stability of interviews, so a large fraction of prospective doctors still end up matched—indeed, in our simulations, participants often end up with the same exact partner as in the centralized match with untruncated preferences. Moreover, the *reported* ranks for match outcomes are greatly inflated.

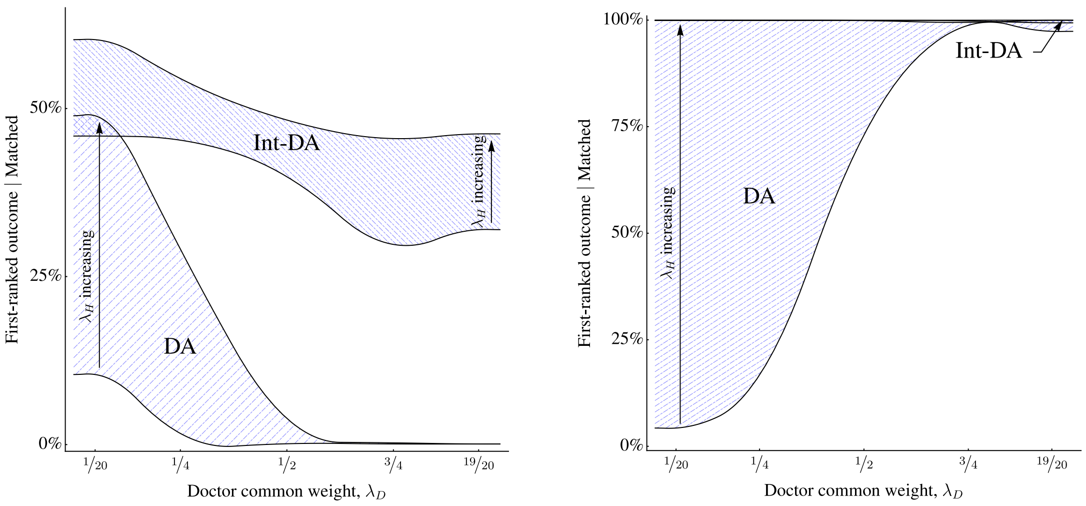
The presence of a common component in the preferences of prospective residents and hospitals is crucial for this conclusion. We show that with sufficient disagreement in doctors' preferences, interviews may cause matched partners' reported rank to *go down*, not up. While perfect agreement among doctors over hospital rankings will clearly lead to inflated rankings for matched programs under interview truncation, this obviously represents an extreme.[3] Our main theoret-

---

[1] See `https://www.nrmp.org/applying-interviewing-residency-programs/`

[2] The 2019 NRMP Applicant Survey (available from `nrmp.org`) provides median respondent data across four applicant-types and 21 medical-specialties. Of the 84 medians reported, 63 have perfectly coincident numbers for interviews attended and programs ranked, where 81 are ±1.

[3] A related idea appears in Beyhaghi and Tardos (2018), who show that interviews may increase the size of a match. See also Kadam (2015).



(a) First-ranked outcome

(b) Unique stable partner

Figure 1: Simulated matched-resident outcomes (NRMP scale)

Note: Simulations under both DA and Int-DA algorithms are used to assess outcomes across many market sizes N. A composite response across 35,704 positions is then imputed across the speciality sub-markets (ranging in size from 22–9,127 positions) listed in the 2020 NRMP report (Table 13).

ical finding is that, in large markets, an *arbitrarily weak* common-component is sufficient for interviews to generate the pattern of high-reported ranks for match partners.

As our most-general result is asymptotic, we complement it with an array of simulations at more-moderate market sizes. Using the listed positions for the sub-markets that residents match within (anesthesiology through vascular surgery), we use our simulations to infer an aggregate-level outcome at the NRMP scale. We illustrate the results in Figure 1, serving both as a guide for the puzzle that motivates us, and the paper's main result.

Figure 1 illustrates two measures for which data is available on NRMP outcomes: (a) The fraction of matched residents getting the hospital at the very top of their rank-order list; (b) Uniqueness, measured as the fraction of doctors with a unique stable match partner given the submitted rank-order lists. The figure illustrates simulated outcomes under both deferred acceptance with the full preferences (the region labeled *DA*) and interview-truncation (the region labeled *Int-DA*). In both figures we sweep across an array of possible preferences that modify the strength of the common component. On the horizontal axis we indicate the effect of increasing the common component weight within doctor's preferences,

λ_D (with a residual weight $1-\lambda_D$ on the idiosyncratic component). The width of each region then indicates the range in effect as we shift the common-component weight in hospital's preferences, λ_H (where an arrow indicates the direction).

As mentioned, across two decades of annual NRMP matches, approximately one half of all matched residents obtain their first-ranked outcome.[4] Figure 1(A) illustrates that, while the high fraction matched to their first-ranked outcome is possible under deferred acceptance, it requires a very particular form of preference: hospitals with a heavy weight on the common-value (λ_H close to one) and doctors with entirely idiosyncratic preferences (λ_D close to zero). For all other preferences, our simulations indicate the fraction of first-ranked matches is considerably lower. In particular, this statistic is close to zero whenever doctors place substantial weight on the common preference component.

While hospitals having a strong common component is consistent with NRMP survey data, the requirement that doctors' preferences are almost completely idiosyncratic contradicts ample survey evidence.[5] Simulation results under interview-truncation in the region labeled *Int-DA* illustrate our explanation. Our model of interviews leads to a near-opposite result to deferred acceptance: Except for a very particular and empirically unlikely preference—hospitals (doctors) with a small (large) common-value component—the NRMP-scale simulations lead to an approximately half of the residents obtaining their first-ranked outcome.

In Figure 1(B) we focus on another documented feature of the NRMP rankings: almost every doctor has a unique stable match partner. Using proprietary rank-order data from the NRMP, Roth and Peranson document small cores, where doctor-proposing and hospital-proposing deferred acceptance produce the same partner for 99.9% of the residents. Our simulation results in Figure 1(B) indicate the type of preferences that can generate this level of uniqueness within DA. The result that almost every participant has a unique stable match partner is a generic feature across all preference weights under the interview-truncated rankings (this is the close-to-degenerate region labeled at the very top of the figure). In contrast, this empirical feature is only possible under the full ranking when one of the two

---

[4] See Figure B.1 in Online Appendix B for the time-series.

[5] From the NRMP Director's Survey, hospitals clearly place substantial weight on common features: test scores, recommendation letters, etc. However, the resident survey has common-value components ("reputation of program," having an "academic medical center program," as well as quality of the residents, faculty, and educational curriculum) cited at similar frequencies to potentially idiosyncratic ones ("perceived goodness of fit") as the reasons for their ranking of programs.

sides' preferences are primarily driven by a common component.

The idea that doctors' reports in the residency match may not reflect true preferences is certainly present in other work. Hassidim, Marciano, Romm, and Shorrer (2017) survey evidence of misreports in the NRMP, suggesting four possible explanations: proposers' failure to identify the dominant strategy, mistrust in the mechanism, non-classical utility, and self-selection. The last of these is closest to ours. In this vein, Chen and Pereyra (2019) consider school-choice problems where students "self select" by only ranking schools they believe will plausibly admit them, showing evidence for this self-selection in Mexican high-school applications. While doctors and hospitals ranking only those they interview with is a manifestation of self-selection, our theoretical analysis offers a constructive process, shedding light on its underpinnings and impacts.[6]

Our results have important implications for the NRMP, and the matching literature more broadly. Doctors participating in the deferred-acceptance algorithm underlying the match have incentives to truthfully report preferences (Roth and Peranson, 1999). Traditionally, economists have viewed the NRMP as an ideal case-study in strategy-proof design. Our findings make clear that because reported preferences in the NRMP are filtered through the interview stage, they should be interpreted with caution. In particular, reported high-rank matches cannot be read literally, and any conclusions drawn about welfare using estimated preferences from the match itself are suspect. This message is particularly stark given that our paper ignores strategic effects at the interview stage.[7] That said, our approach also indicates some constructive ways forward and suggests the potential importance of accounting for interactions preceding centralized clearinghouses.

---

[6] Lee and Schwarz (2017) also consider an interview process that precedes a centralized match. In their setting, workers are fully informed of their preferences, while firms view workers symmetrically at the outset and use costly interviews to infer their own preferences. In the NRMP context, Rees-Jones (2018) uses surveys to illustrate doctors' significant "misreporting" in the match. Rees-Jones and Skowronek (2018) use an online experiment with post-match residents, where 23% misrepresent their preferences in an incentivized NRMP-like matching task.

[7] See Beyhaghi, Saban, and Tardos (2017) for an analysis of some strategic implications of interviews and Manjunath and Morrill (2021) for the effects of changes in the interviewing costs on doctors' welfare.

2 The Model

Our model is a variant of the standard two-sided matching model (see, for example, Roth and Sotomayor, 1990), with an added interview stage.

2.1 Basic Definitions

A *market* is a triple (H, D, U), where: H is a finite set of *hospitals*; D is a finite set of *doctors*; and $U = ((u_d)_{d \in D}, (u_h)_{h \in H})$ is a *utility function* profile (with $u_d : H \cup \{d\} \to \mathbf{R}$ and $u_h : D \cup \{h\} \to \mathbf{R}$ for each d and h).

A utility u_a induces an ordinal preference $\succeq_a$ over the relevant set of alternatives, where we assume throughout that the resulting ordinal preferences are *strict*. The *rank-order* of b in u_a is one plus the number of b' with $u_a(b') > u_a(b)$—a lower rank-order indicates a better ordinal outcome/higher ranking. In particular, agent a's most-preferred match partner has rank-order 1. An agent b is *unacceptable* for a if $u_a(a) > u_a(b)$.

A *matching* is a function $\mu : H \cup D \to H \cup D$ with the properties that $\mu(h) \in D \cup \{h\}$, $\mu(d) \in H \cup \{d\}$, and $\mu(d) = h$ *iff* $\mu(h) = d$. A matching μ is *stable* for a market (H, D, U) if $u_a(\mu(a)) \geq u_a(a)$ for all $a \in D \cup H$, and there is no $(d, h) \in D \times H$ with $u_d(h) > u_d(\mu(d))$ and $u_h(d) > u_h(\mu(h))$.

A *many-to-many* matching is a function $\mu : H \cup D \to 2^{H \cup D}$ with the properties that $\mu(d) \subseteq H$, $\mu(h) \subseteq D$, and $h \in \mu(d)$ *iff* $d \in \mu(h)$. When an agent a is unassigned, we have $\mu(a) = \emptyset$. Given a pair of positive integers (k, k'), a many-to-many matching μ is *pairwise stable* for (k, k') if

- $\left|\mu(d)\right| \leq k$ and there is no $h \in \mu(d)$ with $u_d(h) < u_d(d)$;
- $\left|\mu(h)\right| \leq k'$ and there is no $d \in \mu(h)$ with $u_h(d) < u_h(h)$;
- There is no (h, d) such that $d \notin \mu(h)$ and any one of the following:
    - $u_d(h) > u_d(h')$ and $u_h(d) > u_h(d')$ for some $(h', d') \in \mu(d) \times \mu(h)$;
    - $u_d(h) > u_d(h')$, $u_h(d) > u_h(h)$, and $\left|\mu(h)\right| < k'$ for some $h' \in \mu(d)$;
    - $u_d(h) > u_d(d)$, $u_h(d) > u_h(d')$, and $\left|\mu(d)\right| < k$ for some $d' \in \mu(h)$.

2.2 Interview Schedules

In our model, doctors and hospitals first schedule interviews and then participate in the match.

