SHORT LISTS IN CENTRALIZED CLEARINGHOUSES

NICK ARNOSTI, STANFORD UNIVERSITY

ABSTRACT. Stable matching mechanisms are used to clear many two-sided markets. In most settings, participants' lists tend to be short (even if there are many potentially acceptable matches). This paper studies the consequences of this fact, and focuses on two broad questions. First, when lists are short, what is the quantity and quality of matches formed through the clearinghouse? Second, what are the effects of introducing an aftermarket which allows agents left unmatched by the clearinghouse to find one another?

The answers to these questions depend crucially on the extent and form of correlations in agent preferences. I consider three canonical preference structures: fully independent (or idiosyncratic) preferences, vertical preferences (agents agree on the attractiveness of those on the opposite side), and aligned preferences (potential partners agree on the attractiveness of their match).

I find that when agent preferences are idiosyncratic, more matches form than when agents are vertically differentiated. Perhaps more surprisingly, I show that the case of aligned preferences causes the *fewest* matches to form. When considering quality of matches, the story reverses itself: aligned preferences produce the most high quality matches, followed by correlated preferences, with independent preferences producing the fewest. These facts have implications for the design of priority structures and tie-breaking procedures in school choice settings, as they point to a fundamental tradeoff between matching many students, and maximizing the number of students who get one of their top choices.

Regarding the role of the aftermarket, I find that when preferences are aligned, the aftermarket unambiguously improves the welfare of both sides. In other cases, however, the introduction of an aftermarket has multiple competing effects, and may either raise or lower aggregate welfare. This suggests that when designing an aftermarket, the extent and form of correlations in agent preferences are an important factor to consider.

1. INTRODUCTION

In several high-profile matching settings, centralized clearing houses have emerged as a way to clear the market. Typically, these clearing houses solicit preferences from all agents, and recommend a match that is *stable* with respect to the submitted preferences. The advantages of centralization (and stability in particular) are numerous and well-studied: the use of clearing houses allows agents to evaluate multiple potential partners before having to commit to a match, and when participants truthfully submit their preferences, stable mechanisms select core (and thus Pareto efficient) outcomes.

Most of the literature on stable matching mechanisms assumes that agents list all potential partners whom they find acceptable. This assumption rarely holds in practice, as agents opt instead for short lists, even in very large markets. The reasons for this are numerous, and vary across markets. In some cases, agents do not know their own preferences, and learning them is costly. In other settings, agents may know their preferences, but have to pay a cost for each one they list (for an example of this, consider college admissions, where each additional application incurs a fee).[1] Occasionally, there is even an explicit cap on the number of partners that each agent

Date: March 31, 2015.

[1]Although the American application and admissions process is decentralized, in many other countries, there is a centralized assignment procedure. These assignment mechanisms may charge for applying to many schools (as is the case in Hungary; see Biró (2007)), or cap the number of schools to which each student may apply (as happens in South Korea; see Avery et al. (2014)). Still other clearinghouses may not run stable mechanisms at all (see, for example, the work of Balinski and Sönmez (1999) and Braun et al. (2010)).

may list (for example, Abdulkadiroglu et al. (2009) note that students in NYC may list at most twelve schools to which they would like to match).

Regardless of the reasons for short lists, they are a feature of most centralized clearinghouses, and have important (and understudied) consequences.In this paper, I consider two questions which can only be addressed by a model which explicitly captures the fact that mutually acceptable agents may fail to list one another. First, in markets with short lists, a notable number of agents on both sides will remain (inefficiently) unmatched:[2] I ask how this number depends on the extent (and form) of correlation in agent preferences. As I discuss in Section 5, the answers to this question have direct consequences for the design of priorities and tie-breaking procedures for school assignment algorithms. Second, many centralized clearinghouses feature an "aftermarket" where unmatched agents may find and match to acceptable partners. The final portion of this paper studies the ways in which the presence of an aftermarket shapes agent incentives and utilities, and finds that the answer again depends crucially on the correlation structure of agent preferences.

I model the matching process as a two stage game. In the first stage, one side (the "doctors")[3] schedules a limited number of interviews with agents on the other side ("hospitals"). In the second stage, doctors and hospitals submit ranked lists, with the constraint that hospitals may only list doctors whom they have interviewed (and, WLOG, vice versa). Given the submitted lists, agents are matched according to the doctor-proposing deferred acceptance algorithm.

Agents make decisions at each stage of this process. I simplify the second stage by assuming that agents truthfully rank all potential partners whom they interviewed and found acceptable. This is a weak assumption, as doing so is a dominant strategy for doctors, and is essentially dominant for hospitals in large markets (see Immorlica and Mahdian (2005) and Kojima et al. (2013)). I simplify the first stage by assuming that hospitals have no interviewing constraints (and thus accept any requests that they receive), and that doctors possess no information that *ex-ante* differentiates hospitals (and thus schedule interviews with a set of hospitals chosen uniformly at random). These assumptions, made for technical convenience, are arguably quite strong, and imply that my model cannot provide insight into how agents should optimally schedule interviews in equilibrium.In the conclusion, I discuss each of these assumptions, and the ways in which the conclusions of the model would change if they were relaxed.

The primary focus of this paper is the study of how the preference structure influences the quantity and quality of matches formed by the clearinghouse. Factors underlying agent preferences in the real world are numerous and varied. Doctors may have *idiosyncratic* geographical preferences driven by the locations of their friends or family members. They may differ in their level of expertise, causing *correlation* in preferences, in that some doctors to be highly coveted by all hospitals. Additionally, doctor and hospital preferences may be *aligned*, as a doctor who specializes in a particular field may wish to go to a hospital where his expertise is in demand.[4] The bulk of this paper assumes that doctors have fully idiosyncratic preferences, and studies the way in which the *number* and *quality* of matches depends on the preferences of hospitals. In particular, I consider three cases:

---

[2]For example, in 2014, the NRMP match left roughly 1,000 residency positions unfilled - almost all of these programs matched later through a supplemental procedure (for more information, see `http://www.nrmp.org/wp-content/uploads/2014/04/Main-Match-Results-and-Data-2014.pdf`). In New York City, data from Abdulkadiroglu et al. (2009) suggests that in 2006-2007, approximately 5,500 students (7% of the total) were initially unassigned and forced to reapply in a second matching round.

[3]Although I use terminology from residency matching, the results from this paper apply to any market where a centralized stable matching mechanism is used, and lists are short relative to the number of participants. In particular, they provide insight into the setting of school choice, and so I will occasionally refer to the two sides of the market as students and schools.