An *interview schedule* is a many-to-many matching. Given a pair of integers (k, k'), a (k, k')-*constrained interview schedule* is a many-to-many matching μ with $\left|\mu(d)\right| \leq k$ and $\left|\mu(h)\right| \leq k'$ for all d and h. Each doctor can interview with at most k hospitals, and each hospital can interview at most k' doctors.

Given an interview schedule μ, agents' *interview-truncated* preferences are determined by setting $u_a(b) < u_a(a)$ for all $b \notin \mu(a)$. That is, interview-truncated preferences rank all interviewed agents as in the original preferences, and set all other agents as unacceptable.

The timing in our model is: (i) An interview schedule is determined as the doctor-optimal many-to-many (k, k')-stable matching;[8] (ii) Doctors and hospitals report their interview-truncated preferences as inputs into doctor-proposing DA. The outcome is the doctor-optimal stable matching on the interview-truncated preferences. We term this two-step process Int-DA: the **Int**erview process followed by **D**eferred **A**cceptance.

A doctor-optimal interview schedule can be found algorithmically using the "T-algorithm" (see Blair, 1988; Fleiner, 2003; Echenique and Oviedo, 2006). We assume it is the result of a decentralized interview scheduling process. While modeling explicitly the interview process would certainly be desirable, our focus is on the tension between a "pure" application of DA, and one that is preceded by interviews. Assuming a stable outcome at the interview stage provides us with a simple, tractable model.[9] A richer model might allow for some information to be transmitted at the interview stage. This could further inflate rank differences for two reasons.[10] First, some doctors and hospitals may learn from interviews that they are not acceptable to one another, thereby limiting further the set of ranked

---

[8] Arguably, the doctor-optimal stable matching at the interview stage yields a smaller difference between reported and actual ranks than other selections of stable matchings.

[9] In one-to-one matching markets, experimental evidence suggests decentralized interactions yield stable outcomes at high rates, see Echenique and Yariv (2013). Melcher, Ashlagi, and Wapnir (2018) propose a stable-matching algorithm for internship interviews. For more on the theory of many-to-many matching, see Sotomayor (1999); Konishi and Ünver (2006).

[10] Indeed, while simulations where the idiosyncratic component is only revealed during interviews yield qualitatively similar results, quantitatively they do increase the effects (see Figure B.6 in Online Appendix B).

participants. Second, participants would not discover others whom they would, with full information, rank very highly but that they did not interview with.

We denote the final matching from Int-DA as μ^I. We compare μ^I to the matching obtained from the doctor-proposing DA algorithm using agents' original preferences, μ^{DA}.

3 The Impacts of Interviews

We start by illustrating that when doctors' preferences are misaligned, interviews alone cannot explain the findings in the data: Int-DA does *not* necessarily yield better-ranked partners in submitted preferences. We then show that some preference alignment, in the presence of interviews, leads to patterns similar to those in the data.

3.1 Preference Misalignment

To illustrate that interviews can have the opposite effect when doctors preferences are misaligned, consider a matching market with three doctors, $\{d_1, d_2, d_3\}$, and four hospitals, $\{h_1, h_2, h_3, h_4\}$.[11] Hospitals' preferences are common: they all prefer d_1 to d_2, d_2 to d_3, and d_3 to staying unmatched. Doctors rank all hospitals as acceptable, with preferences given by (first to last):

$$\begin{aligned} d_1 &: \quad h_1, h_3, h_2, h_4; \\ d_2 &: \quad h_2, h_3, h_1, h_4; \\ d_3 &: \quad h_3, h_1, h_4, h_2. \end{aligned}$$

Under DA, d_i matches to h_i. So the rank-order of d_3's match is 1.

Suppose interview constraints are $k = k' = 2$. All doctors want to interview with h_3, but only d_1 and d_2 are able to. The resulting interview schedule is: d_1 with h_1 and h_3; d_2 with h_2 and h_3; and d_3 with h_1 and h_4.

Given the interview-truncated preferences, d_i matches with h_i for $i = 1, 2$, but d_3 is matched with h_4. The Int-DA rank-order of d_3's match is therefore 2, so the presence of interviews leads her to a strictly lower rank-order than under DA

---

[11] It is straightforward to construct more intricate examples with equal numbers of doctors and hospitals.

without interviews. Furthermore, the outcome under Int-DA is unstable for the original preferences.

In this example, there is substantial disagreement between doctors' preferences—there are no pairwise comparisons of hospitals $\{h_1, h_2, h_3\}$ on which doctors agree. In contrast, our discussion of the NRMP data emphasized the role of common components in the doctors' and hospitals' preferences. We now show that some degree of agreement between doctors' ranking of hospitals rules out such examples, and interviews can explain observed high match ranks.

3.2 Aligned Preferences

We start with the extreme case where doctors' preferences are common.

Proposition 1. *Suppose $k = k'$ and that doctors' preferences are identical. For any doctor d, the rank-order of $\mu^I(d)$ in her interview-truncated preference is always weakly lower than the rank-order of $\mu^{DA}(d)$ in her actual preference $\succeq_d$.*

The proof appears in Online Appendix A. Intuitively, when doctors' preferences are common, only one of the doctors under DA is matched to the highest-ranked hospital, one to the second-highest, etc. In particular, $n-k$ doctors are matched to a hospital ranked below their top k. In contrast, interviews allow for presorting of doctors to hospitals they have a chance of matching with. Interviews also limit how low a matched hospital can be ranked in the reported preferences: it can never be lower than k.

The proposition assumes $k = k'$ mainly for expository reasons. In our main result below we allow the two capacities to differ.

3.3 Large Markets

We now demonstrate that perfect preference alignment is not necessary. As long as there is a common-value component in agents' preferences, *however small*, the message of our first result holds in large markets.

We expand the model to account for market size and randomly generated preferences. For each n, let $E_n = (D_n, H_n, U_n, k_n, k'_n)$ denote a market, where $D_n = \{d_1, \ldots, d_n\}$, $H_n = \{h_1, \ldots, h_n\}$, each utility function is randomly drawn with a common-value and idiosyncratic component, and k_n (k'_n) is the limit on doctors' (hospitals') interviews. Specifically, suppose that

$$u_d^n(h) = \lambda_D c_h + (1-\lambda_D)\eta_{d,h} \text{ and } u_h^n(d) = \lambda_H c_d + (1-\lambda_H)\eta_{h,d},$$

for all $d \in D_n$ and $h \in H_n$, where $\lambda_D, \lambda_H \in (0,1)$. Moreover, assume that $u_a^n(a) = 0$. The common-value components c_h and c_d are crucial for our results, but need not dominate doctors' utilities: $\lambda_D, \lambda_H > 0$ may be arbitrarily small.

Suppose that c_h, c_d, $\eta_{d,h}$, and $\eta_{h,d}$ are all drawn from absolutely continuous distributions with support on a convex subset of $\mathbf{R}_+$. Let μ_n^I denote the matching resulting from the Int-DA process in the n-sized market E_n, and μ_n^{DA} the corresponding outcome of the doctor-proposing DA; these matchings are random and depend on realized utilities. The Int-DA procedure determines a matching μ_n^I by choosing a (k_n, k_n')-constrained interview schedule $\hat{\mu}$ as the doctor-optimal many-to-many stable matching, followed by the doctor-proposing DA using the induced preferences.

Proposition 2. *Let $E_n = (D_n, H_n, U_n, k_n, k_n')$ be a sequence of markets satisfying our assumptions. Fix an integer sequence $M_n \geq 1$, and let D_n' be the set of doctors $d \in D_n$ in market E_n for whom the rank of $\mu_n^I(d)$ in their interview-truncated preferences is M_n positions above the rank of $\mu_n^{DA}(d)$ in their true preference. If k_n and M_n are $o(n)$, then, for any $\theta > 0$,*

$$\lim_{n\to\infty} P\left(\frac{1}{n}\left|D_n'\right| \geq 1-\theta\right) = 1.$$

The proof of Proposition 2 appears in the Appendix. The idea underlying it is simple. Consider DA, let $\varepsilon > 0$, and fix $M_n = 1$. By Lee (2016), when n is large, with high probability, the set $A_n(\varepsilon, (c, \eta))$ of doctors that are within ε of their "target" assortative utility in DA accounts for at least $1 - \theta/2$ of all doctors. Let $B(c_d, n)$ be the event that fewer than k_n hospitals provide doctor d a utility greater than d's target utility. We denote by β_n the probability that a fraction of at least $\theta/2$ doctors have a "small" number (at most k_n) of hospitals above their target utility. We show that for n large enough, $\beta_n < \pi/2$, and by Lee (2016), $P\left(\frac{1}{n}\left|A_n(\varepsilon,(c,\eta))\right| \geq 1-\theta/2\right) > 1 - \pi/2$. Thus, the event that $B(c_d, n)$ is false for a fraction $\geq 1 - \theta/2$ of doctors *and* $\frac{1}{n}|A_n(\varepsilon,(c,\varepsilon))| \geq 1 - \theta/2$, has probability $\geq (1-\pi/2) + (1-\pi/2) - 1 = 1 - \pi$. At the intersection of these conditions, for a fraction $\geq (1-\theta/2) + (1-\theta/2) - 1 = 1 - \theta$ of $d \in D_n$, we have that $B(c_d, n)$ is false and $d \in A_n(\varepsilon)$. Hence, for a fraction $\geq 1 - \theta$ of $d \in D_n$, there are more than k_n hospitals above their target utility, and they are within ε of their target utilities.

Convergence rates for the large-market result in Proposition 2 are modest, with (poly-)logarithmic or polynomial growth in the relevant "approximation guarantees" θ and π. In words, the market size needed for Proposition 2 does not grow too quickly with the approximation guarantees. This message complements the simulations in Section 4, which assume (arguably) realistic market sizes, and can be formalized as follows (a proof is in Online Appendix A):

Proposition 3. *Let* E_n *and* D'_n *be as in the statement of Proposition 2. Fix* $\theta, \delta \in (0,1)$. *Then* $P\left(\frac{1}{n}|D'_n| \geq 1-\theta\right) \geq 1-\delta$ *for* $n = \Theta((\ln(1/\pi))^4)$ *as* $\pi \to 0$, *and* $n = \Theta((1/\theta)^4)$ *as* $\theta \to 0$.

4 Simulations

Our theoretical findings raise three important questions. The first regards market size. Proposition 2 is asymptotic, and it is natural to consider whether interviews matter for smaller, more realistic, market sizes. The second question regards unmatched agents. One might worry that interview-truncated preferences give rise to large numbers of unmatched participants, beyond those observed in the NRMP. The final question regards stability. Ideally, the difference between outcomes under DA and the interview-truncated DA procedure would be small.

We address these questions using extensive numerical simulations across the preference parameters λ_D and λ_H. In total, we examine 275 different parameterizations, representing approximately 1 million simulated participants. We also consider a variety of robustness checks. Section B.2 in Online Appendix B provides a summary and more-detailed results. To compactly display our main results, Table 1 aggregates our findings across balanced markets (N doctors and N positions) corresponding to the many smaller sub-markets that compose the NRMP.[12] We examine three algorithmic solutions:[13]

---

[12] For each pair (λ_D, λ_H) and simulation proportion f_N, we fit a linear model for $\log(f_N/(1-f_N))$ against $\log(N)$, where goodness of fit for this approach is shown in Figure B.4 in Online Appendix B. We predict the proportion $\hat{f}(N_j)$ for each sub-market of size N_j listed in Table 13 of the 2020 NRMP report (from 22 positions for Pediatrics/Medical Genetics to 9,127 for Internal Medicine). We then aggregate across these sub-markets to form a composite NRMP measure for f over the $N_T = 35{,}704$ positions as $\sum_j \frac{N_j}{N_T} \hat{f}(N_j)$.

[13] To assess robustness, we also conducted simulations examining: (i) imbalanced markets; (ii)

- Doctor- and hospital-proposing deferred acceptance (DA).
- The stable interview allocation with $k = k' = 5$ slots per position, followed by both doctor- and hospital-proposing deferred acceptance on the interview-truncated preferences (Int-DA).[14]
- Doctor- and hospital-proposing deferred acceptance on preferences truncated to the $k = k' = 5$ top-ranked options (Tr-DA).

Table 1 provides NRMP-scaled simulation outcomes for six different (λ_D, λ_H) pairs. Simulations with Tr-DA were added to distinguish the pure effect of truncation from the interview process our paper focuses on.[15] The first panel in the table provides three characteristics of the match outcome: (i) the fraction of *unmatched* participants; (ii) the fraction of doctors matched to their *first-ranked program*; and (iii) the proportion of doctors matched to a *top-three–ranked program*.

Because our simulated markets have the same participant volume on each side, with all possible matches acceptable, the benchmark for DA with full preferences predicts no unmatched doctors. In contrast, the NRMP data indicates that 5.8% of US seniors are unmatched. The first result from our simulations in Table 1 illustrates that the two-stage Int-DA process leads to a similar unmatched rate as the NRMP. Doctors in our simulations are unmatched after the Int-DA process at a 5.5% rate. Moreover, this proportion does not change substantially with either market size or the common weight. In contrast, a direct truncation to the top-five participants on the other side leads to substantially more unmatched participants. Moreover, the unmatched rate grows sharply with increases to N and λ.

The next pair of results from the Int-DA simulations again match the NRMP data: a large fraction of doctors are matched to top-ranked hospitals. Looking to NRMP data from the past five years, 48% (73%) of US MD Seniors are matched to their first-ranked (top-three–ranked) program. The Int-DA simulations indicate

---

smaller/larger interview capacities; (iii) more extreme values for λ_D and λ_H; (iv) an alternative interview selection method where hospitals only used the doctor's common components. Online Appendix B presents details on these simulations. The main effects of the interview stage are similar across these exercises.

[14] To mirror the theory section, we select the doctor-optimal stable interview schedule.

[15] In simulations where we allowed for information to be realized at the interview stage, hospitals ranked doctors purely on the common component. The idiosyncratic component is observed at the interviews, and used in the DA stage. Results are similar to the Int-DA procedure (see Figure B.6 in Online Appendix B) and suggest that simpler procedures that incorporate both sides preferences can also be used.

Table 1: Simulation Outcomes scaled to NRMP size

	$\lambda_H = 1/4$			$\lambda_H = 3/4$		
	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$
Panel A: Matching outcomes						
	Unmatched [DA: 0.0%, NRMP: 5.4%]†					
Int-DA	6.0%	6.4%	8.1%	8.2%	6.5%	5.6%
Tr-DA	26.1%	72.1%	96.0%	27.1%	71.7%	96.2%
	First-ranked program [NRMP: 48.1%]†					
DA	2.5%	0.2%	0.1%	26.5%	3.0%	0.2%
Int-DA	43.5%	38.7%	32.5%	49.1%	43.7%	41.4%
Tr-DA	22.7%	4.1%	0.2%	31.2%	5.2%	0.2%
	Top-three–ranked program [NRMP: 73.2%]†					
DA	7.4%	0.5%	0.3%	48.4%	8.1%	0.6%
Int-DA	81.6%	79.6%	75.0%	81.6%	81.5%	81.2%
Tr-DA	55.0%	15.1%	1.3%	59.1%	17.2%	1.3%
Panel B: Core size, similarity to DA, and stability						
	Same partner under proposer change | Matched [NRMP: 99.9%]‡					
DA	41.9%	90.0%	99.5%	99.3%	99.2%	98.8%
Int-DA	99.9%	99.9%	100.0%	100.0%	99.9%	99.8%
	Identical partner to DA | Matched					
Int-DA	73.8%	82.6%	79.8%	80.1%	74.6%	81.0%
	Proportion blocking programs in Int-DA					
Matched	0.1%	0.6%	1.0%	0.1%	0.7%	1.8%
Unmatched	9.1%	7.3%	8.3%	19.2%	24.6%	34.3%

Note: Table uses an estimated logit-model for each (λ_D, λ_H) pair to form an estimate at NRMP scale (~35,000 positions) using the sub-market sizes listed for the 2020 NRMP. See Online Appendix for tables at separate N †–Average for US MD Seniors in 2016–20. Source: *Results and Data: 2020 Main Residency Match*, Table 15, available from nrmp.org. ‡–Figure reported for main NRMP match in Roth and Peranson (1999). Smaller thoracic surgery market ($N \simeq 120$) has a 99.6 percent unique match for five reported years in 1991–96 (*ibid*, tables 1 and 3).

similarly-sized effects to the observed NRMP figures, at 40–50% (75–82%).[16] In contrast, the pure DA algorithm on the full preference lists implies substantially lower rates of top-ranked outcomes—matches to the first-ranked program are only found at substantial rates when hospitals have strong common components, and doctors are almost fully idiosyncratic (as depicted in Figure 1).

In the second panel of Table 1 we turn to other observed match outcomes. These outcomes are not part of our explanation of reduced match ranks, but serve to evaluate the empirical relevance of our interviews model. The first outcome is motivated by Roth and Peranson's (1999) finding that NRMP data exhibit small cores. Using NRMP ranking data from the 1990s, they examine the change in outcomes moving from the doctor- to the hospital-proposing DA. They find that 99.9% of doctors receive the same outcome, implying a unique stable partner. In the *same partner under proposer change* rows we mirror this exercise. Our DA simulations get close to the NRMP figure only with heavy weights on the common component. While most participants across each of the simulations do have a unique stable partner, the minority with multiple partners are at least an order of magnitude larger than in Roth and Peranson (1999). However, changing the proposing side over the interview-truncated rankings from Int-DA indicates much-closer effects to the NRMP field study.

Our simulations of the Int-DA procedure show that it can reproduce stylized results reflective of the observed NRMP figures—over unmatched rates, over the fraction of first-ranked outcomes, and over the small cores found in rank-order list data. Moreover, the Int-DA process does so generically, across market sizes and the common-preference weights.[17]

Given the fit with observed data regularities, a natural question regards the difference between outcomes under Int-DA and standard DA. The final set of results in Table 1 speaks to this question.

The *identical partner to DA* row directly contrasts the Int-DA and DA match outcomes. Our results illustrate that the large majority of matched doctors (~75–80%) in the Int-DA procedure are matched to the *exact same partner* they would

---

[16] The Int-DA fraction matched to their first-ranked program does increase slightly as we increase N (see Tables B.1–B.6 in Online Appendix B) and as we decrease λ_D/increase λ_H.

[17] We also examined imbalanced markets, with an excess of doctors. The Int-DA results for these simulations are qualitatively similar. However, we note that DA results allowing imbalance exhibit more frequent unique stable matchings Ashlagi, Kanoria, and Leshno (per 2017) (see Tables B.7 and Figure B.3 in Online Appendix B).

have matched with under DA with full preferences reports.[18] While four of every five doctors are entirely unaffected by the interview process, one in five being affected is clearly far from negligible.[19]

In the last section of Table 1 we evaluate the effects on stability. For each doctor in our simulations we calculate the proportion of programs they form a blocking pair with. We report the average proportion, distinguishing between matched and unmatched doctors. Matched doctors exhibit some instability, despite both stages in the two-stage process being chosen to select stable outcomes. Averaging across parameterizations, a blocking pair is detected for matched doctors 0.7% of the time. Unsurprisingly, instabilities are more substantial for unmatched doctors. A randomly chosen hospital yields a blocking pair between 8% and 34% of the time for each unmatched doctor, depending on the parameterization.

5 Conclusion

Much of the matching literature has focused on the centralized clearinghouse governing the match of newly-minted doctors and residency positions. We illustrate the possibility that *decentralized* interactions preceding the match—namely, interviews—may dramatically impact ultimate outcomes.

For the NRMP, our results imply that empirical estimations based on preferences submitted to the clearinghouse should not be taken at face value. More broadly, beyond the NRMP, our paper suggests that interactions outside of the clearinghouse can have dramatic effects on outcomes.

Our model focuses on the role interviews can have purely in terms of selection. In our setting, there is no constructive role for information acquisition during the interviews (see, however, Figure B.6 in Online Appendix B, for one model variant allowing for information transmission during the interviews). Our algorithmic approach offers a base for future research to explore richer interactions preceding matching clearinghouses.

---

[18] In a series of robustness exercises we also examine balanced-market simulations with different values of k (see Tables B.8–B.11 and Figure B.5 in Online Appendix A), indicating that the DA and Int-DA outcomes become more similar as k increases. Nonetheless, we still find that, among matched doctors, ∼75% get the same exact match as under DA even when $k = 2$.

[19] For those matched to distinct partners under Int-DA and DA, matches tend to be close in ranking terms. See Figures B.7 and B.7 in Online Appendix A for illustration.

References

Agarwal, Nikhil (2015), "An empirical model of the medical match." *The American Economic Review*, 105, 1939–78.

Arnosti, Nick (2015), "Short lists in centralized clearinghouses." In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, EC '15, 751, Association for Computing Machinery, New York, NY, USA, URL `https://doi.org/10.1145/2764468.2764533`.

Ashlagi, Itai, Yash Kanoria, and Jacob D Leshno (2017), "Unbalanced random matching markets: The stark effect of competition." *Journal of Political Economy*, 125, 69–98.

Beyhaghi, Hedyeh, Daniela Saban, and Eva Tardos (2017), "Effect of selfish choices in deferred acceptance with short lists." MATCH-UP.

Beyhaghi, Hedyeh and Eva Tardos (2018), "Two-sided matching with limited interviews." Mimeo, Cornell University.

Blair, Charles (1988), "The lattice structure of the set of stable matchings with multiple partners." *Mathematics of operations research*, 13, 619–628.

Chen, Li and Juan Sebastián Pereyra (2019), "Self-selection in school choice." *Games and Economic Behavior*, 117, 59 – 81, URL `http://www.sciencedirect.com/science/article/pii/S0899825619300843`.

Echenique, Federico and Jorge Oviedo (2006), "A theory of stability in many-to-many matching markets." *Theoretical Economics*, 1, 233–273.

Echenique, Federico and Leeat Yariv (2013), "An experimental study of decentralized matching." California Institute of Technology working paper.

Fleiner, Tamás (2003), "A fixed-point approach to stable matchings and some applications." *Mathematics of Operations Research*, 28, 103–126.

Hassidim, Avinatan, Déborah Marciano, Assaf Romm, and Ran I. Shorrer (2017), "The mechanism is truthful, why aren't you?" *American Economic Review*, 107, 220–24, URL `https://www.aeaweb.org/articles?id=10.1257/aer.p20171027`.

Immorlica, Nicole and Mohammad Mahdian (2015), "Incentives in large random two-sided markets." *ACM Trans. Econ. Comput.*, 3, URL `https://doi.org/10.1145/2656202`.

Kadam, Sangram V (2015), "Interviewing in matching markets." Mimeo Harvard University.

Konishi, Hideo and M Utku Ünver (2006), "Credible group stability in many-to-many matching problems." *Journal of Economic Theory*, 129, 57–80.