[4]In a school choice setting, student preferences may stem from factors such as the school's location, its overall quality (as measured by test scores or graduation rates), or its strength in a particular area (for example, it may have a particularly good orchestral program).

- *Independent* preferences: hospital preferences are entirely idiosyncratic;
- *Correlated* preferences: hospitals share a common ranking of doctors;
- *Aligned* preferences: for each pairing, the doctor and the hospital have a common assessment of its quality. Match qualities are themselves idiosyncratic.

In Section 4, I ask which of these preference structures produces the most matchings (holding fixed the other parameters of the market). I find that correlated hospital preferences result in fewer matches than independent preferences, as low-quality doctors struggle to find employment. I also prove that aligned preferences produce the *fewest* matches of the three - a fact which may seem surprising, as in this case there are no inherently low quality doctors.It is worth noting that these differences may be of "first order" importance. For example, in a balanced market where doctors may schedule $m = 16$ interviews, the model predicts that 5.20% of doctors remain unmatched when doctor and hospital preferences are aligned. By contrast, when preferences are independent, only 1.45% of doctors go unmatched when $m = 16$, and (only) 4.87% go unmatched even when m is reduced to 8.

Of course, the number of matches that form is an imperfect measure of match quality, so Section 5 studies the question of the *quality* of matches that form. Here, the story reverses itself: correlated hospital preferences produce more high-quality matches for doctors than do independent preferences, and aligned preferences produce the most high-quality matches (more precise statements are available in Section 5). I explain why, at a fundamental level, hospital preferences which tend to give many doctors high-quality matches also tend to leave more doctors unassigned (and vice versa).

These results provide practical guidance for the design of school choice algorithms, where school preferences are engineered, not taken as given. In such settings, schools typically assign each student to one of several broad categories, and must use a lottery to break ties among students. One question which has been studied in this literature is how to correlate tie-breaking across schools. In particular, Abdulkadiroglu et al. (2009) consider two possible lottery procedures. A *single tie-breaking* (STB) procedure assigns each student a single lottery number, and breaks all ties according to this number. A *multiple tie-breaking* (MTB) procedure assigns students one lottery number *per school*, and breaks ties for positions at a given school according to the lottery numbers for that school. Abdulkadiroglu et al. (2009) observe that in simulations, the former procedure assigns more students their top choices, but also leaves more students unassigned. The authors offer no compelling explanation for this observation, and indeed Pathak (2011) comments that "there is currently no known stronger ex ante argument for single versus multiple tiebreaking based on the distribution of matchings." My results provide such an argument, and predict the trend observed in the data. I close Section 5 with a brief discussion of alternative priority structures and tie-breaking procedures, and their anticipated consequences.

The final component of this paper, presented in Section 6, studies the effects of introducing an aftermarket to match agents left unassigned by the clearinghouse. I assume the simplest possible structure for such a market: each unmatched agent on the short side is matched to a random agent on the long side. The presence of the aftermarket provides each agent with an outside option whose value is determined by the aggregate behavior of other agents.

I show that the preference structure again crucially affects market outcomes. When preferences are aligned, there is a unique equilibrium, and both sides benefit from the presence of the aftermarket. Neither of these results necessarily holds for other preference structures. I fully characterize the equilibria when match values are independent draws from a binary distribution. In general, there may be multiple equilibria, whereas when the long side has correlated preferences, there is a single equilibrium. Unlike the case of aligned preferences, when preferences are independent or correlated, equilibrium welfare may be lower than welfare when the aftermarket is absent.

2. Related Work

The first work to document the use of a stable centralized clearing house was that of Roth (1984), who noted that the algorithm used to match matches medical students to residency positions was, in fact, equivalent to the deferred acceptance algorithm described by Gale and Shapley (1962).

Pittel (1989) was among the first to study the properties of stable mechanisms in large markets with random preferences; one conclusion of this work is that in large balanced markets with uniform preferences, nearly all agents have multiple stable partners. This contrasts with the empirical findings of Roth and Peranson (1999), who observe that in practice, very few agents have multiple stable partners. One explanation for this was provided by Immorlica and Mahdian (2005), which is to my knowledge the first paper studying a large market model with short preferences lists. They demonstrate that when preference lists are bounded in size, the core of the market becomes essentially unique as the market grows. More recent large market models have shown that this result continues to hold in the presence of a small number of couples Kojima et al. (2013), and in the case when agents have long lists, so long as the market is not perfectly balanced Ashlagi et al. (2013).

Each of the preference structures considered in this paper have appeared repeatedly in previous work (in far too many papers to try to list them all here). Recent papers studying these three preference structures in the context of large random markets include the work of Boudreau and Knoblauch (2010) (who use simulation results to conclude that even a small degree of alignment can significantly reduce the size of the core); and that of Lee and Yariv (2014) (who ask whether stable matchings are asymptotically utilitarian efficient). Unlike our work, these papers assume that each agent has (and submits) complete preferences over the opposite side.

Mathematically, a number of results in this paper are closely related to the computer science literature studying maximal and maximum matching in random graphs. In particular, the work of Wormald (1995) spawned many papers which use differential equations to analyze large graphs (as I do in Theorems 1 and 3 for the case of correlated preferences). One closely-related paper is that of Mastin and Jaillet (2013), which studies the size of the maximal matching produced by greedily assigning each arriving agent to an unmatched partner (although they consider the case of an Erdos-Renyi random graph where each edge present with probability m/n, whereas I assume that each doctor schedules exactly m interviews). The graph structure from this paper has been studied in the context of cuckoo hashing by Dietzfelbinger et al. (2010), Frieze and Melsted (2012) and Fountoulakis and Panagiotou (2012), but they study the size of *maximum* matchings, rather than stable ones (in the language of this paper, they ask how many excess hospital positions are needed in order for it to be possible to match virtually all doctors if one ignores stability constraints).

Although interviews are widely recognized to play an important role in determining match outcomes, little formal work exists on this topic. Two papers that study interviews in matching markets are those of Lee and Schwarz (2012) and Rastegari et al. (2013). These papers, however, focus on questions related to the *optimal coordination* of interviews. This paper instead takes uncoordinated interviewing as a given, and aims to study match outcomes in the resulting market.