Lee, Robin S. and Michael Schwarz (2017), "Interviewing in two-sided matching markets." *The RAND Journal of Economics*, 48, 835–855, URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/1756-2171.12193`.

Lee, SangMok (2016), "Incentive compatibility of large centralized matching markets." *The Review of Economic Studies*, 84, 444–463.

Manjunath, Vikram and Thayer Morrill (2021), "Interview hoarding." ArXiv:2102.06440.

Melcher, Marc L., Itai Ashlagi, and Irene Wapnir (2018), "Matching for Fellowship Interviews." *JAMA*, 320, 1639–1640, URL `https://doi.org/10.1001/jama.2018.13080`.

Rees-Jones, Alex (2018), "Suboptimal behavior in strategy-proof mechanisms: Evidence from the residency match." *Games and Economic Behavior*, 108, 317–330.

Rees-Jones, Alex and Samuel Skowronek (2018), "An experimental investigation of preference misrepresentation in the residency match." *Proceedings of the National Academy of Sciences*, 115, 11471–11476, URL `https://www.pnas.org/content/115/45/11471`.

Roth, Alvin E and Elliott Peranson (1999), "The redesign of the matching market for american physicians: Some engineering aspects of economic design." *American economic review*, 89, 748–780.

Roth, Alvin E and Marilda Sotomayor (1990), *Two-sided Matching: A Study in Game-theoretic Modeling and Analysis*. Econometric Society Monographs volume 18, Cambridge University Press.

Sotomayor, Marilda (1999), "Three remarks on the many-to-many stable matching problem." *Mathematical social sciences*, 38, 55–70.

A Proof of Proposition 2

Let $k_n'' = k_n + M_n$ and note that k_n'' is $o(n)$.

With some notational abuse, we drop λ^D and $1-\lambda^D$, and write c_d for $\lambda^D c_d$, $\eta_{d,h}$ for $(1-\lambda^D)\eta_{d,h}$, etc. This re-scaling implies that utilities are sums of the common and private value components: $u_d^n(h) = c_h + \eta_{d,h}$, and $u_h^n(d) = c_d + \eta_{h,d}$. The probability distributions are re-scaled correspondingly, but remain absolutely continuous. Without loss of generality, we assume distributions have compact support (otherwise, choose a compact set that accumulates large enough probability). Moreover, we take the support to be $[0,1]$.

Let $D = \cup_n D_n$ and $H = \cup_n H_n$. Consider tuples (c,η), with $c = (c_a)_{a\in H\cup D}$ and

$$\eta = ((\eta_{a,b})_{(a,b)\in H\times D}, (\eta_{a,b})_{(a,b)\in D\times H}).$$

The tuples (c,d) are endowed with the product probability measure from the i.i.d. distributions described.

Let G denote the cumulative distribution function corresponding to c_d. Fix $\theta,\pi \in (0,1)$. Choose $c^\star$ and $\varepsilon,\delta \in (0,1)$ such that $1 - G(c^\star) + \delta < \theta/4$ while $0 < P(c_h + \eta_{d,h} > c^\star + 1 + \varepsilon)$. These choices are possible due to the absolute continuity of the distributions of c_d, c_h, and $\eta_{d,h}$. Write $p(c^\star)$ for $P(c_h + \eta_{d,h} > c^\star + 1 + \varepsilon)$.

If agents match assortatively based on the common component, a doctor d should be able to find a hospital h for which it has idiosyncratic utility close to 1, and this hospital should provide d with (approximately) the same utility $c_d + 1$ as it receives from matching with d. Think of $c_d + 1$ as d's "target utility." Let

$$A_n(\varepsilon,(c,\eta)) = \{d \in D_n : c_d + 1 - \varepsilon < u_d(\mu_n^{DA}(d)) < c_d + 1 + \varepsilon\}$$

be the doctors for which this target is achieved (in DA), up to ε. We prove that, when n is large enough, with large probability, a fraction of at least $1-\theta/2$ doctors are in $A_n(\varepsilon,(c,\eta))$.

Consider the number of hospitals ranked above a doctor's adjusted target utility $c_d + 1 + \varepsilon$. Let

$$B(c_d, n) = \{\left|h \in H_n : c_h + \eta_{d,h} > c_d + 1 + \varepsilon\right| \leq k_n''\}$$

be the event that fewer than k_n'' hospitals provide d a utility greater than d's target utility. We denote by β_n the probability that a fraction of at least $\theta/2$ doctors have a "small" number, at most k_n'', of hospitals above their target utility.

We prove that for n large enough, $\beta_n < \pi/2$ and $P\left(\frac{1}{n}\left|A_n(\varepsilon, (c, \eta))\right| \geq 1 - \theta/2\right) > 1-\pi/2$. Thus, the event that $B(c_d, n)$ is false for a fraction $\geq 1-\theta/2$ of doctors *and* the event $\left(\frac{1}{n}|A_n(\varepsilon, (c, \varepsilon))| \geq 1 - \theta/2\right)$ holds, has probability $\geq (1-\pi/2)+(1-\pi/2)-1 = 1-\pi$. At the intersection of these events, it holds for a fraction $\geq (1-\theta/2)+(1-\theta/2)-1 = 1-\theta$ of $d \in D_n$ that $B(c_d, n)$ is false and $d \in A_n(\varepsilon)$. Hence, for a fraction $\geq 1-\theta$ of $d \in D_n$ there are more than k_n'' hospitals above their adjusted target utility, and they are within ε of their target utilities, hence below the adjusted target. The rank-order of any partner in μ^I is at most k_n. Since $M_n = k_n'' - k_n$, these statements prove the proposition.

To complete the proof we present the required calculations.

If $c_d \leq c^\star$, then

$$\begin{aligned}
P(B(c_d, n)) &= P(\sum_{h \in H_n} \mathbf{1}_{c_h + \eta_{d,h} > c_d + 1 + \varepsilon} \leq k_n'') \\
&\leq P\left(\frac{1}{n}\sum_{h \in H_n} \mathbf{1}_{c_h + \eta_{d,h} > c^\star + 1 + \varepsilon} \leq p(c^\star) - (p(c^\star) - \frac{k_n''}{n})\right) \\
&\leq \exp(-2(p(c^\star) - \frac{k_n''}{n})^2 n) \qquad (1)
\end{aligned}$$

by Hoeffding's inequality (observe that, eventually, $p(c^\star) - \frac{k_n''}{n} > 0$).

Let

$$\begin{aligned}
\beta_n &= P\left(|\{d \in D_n : B(c_d, n)\}| > n\theta/2\right) \\
&\leq P\left(\underbrace{\left|\{d \in D_n : B(d, n) \text{ and } c_d \leq c^\star\}\right|}_{Z_n} + \underbrace{\left|\{d \in D_n : c_d > c^\star\}\right|}_{Y_n} > n\theta/2\right) \\
&\leq P(\frac{1}{n}Z_n + 1 - G(c^\star) + \delta > \theta/2) + P(\frac{1}{n}Y_n > 1 - G(c^\star) + \delta).
\end{aligned}$$

The first inequality follows by counting all d with $c_d > c^\star$ as if $B(d,n)$ were true, so the random variable Y_n counts all $d \in D_n$ with $c_d > c^\star$ as if they were in $B(c_d, n)$.

The second inequality is a truncation exercise, partitioning the probability space into two events. The first is $\frac{1}{n}Y_n \leq 1 - G(c^\star) + \delta$ and the second $\frac{1}{n}Y_n > 1 - G(c^\star) + \delta$. Under the second event, $\frac{1}{n}Z_n + \frac{1}{n}Y_n > \theta/2$ as $1 - G(c^\star) + \delta > \theta/2$. Under the first event, the inequality is obtained by raising $\frac{1}{n}Y_n$ to $1 - G(c^\star) + \delta$.

Applying Hoeffding's inequality again,

$$P(\frac{1}{n}Y_n > 1 - G(c^\star) + \delta) \leq \exp(-2\delta^2 n). \tag{2}$$

Now,

$$\begin{aligned} P(Z_n > n(\theta/2 - [1 - G(c^\star) + \delta])) &\leq P(\cup_{d \in D_n} B(d,n))|c_d = c^\star) \\ &\leq \sum_{d \in D_n} P(B(d,n)|c_d = c^\star) \\ &\leq n \exp(-2(p(c^\star) - \frac{k''_n}{n})^2 n), \end{aligned} \tag{3}$$

where the first inequality follows as $n(\theta/2 - (1 - G(c^\star) + \delta)) \geq 1$, and the probability of $B(d,n)$ is maximized when $c_d = c^\star$.

Choose n such that

$$n(\theta/2 - [1 - G(c^\star) + \delta]) > 1, \tag{4}$$

$$\exp(-2\delta^2 n) < \pi/4, \tag{5}$$

$$n \exp(-2(p(c^\star) - \frac{k''_n}{n})^2 n) < \pi/4, \tag{6}$$

$$\text{and } P\left(\frac{1}{n}\left|A_n(\varepsilon, (c, \eta))\right| \geq 1 - \theta/2\right) > 1 - \pi/2. \tag{7}$$

Observe that (4) is possible as $\theta/2 - [1 - G(c^\star) + \delta] > 0$. Inequality (6) requires k''_n be $o(n)$, which holds by hypothesis, and our choice of $c^\star$ to ensure that $p(c^\star) - k''_n/n > 0$ is eventually bounded away from zero. Inequality (7) is possible by Theorem 1 of Lee (2016).

By (2),(3),(5), and (6), we obtain that

$$\beta_n \leq n \exp(-2(p(c^\star) - \frac{k''_n}{n})^2 n) + \exp(-2\delta^2 n) < \pi/2. \tag{8}$$

Statements (7) and (8) provide the required bounds.

For Online Publication–Appendix: Omitted Proofs

A.1 Proposition 1

Proof. Let $\succeq$ be the common preference that doctors have over hospitals. Note that DA is the same as serial dictatorship (SD) with the order dictated by hospital rank in $\succeq$.

Consider a doctor d assigned to $h = \mu^{DA}(d)$ in the rth round of SD. The rank-order of h in d's preference is therefore r. If $k \leq r$ then we are done, as the rank-order of $\mu^I(d)$ in d's truncated preference is at most k.

Suppose that $r < k$. Two observations follow. First, consider the interview stage and a hospital $h = \mu^{DA}(d)$ matched to d in stage $r' < k$ of DA. When choosing whom to interview, h can choose any doctor, as all of them would have received strictly fewer than k interview requests when they get a request from h. So the hospital choosing at stage r' of DA will interview the highest k doctors in her preference.

Second, $\mu^{DA}(h) = \mu^I(h)$ for the hospital h choosing at round r.[20] This is shown by induction: The statement is obviously true for the highest ranked hospital. Suppose that $\mu^{DA}(h) = \mu^I(h)$ for all hospitals choosing at any stage $r' < r$. If h is the r-ranked hospital then the set of doctors available to h in the DA stage of Int-DA is D, by our first observation, minus the choices of hospitals with rank-order $r' < r$. By the inductive hypothesis the doctors chosen by the hospitals with rank-order $r' < r$ is the same as DA. So the set of available doctors to hospital h is the same in Int-DA as in DA. Thus $\mu^{DA}(h) = \mu^I(h)$. ||

A.2 Proposition 3

Proof. Specifically, we show that there are constants N, K, K', K'' and K''' that do not depend on θ and π, such that for all

$$n \geq \max\{\bar{N}, \frac{\ln(\pi/4)}{K}, \frac{\ln(4/\pi)}{2\delta^2}, (\frac{\theta}{2} + K')^{-1}, (\frac{12}{\theta})^4, \left(\frac{\log(1-\frac{\pi}{2})}{\log K''} + 3\right)^4 K''', \}$$

the statement in Proposition 2 holds.