3. Model

I consider a sequence of two-sided matching markets indexed by the parameter n. In the n^{th} market, there are n hospitals and $\lfloor rn \rfloor$ doctors. For ease of exposition, I assume that each hospital has a single position; the case where hospitals have multiple positions is treated in Appendix 9.1, and most of the conclusions reached in this paper continue to apply. Note that r determines the market imbalance: $r > 1$ implies that there are more doctors than there are hospital positions. For a given market, I use $\mathcal{D}$ to denote the set of doctors, and $\mathcal{H}$ to denote the set of hospitals.

Matches are formed through a two-period process:

- First, each doctor d (simultaneously) selects a subset $H_d \subseteq \mathcal{H}$ of hospitals with whom they wish to interview (hospitals accept all requests).[5]
- In the second stage, doctors and hospitals submit preference lists, and may only list agents whom they interviewed. The doctor-optimal deferred acceptance algorithm is run on the submitted preference lists.

When doctor d interviews with hospital h, both agents learn the cardinal utility that they would receive from matching to each other. I denote doctor d's utility for matching to h by u_{dh} and h's utility for matching to d by v_{dh}. I assume that u_{dh} and v_{dh} are positive for all d, h, that unmatched agents get a utility of zero, and that agents truthfully report their (ordinal) preferences to the matching mechanism (the Appendix considers the case where agents may view each other as unacceptable).

I capture the inherent frictions of interviewing by assuming that each doctor is capacity constrained, and may only select at most $m < n$ hospitals with whom they can interview; that is, for all d, $|H_d| \leq m$. Furthermore, I make the simplifying assumptions that interviews are costless, that hospitals appear ex-ante identical to doctors, and that doctors cannot coordinate their interviews with one another. Taken together, these assumptions apply that doctors select m hospitals with which they interview uniformly at random, and independently from one another.

In this paper, I consider several different preference models; that is, several joint distributions for the values u_{dh} and v_{dh}. In all cases, I assume that the values u_{dh} are iid draws from a distribution F;[6] that is, that doctors have idiosyncratic preferences. This paper considers three possibilities for the joint distribution of u, v:

- *Independent* preferences: The values v_{dh} are iid $\sim F$ (independently from u).
- *Correlated* preferences: Each doctor d has a quality q_d, drawn iid from the distribution F. For all d, h, we have $v_{dh} = q_d$.
- *Aligned* preferences: For all d, h, we have $v_{dh} = u_{dh}$.

For fixed interviews and fixed values of u_{dh} and v_{dh}, the clearing house selects a matching μ of size $|\mu|$. Let $\mu_n^I = \mu_n^I(r, m)$ denote the (random) matching that results in the market with parameters r, m, n and independent preferences. Similarly, define μ_n^C and μ_n^A to be (random) matchings produced when (hospital) preferences are correlated and aligned, respectively. Note that by the rural hospital theorem, the size of the stable matching μ does not depend on which stable match is selected.

In the following sections, I fix the market imbalance (r) and the number of interviews scheduled by each doctor (m), let the size of the market (n) grow, and study differences in match outcomes between the cases of independent, correlated, and aligned hospital preferences. I begin by analyzing the number of matches that form in each case, before moving on to discuss the quality of these matches in Section 5.

[5]As mentioned in the introduction, this arguably a strong assumption. I discuss it in detail in the conclusion.

[6]The assumption that F is the same for all agents is merely a notational convenience. All results from Sections 4 and 5 carry through if we allow each agent a to have a personal distribution function F_a. In this case, the definition of correlated and aligned preferences should read that $F_h(v_{dh}) = F_d(q_d)$ and $F_d(u_{dh}) = F_h(v_{dh})$, respectively.

4. Results: Number of Matches

I begin with a result which allows for explicit computation of the fraction of hospitals that match in large markets (as a function of r, m, and the preference structure). This is used to prove Theorem 2, which states that for any r, m, independent preferences produce the most matches, and aligned preferences produce the fewest.

Theorem 1.

(1) Independent preferences:

$$\lim_{n\to\infty} \frac{E[|\mu_n^I|]}{n} = x^*,$$

where x^* is the unique solution to

$$r = \frac{x^*}{1 - \left(1 + \frac{x^*}{\log(1-x^*)}\right)^m}. \tag{1}$$

(2) Perfectly correlated preferences:

$$\lim_{n\to\infty} \frac{E[|\mu_n^C|]}{n} = X(r),$$

where $X(t)$ is the solution to the differential equation[7]

$$X'(t) = 1 - X(t)^m, \qquad X(0) = 0. \tag{2}$$

(3) Aligned preferences:

$$\lim_{n\to\infty} \frac{E[|\mu_n^A|]}{n} = 1 - \tilde{G}(0),$$

where $\tilde{G}$ is the solution to the differential equation

$$\tilde{G}'(t) = rm\tilde{G}(t)\left(1 - \int_t^1 \tilde{G}(u)du\right)^{m-1}, \qquad \tilde{G}(1) = 1. \tag{3}$$

I provide proofs for the cases of independent and correlated preferences in Appendix 8.1.[8] Readers seeking intuition for the results corresponding to independent and correlated preferences can find it in Section 4.1 (similarly, intuition for Theorem 2 is provided in Section 4.2); those uninterested in this intuition may skip that portion of the paper.

[7]For $m \in \{1, 2\}$, this differential equation has a closed-form solution: $X(r) = 1 - e^{-r}$ and $X(r) = \frac{e^{2r}-1}{e^{2r}+1}$, respectively. For $m = 3$, it is possible to compute that the solution $x = X(r)$ satisfies

$$6r = \log\left(1 + \frac{3x}{(1-x)^2}\right) + 2\sqrt{3}\tan^{-1}\left(\frac{2x+1}{\sqrt{3}}\right) - \frac{\pi}{\sqrt{3}},$$

and for $m = 4$, the value $x = X(r)$ satisfies

$$4r = \log\left(\frac{1+x}{1-x}\right) + 2\tan^{-1}(x).$$

Similar expressions may be derived for $m \geq 5$, but the value of doing so is questionable, as the differential equation (2) is arguably far more illuminating.

[8]The current version of the Appendix provides derivations for the aligned preference results from Theorems 1 and 3, but does not provide a formal proof that this derivation is correct. Simulations confirm the correctness of the given expressions for aligned preferences, and I expect to post proofs for their correctness in an online appendix by the end of the month.

Proofs for the remaining theorems in the paper appear in complete form in the appendix; furthermore, these proofs *do not rely on the correctness of the expressions for aligned preferences*. For example, the conclusion that $M^C \geq M^A$ is established using a coupling argument which demonstrates that $E[|\mu_n^C|] \geq E[|\mu_n^A|]$ for any finite n (and thus this conclusion must continue to hold in the limit).

Motivated by Theorem 1, define $M^I = M^I(r,m) = \lim_{n\to\infty} \frac{1}{n} E[|\mu_n^I|]$, and similarly for M^C and M^A. While Theorem 1 makes it possible to compute, for any r and m, the values M^I, M^C, M^A, it is worthwhile to gain a structural understanding of the given expressions. For example, it is straightforward to show that these quantities are increasing in r and m, and approach $\min(1,r)$ as m grows (meaning that when m is large, nearly the entire short side of the market finds a match).[9] Furthermore, as the following Theorem shows, it is in general possible to order the number of matches that form under each of the three preference structures.

Theorem 2. *For any r, m, $M^I \geq M^C \geq M^A$, with strict inequality unless $m = 1$.*

In assessing the importance of Theorems 1 and 2, one should note that the number of matches that form is depends not only of the preference structure, but also on the market imbalance r, the number of interviews scheduled by each doctor m, and on the assumption that doctors schedule interviews with hospitals uniformly at random. One might wonder whether these factors, rather than the preference structure, are "first order" determinants of the number of matches that form: perhaps differences between M^I, M^C, and M^A are negligible in comparison. This turns out not to be the case.

First, when $r = 1$ and $m = 8$, the expressions from Theorem 1 imply that under aligned preferences, 9.30% of agents go unmatched, whereas 4.87% go unmatched when preferences are independent. For comparison, when preferences are *anti-aligned*[10] fewer than 1% of all agents go unmatched; if we abandon stability constraints altogether, it is possible to reduce this number to 0.03%.[11] This suggests that the random interviewing assumption is *not* a significant driver of the number of unmatched agents; rather, this number depends crucially on the structure relating agents' preferences.

Furthermore, differences across preference structures are significant when compared to differences that result from changes in r and m. For example, if m is increased to 16 and preferences are aligned, then 5.20% of hospitals go unmatched (note that this more than go unmatched when $m = 8$ and preferences are independent). If m is held at 8 and the number of doctors is increased by ten percent (so that there are 1.1 doctors per vacant position), then under aligned preferences 5.40% of hospitals go unmatched (again, more than the case $r = 1$ with independent preferences).

I also want to note that the result $M^I \geq M^C \geq M^A$ does not directly extend to the case where agents have exogenous outside options and thus consider some partners unacceptable. To see why, define the *effective list length* of doctor d to be the number of partners listed by d who also choose to list d. Suppose that doctors decide to list only matches for which $F(u_{dh}) \geq 1 - \varepsilon$, and similarly for hospitals. Then under aligned preferences, d's expected effective list length is $m\varepsilon$, whereas under independent preferences, it is $m\varepsilon^2$; if ε is small, it follows that the former situation produces more matches than the latter. Thus, the right informal reading of Theorem 2 is that for comparable values of r and comparable *effective list lengths* (rather than number of interviews conducted), independent preferences produce the most matches and aligned preferences produce the fewest.

4.1. **Intuition for Theorem 1.**

Suppose that preferences are independent. Because the outcome of the deferred acceptance algorithm does not depend on the order in which doctors propose, we may hold out a single doctor d and run the algorithm on the remainder of the market. The set of hospitals to which d could match is the set of hospitals which interviewed d, and are left with a match that they find inferior

[9]For more careful analysis of the rate at which this convergence occurs, refer to Appendix 9.2.

[10]By this, I mean informally that hospital and doctor evaluations of a match are negatively correlated; formally, I mean that $F(u_{dh}) + F(v_{dh}) = 1$ for all d, h. The claim that fewer than 1% of agents go unmatched in this case comes from simulation results. For a discussion of factors which might cause such negative correlation, refer to Boudreau and Knoblauch (2010).

[11]The number 0.03% comes from a result by Frieze and Melsted (2012).

to d.[12] In a large market, learning that d's first k choices were unavailable to d provides little to no information about the availability of other choices. Thus, from d's perspective, each hospital with which he interviewed should be available to him with some probability p, and their availability should be independent.

Once we know that the market has this structure, all remaining analysis is greatly simplified: for example, the probability that d matches is $1-(1-p)^m$, and it is easily verified that the expected number of proposals made by d is $(1-(1-p)^m)/p$. All that remains is to compute the value p. Because d sends an average of $(1-(1-p)^m)/p$ proposals, we expect a total of $rn(1-(1-p)^m)/p$ proposals to be sent throughout the course of the algorithm. From the point of view of each hospital, each of these proposals is sent to them roughly with probability $1/n$; thus, the number of proposals received by a given hospital should be distributed as a Poisson random variable with mean $\lambda = r(1-(1-p)^m)/p$. Because a hospital matches if and only if it receives at least one proposal, the fraction of hospitals that match should equal $1-e^{-\lambda}$. Since doctors match with probability $1-(1-p)^m$ and the number of doctors and hospitals that match must be equal, we must have that

$$r(1-(1-p)^m) = 1 - e^{-r(1-(1-p)^m)/p}. \tag{4}$$

I show in Appendix 8 that this consistency equation for p has a unique solution for any r, m. Performing a change of variables, define $x = r(1-(1-p)^m)$ to be the fraction of hospitals who match. After expressing p as a function of x, r, m, substituting into (4) and solving for r yields (1).

I now discuss the intuition when hospitals have perfectly correlated preferences. In this case, the deferred acceptance algorithm is equivalent to a serial dictatorship, in which the top-ranked doctor matches to their most preferred hospital (among those interviewed), the second-ranked doctor matches to their most preferred hospital among those remaining, and so on. Let $Z_n(k)$ be the number of hospitals (out of n) who match to doctors ranked k^{th} or above. The doctor ranked $k+1$ fails to match only if all of his interviews were with one of these hospitals, which occurs with probability close to $(Z_n(k)/n)^m$. Thus, we have

$$E[Z_n(k+1) - Z_n(k) \,|\, Z_n(k)] \approx (Z_n(k)/n)^m,$$

If we define $X_n(t) = \frac{1}{n} Z_n(\lfloor nt \rfloor)$, then the above equation suggests that as $n \to \infty$ we should have $E[X_n(t)] \to X(t)$, where X is given by (2).