---

[20] Incidentally this may not happen for hospitals choosing at round $r' > k$. It is easy to come up with examples.

The market size in the proof of Proposition 2 is determined from inequalities (4)-(7). These are the starting point of the proof. Using the bounds in Lee (2016), these mean that we need to choose n such that

$$-2[p(c^\star) - \frac{k_n}{n}]^2 n \leq \ln(\frac{\pi}{4n}) \tag{9}$$

$$-2\delta^2 n < \ln(\frac{\pi}{4}), \tag{10}$$

$$\frac{1}{\frac{\theta}{2} - (1 - G(c^\star) + \delta)} < n \tag{11}$$

$$\frac{2}{n}\left(\frac{1}{n^{1/4}} - 3\right)\sqrt{n}\log(n) + \frac{6}{n^{1/4}} > \frac{\theta}{2} \tag{12}$$

$$(1 - g_n)^{2n^{1/4}-4} \geq 1 - \frac{\pi}{2}, \tag{13}$$

where g_n is $o(e^{-\sqrt{n}\log n})$

For (9), choose N_0 and K_0 such that if $n \geq N_0$ then $(p(c^\star) - k_n/n)^2 \leq K_0$. This is possible given the hypothesis that $\limsup k_n/n < 1$. Next, let $N_1 \geq N_0$ and K_1 be such that, for all $n \geq N_1$, $2K_0 n - \ln n \geq K_1 n$. Then we need that

$$K_2 n \geq \ln(\frac{4}{\pi}) \tag{14}$$

For (10) and (11), we have

$$n > \frac{\ln(4/\pi)}{2\delta^2} \tag{15}$$

$$n \geq \frac{1}{\frac{\theta}{2} - (1 - G(c^\star)) - \delta} \tag{16}$$

For (12) we need that

$$\frac{2\sqrt{n}\log n}{n^{5/4}} - \frac{6\sqrt{n}\log n}{n} + \frac{6}{n^{1/4}} < \frac{\theta}{2}$$
$$\iff \frac{2\log n}{n^{1/4}}\left(\frac{1}{\sqrt{n}} - \frac{3}{n^{1/4}}\right) + \frac{6}{n^{1/4}} < \frac{\theta}{2}$$

Let $N_2 \geq N_1$ be such that for all $n \geq N_2$, $\frac{1}{\sqrt{n}} - \frac{3}{n^{1/4}} \leq 0$. Then all we need is that

$\frac{6}{n^{1/4}} < \frac{\theta}{2}$, or that

$$n \geq (\frac{12}{\theta})^4. \tag{17}$$

For (13), fix $N_3 \geq N_2$ and K_4 such that for all $n \geq N_3$ $1 - g_n \geq K_4$. So we need to obtain $\log(1 - \frac{\pi}{2})) \leq (2n^{1/4} - 3)\log K_3$. That is,

$$n \geq \left(\frac{\log(1 - \frac{\pi}{2})}{\log K_3} + 3 \right)^4 \frac{1}{16} \tag{18}$$

Set $\bar{N} = N_3$, $K = K_2$ $K' = (1 - G(c^\star)) + \delta$, $K'' = K''' = 1/16$. Then the calculations above correspond to (14), (15), (16), (17), and (18). ||

B For Online Publication–Appendix: Additional Figures

B.1 Historical data form the NRMP

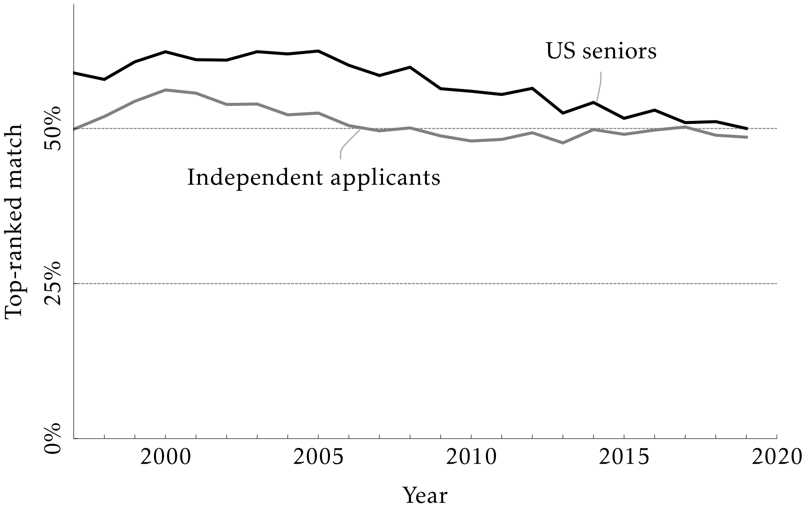


Figure B.1: NRMP residents matched to first-ranked program (conditional on matching)

B.2 Simulation summary

In each of our simulations we considered all possible combinations of $(\lambda_D, \lambda_H) \in \{1/20, 1/4, 1/2, 3/4, 19/20\}^2$ to get a representative span for the effect. In our balanced simulations we consider matching N doctors to N hospitals with $k = k' = 5$ for all combinations of $N \in \{50, 100, 200, 500, 1000, 1700\}$. In each balanced simulation we draw preferences for a total of 17,000 doctors and 17,000 hospitals (so 340 simulations at $N = 50$ and 10 simulations total at $N = 1700$). All simulation data and the code to generate and analyze it are available from the paper's ICPSR repository.

In addition to the balanced markets, we also analyzed the effects of (i) imbalanced markets; (ii) the effects of the interview capacities; (iii) and alternative procedure where only the doctor's common component was used to allocate interviews. Representative results from the simulations carried out are included in this appendix as follows:

Balanced The effects of N are illustrated for our balanced market in Figure B.2. In Figure B.4 we show how the difference market sizes are estimated with our logit model to allow for aggregation. In Figures B.7 and B.8 we outline the difference in outcomes for the minority of matched doctors for do not get the DA partner. More-detailed statistics for each separate market size are given in:

- $N = 50$: Table B.1
- $N = 100$: Table B.2
- $N = 200$: Table B.3
- $N = 500$: Table B.4
- $N = 1000$: Table B.5
- $N = 1700$: Table B.6

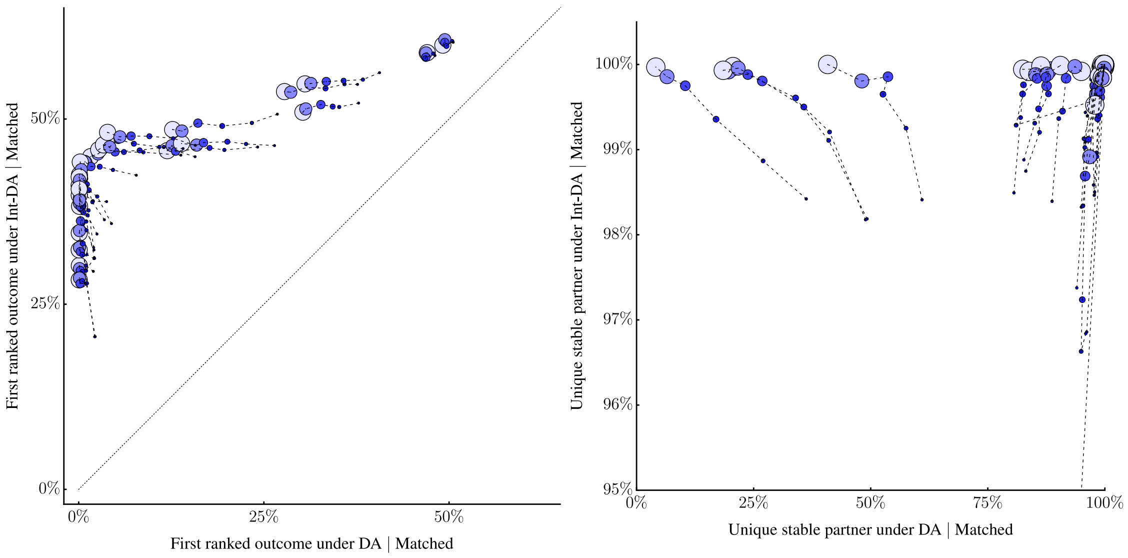
Imbalanced We examine the effects of market imbalance with $N_D = 600$ doctors being matched to $N_H = 500$ positions. Figure B.3 indicates the effects of imbalance relative to the balanced market at $N = 500$. In Table B.7 we present more-detailed statistics.

k-Effects For a balanced market with $N = 500$ we examine the effects of the interview capacities where we set $k = k'$. We consider $k \in 2, 5, 10, 20$, where we graph the effects in B.5. We present more-detailed statistics in:

- $k = 2$: Table B.8
- $k = 10$ =: Table B.10
- $k = 20$: Table B.11

Information We consider an alternative interview selection procedure where we calculate the stable allocation with $k = k' = 5$ when hospitals do not realized the independent idiosyncratic component $\eta_{h,d}$ until after the interviews. As such, interviews are allocated using a much simpler procedure where doctors propose in turn according to their ranking by hospitals over the common component. Each doctor then proposes to their k favorite hospitals that have not yet filled their quota k'. Following this interview stage, the full preferences are then used on the interview-truncated list within DA.

Figure B.6 indicates the effect of this procedure over the full-information Int-DA procedure for first-ranked and unmatched rates.

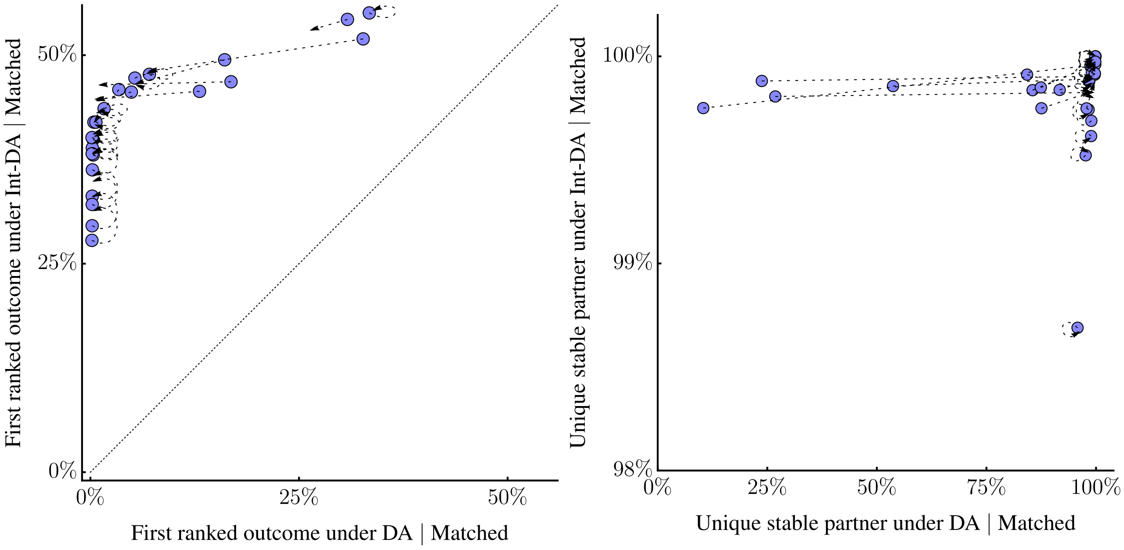


(a) First-ranked program

(b) Same partner | Match

Figure B.2: Effects from market size N on top-ranked proportion

Note: Lines show effects over N for balanced market simulations at $k = 5$ across the following market sizes $N \in \{50, 100, 200, 500, 1000, 1700\}$, where each chain indicates a parameter pair $(\lambda_D, \lambda_H) \in \{1/20, 1/4, 1/2, 3/4, 19/20\}^2$. Areas of each point proportional to market size.

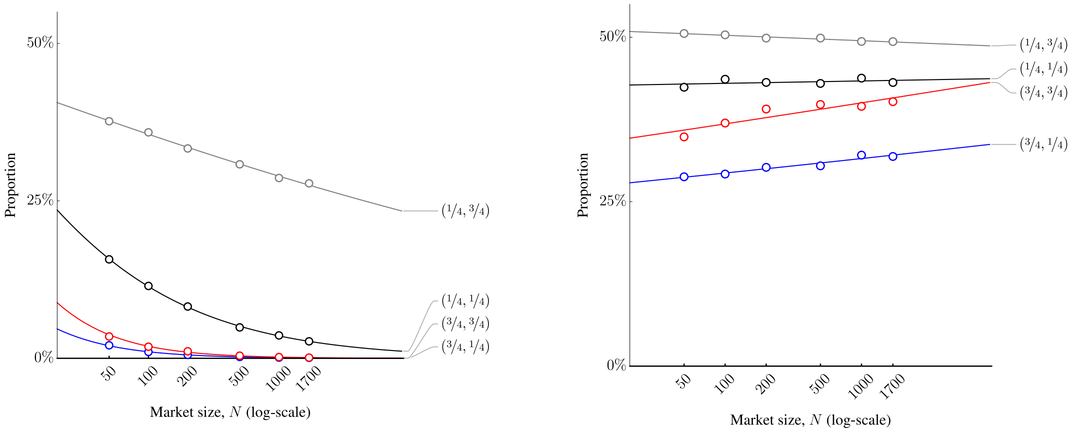


(a) First-ranked program

(b) Same partner | Match

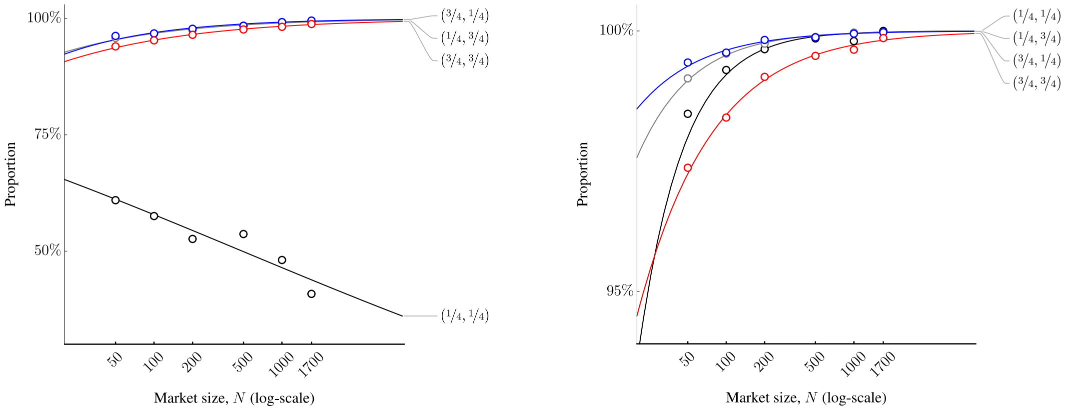
Figure B.3: Imbalanced market effects for DA vs Int-DA

Note: Arrows show effect in moving from a balanced market with $N_D = N_H = 500$ market (blue point) to a market with $N_D = 600$, $N_H = 500$ doctors (20% excess) for all values $(\lambda_D, \lambda_H) \in \{1/20, 1/4, 1/2, 3/4, 19/20\}^2$. Horizontal axis shows the effect under DA, vertical under Int-DA.



(a) First-ranked outcome under DA

(b) First-ranked outcome under Int-DA



(c) Unique stable partner under DA

(d) Unique stable partner under DA

Figure B.4: Fitted models in market size N for preference weights (λ_D, λ_H)

Note: Lines indicate the fitted models (a linear model for the log-odds ratio for the relevant variable against the $\log(N)$ on the right-hand side) while points indicate simulation data across 170,000 doctors at the relevant (λ_D, λ_H) preference pair. All simulations shown are for balanced markets with $k = k' = 5$

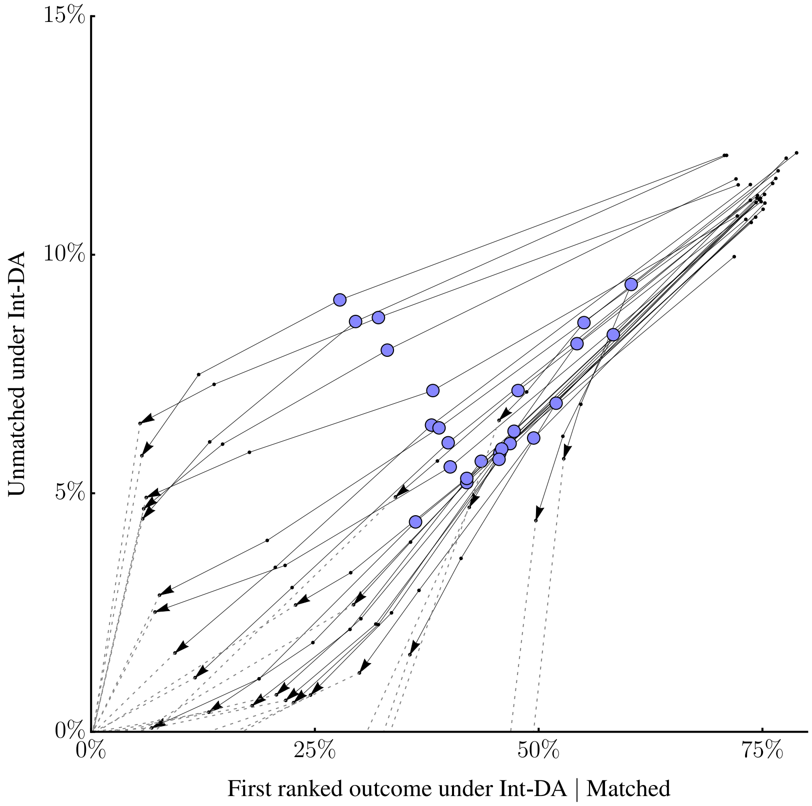
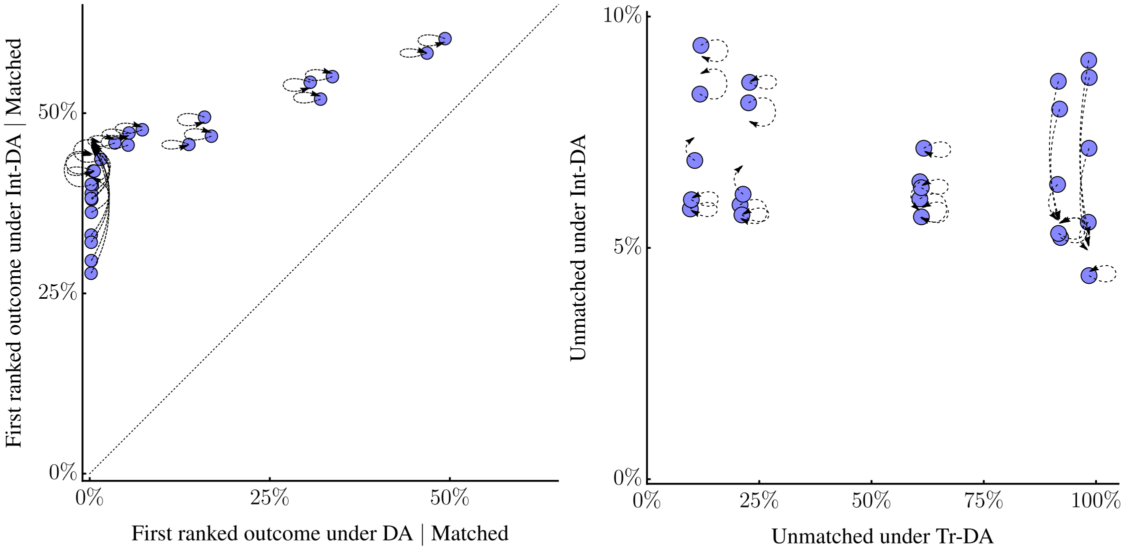


Figure B.5: Effects from the number of interview slots $k(=k')$ on top-rank and unmatched proportions

Note: Arrows show transitions for balanced market with $N = 500$ participants as we increase $k \in \{2, 5, 10, 20\}$. Each arrow shows simulations under preference parameters $(\lambda_D, \lambda_H) \in \{1/20, 1/4, 1/2, 3/4, 19/20\}^2$. Dotted lines show effect at $k \to \infty$ (the proportion under pure DA without interviews). Blue points show the locations under our core simulations at $k = 5$.



(a) First-ranked program

(b) Unmatched rates

Figure B.6: Alternative Interview Process: Full-Information vs. Doctor-Common–ranking

Note: Blue points indicate results from our standard interview-selection model. Curved arrows indicate the effect in the vertical direction from changing the interview selection procedure from Int-DA to Sim-Int-DA. In this alternative implementation, we assume that hospitals initial preferences are entirely driven by the common-value component (test scores, letters, etc.). At the interviews, the hospitals acquire information on the idiosyncratic component, subsequently ranking interviewed doctors using the full preference in DA. While this distinct algorithmic process allows for information to have a productive role in the interviews, it also substantially simplifies the computational burden of finding a stable interview matching. Hospitals' k interview slots are filled sequentially using doctors' preferences, acting in turn from the doctor with the highest common-rank, to the lowest.