Although throughout this paper, we maintain the assumption that doctor preferences are fully idiosyncratic, this assumption is made primarily to facilitate a comparison between the three preference structures considered. In particular, the logic behind the derivation of (2) does not in any way rely on the correlation of preferences among doctors, implying that $X(r)$ gives the expected fraction of doctors who match for *any* model of doctor preferences.[13]

4.2. **Intuition for Theorem 2.**

I will illustrate the intuition for Theorem 2 using an example. Consider a market with three doctors and three hospitals in which each doctor interviews with two hospitals. Suppose that interviews are scheduled as illustrated in Figure 1, so that each hospital interviews two doctors. By symmetry, each doctor goes unmatched with equal probability, so it is enough to consider d_2. Note that d_2 goes unmatched if and only if d_1 is matched to h_1, and d_3 to h_3. In order for this to occur, it must be the case that

(A) d_1 and d_3 prefer h_1 and h_3 (respectively) to h_2, and
(B) h_1 and h_3 prefer d_1 and d_3 (respectively) to d_2.

[12]Technically, d's application could trigger a "rejection chain" returning to one of the hospitals with which d interviewed, meaning that the described set of hospitals may be a strict superset of those available to d. However, the probability of such a rejection chain vanishes at the market grows.

[13]This equivalence does not hold when each hospital has the capacity to match to multiple doctors, as I discuss in Appendix 9.1. In this case, introducing correlation among doctor preferences causes still fewer matches to form.

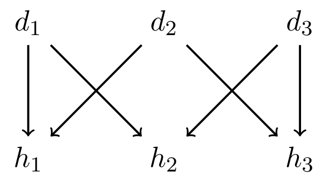


FIGURE 1. A possible interview graph when $r = 1, m = 2, n = 3$. A stable matching may produce either two or three matches, depending on agent preferences.

When preferences are independent, the probability of each event above is 1/4, and thus the probability that d_2 goes unmatched is $\frac{1}{16}$. When hospitals have correlated preferences, (B) occurs whenever d_2 is the lowest-ranked doctor (i.e. with probability $\frac{1}{3}$), and the ex-ante probability that d_2 goes unmatched rises to $\frac{1}{12}$.

Intuitively, the difference between independent and correlated preferences can be explained by noting that doctors can only go unmatched if all of their proposals are rejected, and this should be more likely when some doctors are undesirable to all hospitals.[14] This intuition, however, does not help to explain the result that $M^C \geq M^A$, as under the model of aligned preferences, a doctor who is disfavored by one hospital may nevertheless be very desirable to a second hospital. Indeed, the ranking $M^I \geq M^C \geq M^A$ may seem a bit of a puzzle. After all,

- With correlated preferences, hospitals tend to get proposals from undesirable doctors (since desirable doctors need not apply to many places).
- With aligned preferences, hospitals tend to get proposals from doctors that they find desirable (since doctors start with better matches).
- With independent preferences, the arrival of each proposal is independent of the hospital's assessment of the doctor.

Why, then, is M^I not "in the middle"? The flaw in this logic is that the number of unmatched agents depends not on the desirability of the *typical* applicant, but rather on the desirability of a doctor *who is applying to one of their final choices*. For both correlated and aligned preferences, such a doctor is likely to be turned down (in the first case, because they are likely a low-quality doctor, and in the second because it is likely a low-quality match).

I now provide intuition (and a sketch of a proof) for the result $M^C \geq M^A$. Note that under both correlated and aligned preferences, there is a unique stable matching, which can be constructed through a simple greedy procedure.

P1 Correlated preferences: the top-ranked doctor must receive their most preferred option. Form this match, and apply this principle to the remaining subgraph.

P2 Aligned preferences: the match of highest quality must be in any stable matching. Form this match, and apply this principle to the remaining subgraph.

Returning to the example in Figure 1, by symmetry, each edge is equally likely to be the first one chosen by either greedy procedure; suppose without loss of generality that we begin by matching d_1 and h_1. Now, only two matches form if and only if d_3 and h_3 are matched. Under aligned preferences, this occurs with probability $\frac{1}{3}$, whereas with correlated hospital preferences (and independent doctor preferences) the probability that this occurs is only $\frac{1}{4}$.

Why the difference? Note that procedure P1 selects at each step a random remaining doctor to match, whereas procedure P2 selects each doctor *in proportion to the number of yet-unmatched hospitals with whom he has interviewed*. In other words, the greedy procedure corresponding to

[14]Of course, high quality doctors are *more* likely to match than "typical" doctors in a world with independent hospital preferences. It's not obvious that the decreased matching probabilities for low-quality doctors dominates the increased probabilities for high-quality doctors, though I argue that it is at least intuitive.

aligned preferences tends to match doctors with many remaining options, forcing doctors with few remaining options to wait (and risk that their final option gets matched). In the proof (see Appendix 8), I formalize this idea by using the principle of deferred decisions to define two Markov chains (one corresponding to correlated preferences and the other to aligned preferences), and then coupling these chains to show that fewer matches form under aligned preferences.[15]

5. Results: Quality of Matches

Of course, if the only goal were to maximize the number of matches that form, this could trivially be accomplished. It is also important to consider the *quality* of matches that form. The measure of quality used in this paper is the *quantile* that each agent assigns to their match partner. In other words, we answer questions of the form "What percent of doctors are matched to a hospital which they would consider to be among the top $y\%$ of hospitals?" Given the answer to this question for all y, it is straightforward to compute the expected welfare of each agent, for any given distribution F.

More formally, given a matching μ, and $s \in [0,1]$, let $N_d(s)$ to be the number of doctors who receive utility at most $F^{-1}(s)$ under μ. Define $G_d(s) = \lim_{n\to\infty} \frac{1}{rn} E[N_d(s)]$ to be the large-market fraction of doctors who receive utility at most $F^{-1}(s)$. Similarly, let $N_h(t)$ be the number of hospitals that receive utility at most $F^{-1}(t)$ under μ, and define $G_h(t) = \lim_{n\to\infty} \frac{1}{n} E[N_h(t)]$.

Theorem 3 provides, for each of the three preference structures, expressions for G_d and G_h. Note that this theorem subsumes Theorem 1, as $M^I = G_h^I(0)$, $M^C = G_h^C(0)$, and $M^A = G_h^A(0)$. An important result in this section is Theorem 4, which states that although more matches form under independent preferences, more "high quality" matches form when hospital preferences are correlated. In other words, there exists $\hat{s} < 1$ such that $G_d^I(s) > G_d^C(s)$ for $s \in (\hat{s}, 1)$.

Theorem 3.