Table B.1: Simulation Outcomes, $N = 50$

	$\lambda_H = 1/4$			$\lambda_H = 3/4$		
	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$
Panel A: Matching outcomes						
Unmatched [DA: 0.0%]						
Int-DA	5.6%	6.1%	7.6%	7.5%	5.9%	4.3%
Tr-DA	14.7%	38.9%	71.2%	16.7%	40.1%	71.7%
First-ranked program						
DA	15.7%	3.8%	2.1%	37.6%	14.1%	3.5%
Int-DA	42.4%	36.5%	28.8%	50.6%	43.3%	34.8%
Tr-DA	30.9%	10.7%	2.4%	40.6%	16.5%	3.6%
Top-three–ranked program						
DA	40.4%	12.4%	6.3%	64.4%	32.2%	10.3%
Int-DA	81.8%	78.3%	72.6%	82.8%	82.6%	81.1%
Tr-DA	67.9%	37.4%	10.9%	70.9%	41.2%	13.2%
Panel B: Core size, similarity to DA, and stability						
Same partner under proposer change | Matched						
DA	61.0%	83.1%	96.3%	95.8%	95.5%	94.0%
Int-DA	98.4%	98.7%	99.4%	99.1%	98.7%	97.4%
Identical partner to DA | Matched						
Int-DA	78.9%	85.0%	87.1%	86.9%	83.9%	86.7%
Proportion blocking programs in Int-DA						
Matched	0.4%	0.6%	0.7%	0.6%	1.2%	1.2%
Unmatched	16.1%	14.2%	15.1%	17.8%	19.7%	33.3%

Table B.2: Simulation Outcomes, $N = 100$

	$\lambda_H = 1/4$			$\lambda_H = 3/4$		
	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$
Panel A: Matching outcomes						
			Unmatched	[DA: 0.0%]		
Int-DA	5.7%	6.3%	8.0%	8.0%	6.3%	4.7%
Tr-DA	16.8%	47.0%	80.6%	18.1%	47.2%	80.3%
			First-ranked program			
DA	11.5%	1.9%	1.0%	35.9%	10.9%	1.8%
Int-DA	43.6%	36.5%	29.2%	50.4%	43.3%	37.0%
Tr-DA	28.2%	8.7%	1.4%	39.1%	13.5%	2.2%
			Top-three–ranked program			
DA	31.3%	6.2%	3.1%	61.8%	25.7%	5.7%
Int-DA	82.1%	78.2%	73.0%	82.4%	81.9%	80.9%
Tr-DA	65.3%	31.5%	6.7%	68.8%	35.6%	8.5%
Panel B: Core size, similarity to DA, and stability						
			Same partner under proposer change | Matched			
DA	57.5%	86.0%	96.8%	96.8%	95.5%	95.3%
Int-DA	99.3%	99.2%	99.6%	99.6%	99.1%	98.3%
			Identical partner to DA | Matched			
Int-DA	78.6%	83.6%	84.8%	84.9%	81.0%	84.4%
			Proportion blocking programs in Int-DA			
Matched	0.3%	0.7%	0.8%	0.4%	1.2%	1.4%
Unmatched	14.9%	12.4%	13.4%	17.9%	21.6%	34.7%

Table B.3: Simulation Outcomes, $N = 200$

	$\lambda_H = 1/4$			$\lambda_H = 3/4$		
	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$
Panel A: Matching outcomes						
			Unmatched [DA: 0.0%]			
Int-DA	5.8%	6.4%	7.8%	7.9%	6.2%	5.1%
Tr-DA	19.0%	53.5%	86.2%	20.6%	53.7%	86.4%
			First-ranked program			
DA	8.2%	0.8%	0.5%	33.3%	7.5%	1.1%
Int-DA	43.1%	37.3%	30.2%	49.9%	43.8%	39.1%
Tr-DA	27.2%	6.9%	0.9%	37.3%	9.9%	1.4%
			Top-three–ranked program			
DA	23.1%	3.0%	1.5%	57.7%	19.3%	3.2%
Int-DA	81.6%	78.5%	73.7%	82.2%	82.2%	80.7%
Tr-DA	63.2%	26.9%	5.0%	66.2%	29.7%	5.6%
Panel B: Core size, similarity to DA, and stability						
			Same partner under proposer change | *Matched*			
DA	52.6%	85.8%	97.8%	97.5%	97.4%	96.5%
Int-DA	99.7%	99.5%	99.8%	99.8%	99.6%	99.1%
			Identical partner to DA | *Matched*			
Int-DA	77.0%	82.9%	83.5%	84.3%	80.1%	82.8%
			Proportion blocking programs in Int-DA			
Matched	0.3%	0.7%	0.9%	0.3%	1.1%	1.6%
Unmatched	13.2%	10.9%	11.7%	18.6%	22.5%	36.1%

Table B.4: Simulation Outcomes, $N = 500$

	$\lambda_H = 1/4$			$\lambda_H = 3/4$		
	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$
Panel A: Matching outcomes						
			Unmatched [DA: 0.0%]			
Int-DA	5.7%	6.1%	8.0%	8.1%	6.3%	5.2%
Tr-DA	21.1%	60.8%	91.9%	22.6%	61.1%	92.1%
			First-ranked program			
DA	4.9%	0.3%	0.2%	30.8%	5.3%	0.4%
Int-DA	43.0%	37.5%	30.4%	49.9%	44.3%	39.8%
Tr-DA	25.6%	6.2%	0.4%	34.7%	7.8%	0.6%
			Top-three–ranked program			
DA	14.6%	1.2%	0.6%	55.3%	14.1%	1.2%
Int-DA	81.4%	79.0%	74.2%	82.1%	81.6%	80.9%
Tr-DA	60.1%	21.9%	2.5%	64.0%	24.3%	2.7%
Panel B: Core size, similarity to DA, and stability						
			Same partner under proposer change | Matched			
DA	53.7%	87.4%	98.4%	98.5%	98.3%	97.6%
Int-DA	99.9%	99.8%	99.9%	99.9%	99.7%	99.5%
			Identical partner to DA | Matched			
Int-DA	76.0%	83.9%	82.4%	82.6%	78.3%	82.1%
			Proportion blocking programs in Int-DA			
Matched	0.2%	0.6%	0.9%	0.2%	0.9%	1.7%
Unmatched	11.2%	9.2%	10.2%	18.6%	23.2%	34.7%

Table B.5: Simulation Outcomes, $N = 1000$

	$\lambda_H = 1/4$			$\lambda_H = 3/4$		
	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$
Panel A: Matching outcomes						
			Unmatched	[DA: 0.0%]		
Int-DA	6.0%	6.4%	8.2%	8.0%	6.4%	5.1%
Tr-DA	22.6%	65.7%	93.9%	24.5%	64.9%	94.5%
			First-ranked program			
DA	3.6%	0.2%	0.1%	28.6%	4.2%	0.2%
Int-DA	43.8%	38.2%	32.1%	49.4%	43.3%	39.5%
Tr-DA	25.2%	4.8%	0.3%	33.3%	6.9%	0.3%
			Top-three–ranked program			
DA	11.0%	0.6%	0.3%	51.9%	11.1%	0.8%
Int-DA	81.7%	79.5%	74.3%	81.8%	81.7%	81.3%
Tr-DA	58.2%	18.7%	2.0%	62.1%	21.5%	2.0%
Panel B: Core size, similarity to DA, and stability						
			Same partner under proposer change | Matched			
DA	48.1%	87.6%	99.2%	99.2%	98.8%	98.2%
Int-DA	99.8%	99.9%	99.9%	100.0%	100.0%	99.6%
			Identical partner to DA | Matched			
Int-DA	71.3%	82.2%	81.9%	81.9%	76.2%	82.2%
			Proportion blocking programs in Int-DA			
Matched	0.2%	0.6%	1.0%	0.1%	0.8%	1.6%
Unmatched	10.3%	8.6%	9.6%	18.9%	23.1%	34.0%

Table B.6: Simulation Outcomes, $N = 1700$

	$\lambda_H = 1/4$			$\lambda_H = 3/4$		
	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$
Panel A: Matching outcomes						
			Unmatched [DA: 0.0%]			
Int-DA	6.0%	6.4%	7.9%	8.1%	6.5%	5.3%
Tr-DA	24.4%	68.6%	95.9%	25.1%	68.6%	95.9%
			First-ranked program			
DA	2.7%	0.2%	0.1%	27.8%	3.3%	0.1%
Int-DA	43.1%	38.6%	31.9%	49.3%	43.5%	40.2%
Tr-DA	22.9%	4.5%	0.2%	32.2%	5.7%	0.2%
			Top-three–ranked program			
DA	8.0%	0.3%	0.2%	50.7%	9.0%	0.5%
Int-DA	81.8%	79.3%	74.8%	81.9%	81.8%	81.2%
Tr-DA	56.8%	16.7%	1.3%	60.8%	19.1%	1.3%
Panel B: Core size, similarity to DA, and stability						
			Same partner under proposer change | Matched			
DA	40.8%	90.5%	99.6%	99.2%	99.3%	98.8%
Int-DA	100.0%	100.0%	100.0%	100.0%	100.0%	99.9%
			Identical partner to DA | Matched			
Int-DA	77.3%	83.5%	81.2%	81.3%	76.8%	82.7%
			Proportion blocking programs in Int-DA			
Matched	0.1%	0.6%	0.9%	0.1%	0.7%	1.6%
Unmatched	9.8%	7.8%	9.0%	18.9%	23.9%	33.8%

Table B.7: Simulations for Unbalanced Market, $N_D = 600$, $N_H = 500$

	$\lambda_H = 1/4$			$\lambda_H = 3/4$		
	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$
Panel A: Matching outcomes						
	Unmatched [DA: 16.7%]					
Int-DA	19.1%	19.2%	19.5%	22.5%	21.2%	20.2%
Tr-DA	29.7%	65.8%	92.8%	30.7%	65.6%	93.2%
	First-ranked program					
DA	0.5%	0.3%	0.2%	21.9%	4.5%	0.4%
Int-DA	36.0%	32.5%	28.6%	41.0%	36.6%	33.3%
Tr-DA	20.6%	4.9%	0.3%	29.3%	7.1%	0.5%
	Top-three–ranked program					
DA	1.7%	0.9%	0.5%	38.5%	11.1%	1.1%
Int-DA	70.3%	68.1%	65.5%	68.9%	68.5%	68.2%
Tr-DA	51.0%	18.6%	2.2%	55.7%	21.3%	2.5%
Panel B: Core size, similarity to DA, and stability						
	Same partner under proposer change | Matched					
DA	99.7%	99.8%	99.9%	99.0%	98.4%	98.4%
Int-DA	99.9%	100.0%	100.0%	99.9%	99.7%	99.5%
	Identical partner to DA | Matched					
Int-DA	86.8%	86.4%	86.7%	84.7%	82.4%	82.1%
	Proportion blocking programs in Int-DA					
Matched	0.6%	0.5%	0.5%	0.4%	0.9%	1.3%
Unmatched	2.5%	2.7%	3.0%	7.6%	7.1%	8.2%

TABLE B.8: Simulation Outcomes, $k = k' = 2$

	$\lambda_H = 1/4$			$\lambda_H = 3/4$		
	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$
Panel A: Matching outcomes						
	Unmatched [DA: 0.0%]					
Int-DA	11.3%	11.2%	11.6%	11.6%	11.2%	10.8%
Tr-DA	36.6%	74.5%	96.1%	37.4%	73.2%	95.0%
	First-ranked program					
DA	5.0%	0.4%	0.2%	30.6%	5.5%	0.5%
Int-DA	66.8%	66.1%	63.7%	67.6%	66.4%	66.2%
Tr-DA	38.3%	11.2%	1.0%	41.2%	13.4%	1.4%
Panel B: Core size, similarity to DA, and stability						
	Same partner under proposer change | Matched					
DA	50.1%	84.8%	98.5%	98.5%	98.1%	97.2%
Int-DA	100.0%	100.0%	99.9%	100.0%	100.0%	100.0%
	Identical partner to DA | Matched					
Int-DA	73.8%	78.3%	80.7%	80.5%	77.1%	77.3%
	Proportion blocking programs in Int-DA					
Matched	0.4%	1.8%	3.0%	0.2%	1.3%	3.5%
Unmatched	18.1%	14.0%	13.6%	35.2%	39.2%	40.2%