(1) Independent preferences:

$$G_d^I(s) = (1 - p(1-s))^m, \qquad G_h^I(t) = e^{-\lambda(1-t)} \tag{5}$$

where $\lambda = r(1-(1-p)^m)/p$ *and* $p \in (0,1]$ *is the unique solution to* (4).

(2) Perfectly correlated preferences:

$$G_d^C(s) = \int_0^1 (s + (1-s)X(rt))^m dt, \qquad G_h^C(t) = 1 - X(r(1-t)), \tag{6}$$

where $X(t)$ *is the solution to the differential equation*

$$X'(t) = 1 - X(t)^m, \qquad X(0) = 0.$$

(3) Aligned preferences:

$$G_d^A(s) = 1 - \frac{1 - \tilde{G}(s)}{r}, \qquad G_h^A(t) = \tilde{G}(t).$$

where $\tilde{G}$ *is the solution to the differential equation*

$$\tilde{G}'(t) = rm\tilde{G}(t)\left(1 - \int_t^1 \tilde{G}(u)du\right)^{m-1}, \qquad \tilde{G}(1) = 1. \tag{7}$$

[15]Thus, this proof holds for markets of arbitrary size (not only in the large market limit), and implies that the distribution of the number of matches that form under correlated preferences stochastically dominates the corresponding distribution for aligned preferences.

While Theorem 3 makes it possible to compute aggregate welfare for any given F, it can also be used to derive structural insights, and provide guidance in the design of school choice procedures (as school preferences are engineered, rather than given exogenously). Several cities have recently redesigned their assignment algorithms to resemble the student-proposing deferred acceptance algorithm. Generally, these procedures sort students into one of several broad classes, and use lotteries to break ties among students within a class.

Two natural tie-breaking procedures are *single tie-breaking*, or STB (which assigns each student a single lottery number, and breaks all ties according to this number), and *multiple tie-breaking*, or MTB (which assigns students one lottery number *per school*, and breaks ties for positions at a given school according to the lottery numbers for that school). In the context of this paper, these procedures naturally correspond to the cases of correlated and independent school preferences, respectively. Theorem 4 states that there is no stochastic dominance relation between the allocations that result from STB and MTB. Instead, the functions G_d^I and G_d^C have a unique intersection point $\hat{s}$: STB produces more matches of quality above $F^{-1}(\hat{s})$, and MTB produces more matches overall.

Theorem 4. *If $m = 1$, $G_d^C = G_d^I$. For $m \geq 2$ and all r, there exists $\hat{s} \in (0,1)$ such that $G_d^C(s) > G_d^I(s)$ for $s \in [0, \hat{s})$ and $G_d^C(s) < G_d^I(s)$ for $s \in (\hat{s}, 1)$.*

Theorem 4 considers the somewhat abstract notion of student utilities, which are hard to observe directly. A more easily measured proxy is the number of students getting their first, second, third choices and so on. It is straightforward to show that a discrete analog of Theorem 4 holds: for any r and $m \geq 2$, there exists $k' < m$ such that STB results in more students getting one of their top k choices for $k \leq k'$, and MTB results in more students getting one of their top k choices for $k > k'$ (for details, see Appendix 8.2). This trend was observed empirically by Abdulkadiroglu et al. (2009); to my knowledge, this paper is the first to provide a theoretical explanation.

Theorem 4 implies that there are tradeoffs when choosing between STB and MTB, and that the correct choice depends on the form of the distribution F. In settings where a small fraction of high-quality matches are of very high value, STB is likely to be preferable; MTB is better suited to settings where the variation in match quality is small relative to the difference between being matched and going unmatched.

In principle, there is no need to restrict attention to these two tie-breaking procedures; one might hope for a procedure which produces match outcomes which stochastically dominates those produced by STB and MTB. I believe that this is an unattainable goal, and that the tradeoff between forming high quality matches (i.e. giving many students their top choices) and forming a large number of total matches is fundamental. My reasoning is best presented through an example. Suppose that there are two candidates, d_1 and d_2, for a position at hospital h. This hospital is the first choice for d_1, whereas it is the final acceptable choice for d_2 (who has already been rejected by his more preferred options). If preferences are such that h is likely to select d_1, this increases the number of doctors who receive their first choice, while causing fewer doctors to match overall; selecting d_2 has the opposite effect.

Even if stochastic dominance of STB and MTB is unattainable, for a given distribution F, other natural mechanisms may outperform either of these procedures. One well-studied alternative assignment algorithm is the so-called "Boston" mechanism, which explicitly maximizes the number of students who receive their first choice, then maximizes the number of students receiving their second choice (given the constraints from previous assignments), and so on. If the distribution of cardinal utilities has a heavy right tail, so that STB is in fact socially preferable to MTB, then it seems likely that the outcome of the Boston mechanism may be socially preferable to either.[16]

[16] On the opposite extreme, if minimizing the number of unmatched students were paramount, then it might be worthwhile to run a multiple tie-breaking procedure in which lottery numbers for each student were *negatively* correlated across schools.

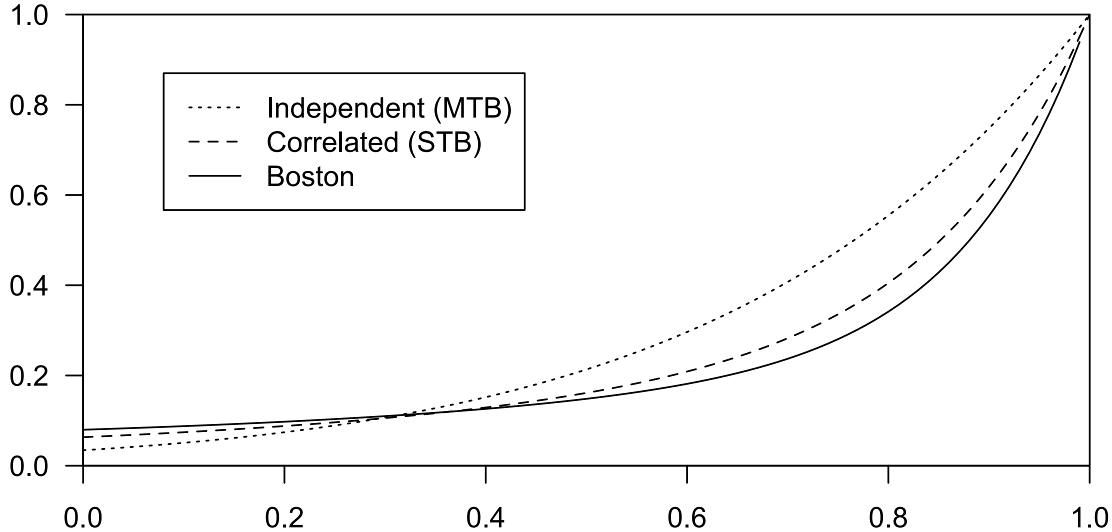


FIGURE 2. Plotting G_d for $r = 1, m = 10$. Note that each pair of curves cross exactly once: single tie-breaking produces the most matches but the fewest high-quality matches, whereas the Boston mechanism produces more high quality matches, at the cost of fewer matches overall.