Table B.9: Simulation Outcomes, $k = k' = 5$

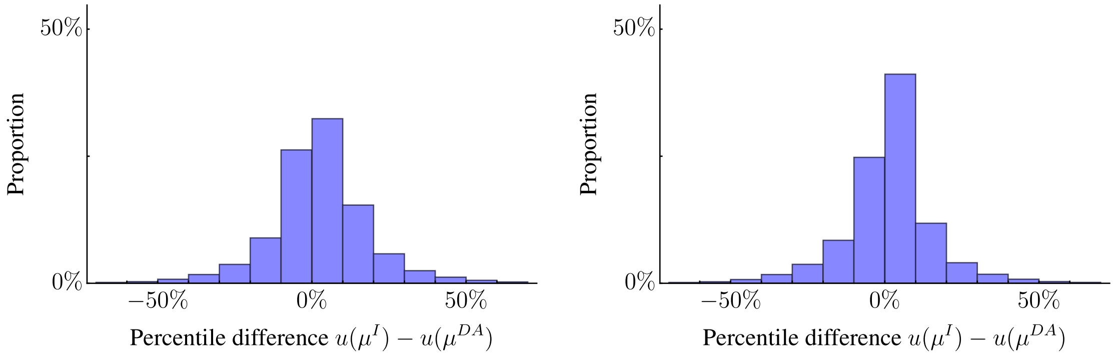
	$\lambda_H = 1/4$			$\lambda_H = 3/4$		
	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$
Panel A: Matching outcomes						
	Unmatched [DA: 0.0%]					
Int-DA	5.7%	6.1%	8.0%	8.1%	6.3%	5.2%
Tr-DA	21.1%	60.8%	91.9%	22.6%	61.1%	92.1%
	First-ranked program					
DA	4.9%	0.3%	0.2%	30.8%	5.3%	0.4%
Int-DA	43.0%	37.5%	30.4%	49.9%	44.3%	39.8%
Tr-DA	25.6%	6.2%	0.4%	34.7%	7.8%	0.6%
	Top-three–ranked program					
DA	14.6%	1.2%	0.6%	55.3%	14.1%	1.2%
Int-DA	81.4%	79.0%	74.2%	82.1%	81.6%	80.9%
Tr-DA	60.1%	21.9%	2.5%	64.0%	24.3%	2.7%
Panel B: Core size, similarity to DA, and stability						
	Same partner under proposer change | *Matched*					
DA	53.7%	87.4%	98.4%	98.5%	98.3%	97.6%
Int-DA	99.9%	99.8%	99.9%	99.9%	99.7%	99.5%
	Identical partner to DA | *Matched*					
Int-DA	76.0%	83.9%	82.4%	82.6%	78.3%	82.1%
	Proportion blocking programs in Int-DA					
Matched	0.2%	0.6%	0.9%	0.2%	0.9%	1.7%
Unmatched	11.2%	9.2%	10.2%	18.6%	23.2%	34.7%

Table B.10: Simulation Outcomes, $k = k' = 10$

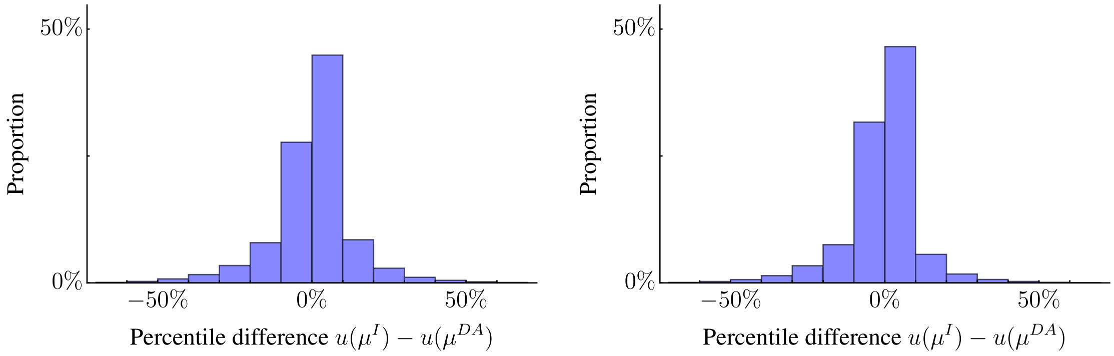
	$\lambda_H = 1/4$			$\lambda_H = 3/4$		
	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$
Panel A: Matching outcomes						
	Unmatched [DA: 0.0%]					
Int-DA	2.3%	3.0%	6.0%	6.0%	4.0%	1.9%
Tr-DA	11.5%	48.8%	86.9%	13.9%	48.9%	87.5%
	First-ranked program					
DA	5.1%	0.3%	0.2%	30.3%	5.1%	0.4%
Int-DA	31.1%	21.8%	13.8%	43.4%	34.3%	24.3%
Tr-DA	18.1%	2.9%	0.2%	32.0%	6.1%	0.4%
	Top-three–ranked program					
DA	15.6%	1.2%	0.6%	54.4%	14.0%	1.3%
Int-DA	67.5%	57.1%	43.3%	74.3%	70.8%	61.9%
Tr-DA	46.7%	12.4%	1.0%	58.9%	18.2%	1.6%
Panel B: Core size, similarity to DA, and stability						
	Same partner under proposer change | *Matched*					
DA	46.1%	88.4%	98.5%	98.9%	98.1%	97.7%
Int-DA	99.3%	99.5%	99.9%	99.9%	99.5%	99.0%
	Identical partner to DA | *Matched*					
Int-DA	78.2%	85.0%	82.1%	81.1%	76.4%	88.6%
	Proportion blocking programs in Int-DA					
Matched	0.1%	0.2%	0.2%	0.2%	0.6%	0.5%
Unmatched	7.5%	6.6%	8.8%	9.5%	8.8%	28.9%

Table B.11: Simulation Outcomes, $k = k' = 20$

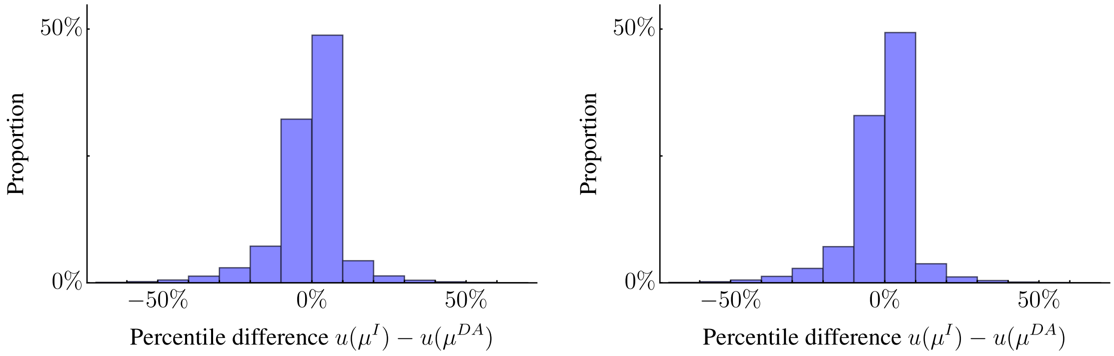
	$\lambda_H = 1/4$			$\lambda_H = 3/4$		
	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$	$\lambda_D = 1/4$	$\lambda_D = 1/2$	$\lambda_D = 3/4$
Panel A: Matching outcomes						
	Unmatched [DA: 0.0%]					
Int-DA	0.7%	1.1%	4.7%	4.7%	2.7%	0.4%
Tr-DA	5.6%	34.8%	79.5%	7.7%	36.5%	80.2%
	First-ranked program					
DA	4.7%	0.4%	0.2%	30.3%	5.3%	0.5%
Int-DA	21.6%	11.5%	5.6%	40.3%	28.5%	13.1%
Tr-DA	12.7%	1.7%	0.2%	30.8%	5.7%	0.5%
	Top-three–ranked program					
DA	14.1%	1.3%	0.6%	54.7%	13.7%	1.5%
Int-DA	51.6%	35.0%	19.1%	70.5%	63.5%	39.7%
Tr-DA	35.1%	7.0%	0.7%	56.2%	15.2%	1.5%
Panel B: Core size, similarity to DA, and stability						
	Same partner under proposer change | Matched					
DA	48.5%	86.5%	98.6%	98.5%	98.3%	97.2%
Int-DA	97.8%	98.7%	99.6%	99.5%	99.3%	97.9%
	Identical partner to DA | Matched					
Int-DA	75.2%	87.6%	80.8%	80.0%	78.8%	95.2%
	Proportion blocking programs in Int-DA					
Matched	0.0%	0.0%	0.0%	0.1%	0.4%	0.1%
Unmatched	4.1%	5.1%	7.7%	5.7%	3.6%	19.5%



(a) $N = 50$ (b) $N = 100$



(c) $N = 200$ (d) $N = 500$



(e) $N = 1,000$ (f) $N = 1,700$

FIGURE B.7: N-effects on Difference in match outcomes between Int-DA and DA | Matched

Note: Sample pools across all preference parameters (λ_D, λ_H) pairs. Data conditional on both a match in Int-DA, and a distinct outcome from the DA partner.

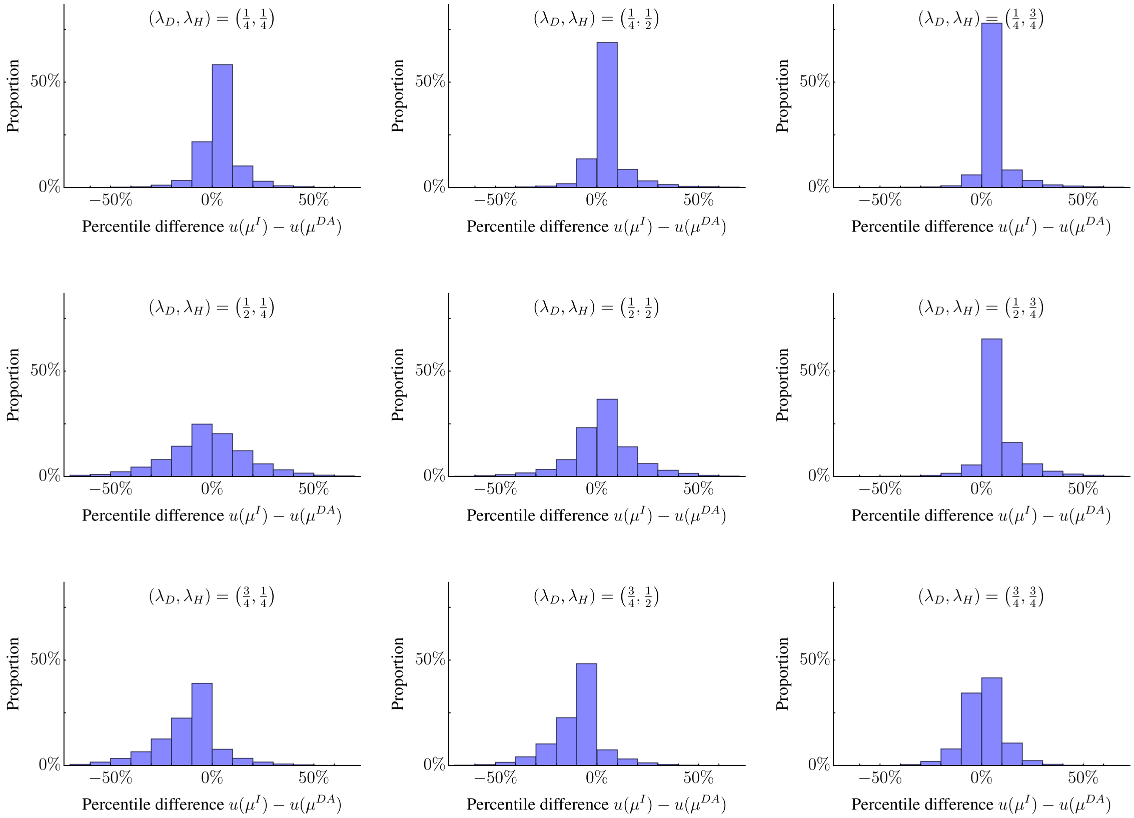


Figure B.8: λ-Effects on Difference in match outcomes between Int-DA and DA | Matched

Note: Sample pools all across simulation sizes N. Data conditional on both a match in Int-DA, and a distinct outcome from the DA partner.