Within the model discussed in this paper, we may compute the fraction of students matched to one of their top k choices by the Boston mechanism, as well as an equation for $G_d^B(s)$ (the fraction of students assigned by the Boston procedure to a school which they value at $F^{-1}(s)$ or lower). We provide expressions for these quantities in Appendix 8.2, and plot G_d^I, G_d^C, and G_d^B against one another in Figure 2. As expected, Boston provides more top matches than either of the other procedures, while matching fewer students overall.

One well-known flaw of the Boston mechanism is that (unlike STB and MTB deferred acceptance procedures) it induces significant incentive for students to misreport their preferences. For example, a student who covets a position at a highly-ranked school may choose not to list that school in order to ensure a spot at their second choice. Within the model discussed in this paper, this is not a concern: I assume that schools are indistinguishable, so students will not be able to predict which schools will be most competitive, and thus might as well report truthfully. In reality, however, this is a major concern which might prevent the implementation of the Boston mechanism.

It may, however, be possible to partially align preferences by giving students priority at schools that they covet, *so long as these schools can be predicted in advance.* For example, if students tend to prefer schools close to where they live or schools that their siblings attend, then policies which give priorities to these students will tend to increase the number of students who get one of their top choices (likely at the cost of some decrease in the total number of students who match).

6. After-Market

In practice, agents left unmatched by the centralized process may be able to match in some sort of after-market. It seems intuitive that in a large market, the quality of matches available in the aftermarket is insensitive to the behavior of any individual. If this is the case, we can think of this market as providing an outside option of some value $\overline{u}$ to doctors and value $\overline{v}$ to hospitals. The optimal doctor response in this case is to interview as before, and list only hospitals for which $u_{dh} \geq \overline{u}$ (and analogously for hospitals). Of course, the perceived quality of the after-market will in general depend on the "pickiness" of each side, as well as the mechanism through which the aftermarket clears.

In this section, I suppose a very simple clearing mechanism: each unmatched agent on the short side is matched to a random unmatched agent on the long side. We can then define an equilibrium of the game to be a pair $(\overline{u}, \overline{v})$ such that when agents treat these values as their outside options, the resulting aftermarket in fact generates these expected utilities. As a first step towards solving for an equilibrium, it is necessary to study the behavior of the centralized clearinghouse when agents list as unacceptable partners below a given threshold. I do so in Appendix 8.3: given any values $(F(\overline{u}), F(\overline{v})) \in [0,1]^2$, I provide equations for the match quality distribution functions G_h, G_d.

For any distribution F, and any of the three correlation structures studied in this paper, these expressions make it possible to solve for equilibria of the induced game. The set of equilibria depend on the distribution F, as candidates must decide whether to accept a low-quality match or gamble for a better one in the aftermarket. This paper focuses primarily on questions whose answers are insensitive to the choice of F, and I leave a full examination of the possible effects and implications of the aftermarket to future work. I do, however, want to illustrate that the preference structure crucially affects the set of equilibria and the welfare implications of the aftermarket.

First, I note that when preferences are aligned there is a unique (and simple) equilibrium, and that the presence of an aftermarket necessarily increases aggregate welfare for both sides.

Theorem 5. *When preferences are aligned, there is a unique equilibrium outcome of the game with an aftermarket. In this equilibrium, doctor d and hospital h list each other if and only if they interviewed and u_{dh} exceeds the mean match quality. Aggregate welfare for each side increases relative to the game without an aftermarket.*

To see that this is the unique equilibrium, note that every agent will choose to list an above-average partner rather than participate in the aftermarket, and agents on the short side will prefer the aftermarket to accepting a below-average partner. To see that the presence of the aftermarket increases aggregate welfare, consider a case where d and h would have matched absent an aftermarket, but in the presence of an aftermarket, h declines to list d and instead finds d'. Then it must be the case that the expected match quality between h and d' exceeds u_{dh}, and thus both sides of the market benefit (in aggregate) from the change (although the individual doctor d may be worse off).

When preferences are not aligned, however, the welfare effects of the aftermarket are more ambiguous. Suppose, for example, that absent an aftermarket, d and h would match, and u_{dh} is very high. If v_{dh} is not particularly high, the presence of an aftermarket might cause h not to list d. Even if h goes on to match with another doctor d', there is no reason to believe that $u_{d'h} \geq u_{dh}$; thus, aggregate doctor welfare may be lower in the presence of the aftermarket. Example 1 demonstrates that the presence of an aftermarket might decrease welfare for both sides. The example makes use of a class of distributions which I call *binary*: these are distributions such that $u_{dh} \in \{s, 1\}$ (with $s < 1$, and $P(u_{dh} = 1) = P(v_{dh} = 1) = \gamma \in (0, 1)$). I say that hospital h is an *excellent match* for doctor d if $u_{dh} = 1$.

Example 1. Suppose that preferences are independent, and that $u_{dh}, v_{dh} \in \{s, 1\}$, with $s < 1$ and $P(u_{dh} = 1) = \gamma$. If $r = 1, m = 10$, and $\gamma = 1/10$, then[17]

- Without an aftermarket, 96.5% of participants match; over 25% of doctors, and over 28% of hospitals, receive excellent matches.
- With an aftermarket, there is a unique equilibrium in which both sides list only excellent matches. Fewer than 10% of participants match through the clearinghouse; fewer than 19% of participants on each side eventually receive excellent matches.

[17] Here and in Theorem 6, I assume that agents break ties in uniform random order, independent from one another. This is effectively equivalent to taking F to be a continuous perturbation of the binary distribution above.

If $s < 2/3$, the presence of an aftermarket lowers welfare of both sides.

Furthermore, when preferences are not aligned, there may be multiple equilibria. To see the intuition behind this fact, suppose that there is an excess of doctors ($r > 1$) and preferences are independent. Then the utility that each doctor expects to receive from the aftermarket is proportional to the probability that they will find a match there. This probability is higher when many agents on both sides go unmatched, and lower when most agents match. Thus, it may be the case that there are multiple equilibria:

- If doctors are unselective (list many hospitals), then most hospitals match through the clearing house, and the after-market is relatively unappealing.
- If doctors are selective (list relatively few hospitals), then many hospitals remain unmatched by the clearinghouse, and the after-market is relatively appealing.

Theorem 6 fully characterizes the set of equilibria for binary independent preferences. In particular, for any r, m, γ there exist $\underline{s}, \overline{s} \in [0, 1]$ such that there are two equilibria if and only if $s \in (\underline{s}, \overline{s})$.

Theorem 6. *Suppose that preferences are binary and independent. When $r = 1$, there is a unique equilibrium in which agents on both sides accept only excellent matches. Otherwise, in every equilibrium the short side of the market accepts only excellent matches. For any $r, m, \gamma, \exists\ \underline{s} < \overline{s}$ such that*

- *It is an equilibrium for the long side to list only excellent matches iff $s \leq \overline{s}$.*
- *It is an equilibrium for the long side to list all interview partners iff $s \geq \underline{s}$.*

Furthermore, the thresholds $\underline{s}, \overline{s}$ are decreasing in m, increasing in r when $r < 1$, and decreasing in r when $r > 1$.

In the case where preferences are correlated, there is an additional consideration of adverse selection: those agents left in the aftermarket are likely to be undesirable. It turns out that when preferences are correlated and binary, there always exists a unique equilibrium.[18] Aggregate welfare for the long side in this equilibrium may be either higher or lower than welfare without an aftermarket. I leave a treatment of more general (non-binary) distributions to future work.

7. Conclusion/Discussion

This paper examines the role of preference correlation in determining the number and quality of matches formed by centralized clearinghouses. I focus on three canonical preference structures: independent, correlated, and aligned preferences. The first of these assumes no correlations between agent preferences. The second assumes strong positive correlations in the preferences of agents on a single side of the market. The third assumes correlations in preferences *across* sides of the market; that is, agents tend to agree on whether they would be a suitable match.

My first finding is that more matches form when preferences are independent than when preferences on at least one side are perfectly correlated. Perhaps more surprisingly, I show that the case of aligned preferences generates the *fewest* matches, and that the differences across preference structures may be substantial. The intuition for these results is that the primary factor determining the number of agents who go unmatched is the probability that an agent applying to one of their final choices will be accepted. Under both correlated and aligned preferences, this probability is relatively low (in the first case, because such an agent is likely to be of low quality, and in the latter case, because the match is likely to be a marginal fit).

Of course, we are interested not only in the number of matches that form, but also in their quality. In Section 5, I provide methods for computing, for each preference structure, the fraction of agents on each side who receive partners that they rank above any given threshold. I apply these results in the context of school assignment. In particular, I show that when compared to a

[18]A proof of this fact is available upon request.

lottery with multiple tie-breaking, single tie-breaking results in more students being assigned to one of their top few choices, but fewer students being assigned overall. I provide reasons to believe that the tradeoff between forming more *high quality* matches and forming more *total* matches is fundamental.

In many centralized markets, agents left unmatched by the clearinghouse have an opportunity to match to one another (through formal or informal channels). The presence of such an aftermarket has direct benefits to the matched agents, but also changes incentives during the primary matching stage, as the opportunity to match later provides agents with an outside option. I demonstrate that when agent preferences are aligned, the introduction of an aftermarket always increases welfare, but for other preference structures the aggregate welfare consequences of an aftermarket may be either positive or negative. Furthermore, multiple equilibria may emerge, as there exist same-side complementarities: a given doctor can more safely be "selective" in who they list when other doctors do the same.

The model used in this paper makes several simplifying assumptions. In the remainder of this conclusion, I discuss two of the most salient: the assumption that hospitals accept all interviews, and the assumption that agents appear identical ex-ante.

7.1. **Alternate Interview Scheduling Procedures.**

For technical convenience, the process of scheduling of interviews assumed in this paper is very simple, and not descriptively accurate when applied to the NRMP.[19] Even if agents possess no information about one another ex-ante, hospitals may only be able to interview a limited number of candidates, and agents on both sides may have incentives to coordinate interviews in order to avoid the situation where some programs receive few interview requests and others receive many.

One possible revision would be to assume a two-round interview scheduling process whereby doctors request up to m interviews, and hospitals with too many requests decline some applicants. In one extreme of this model, doctors apply broadly, and most hospitals have a full schedule of interviews and reject many applicants; this case should resemble the model presented above, with the role of doctors and hospitals reversed. One could also specify more complicated interview formation processes involving multiple rounds of choices by agents on both sides.

A more realistic and/or optimized interview scheduling process might plausibly affect the number of matches that form in the market, and the challenges of modeling and optimizing interview formation are quite interesting. Thus, I do not consider the expressions derived in Theorems 1 and 3 to be sacrosanct. Instead, I emphasize the fact that for a *given* interview procedure, different preference structures have very different implications for the quantity and caliber of matches that form. Furthermore, these effects may be as large or larger than the differences between different procedures for scheduling interviews.[20] None of the intuition underlying these differences relies heavily on the details of the interview formation process, and I expect that qualitatively similar results would emerge from other models of interview formation.

7.2. **Ex-Ante Heterogeneity and Targeted Interviews.**

The previous section discussed alternate interview scheduling procedures, holding fixed the assumption of low information (i.e. the fact that agents appear identical ex-ante). In the model above, even if if both sides of the market are vertically differentiated ex-post and doctors know their own quality, they cannot tailor their strategy to reflect this information (as it is costless to

[19]For information on the number of interviews requested and accepted by residency programs in 2014, see `http://www.nrmp.org/wp-content/uploads/2014/09/PD-Survey-Report-2014.pdf`.

[20]Recall that with $r = 1, m = 8$, and uncoordinated interview scheduling, the model predicts that 4.87% of agents go unmatched when preferences are independent. This number rises to 9.30% when preferences are aligned, and falls to less than 1% when preferences are anti-aligned. By comparison, when preferences are independent, coordinating interviews so that each hospital interviews eight candidates (without further coordinating the interview schedule in the manner studied by Lee and Schwarz (2012)) causes a significant but more modest drop in the unemployment rate: it falls from 4.87% to 3.20%.

interview at the maximum allowable number of hospitals, and they do not initially observe hospital quality). The doctors who fail to match tend to be low-ranked doctors who happen to have applied primarily to highly ranked programs. In practice, we might expect that some signals of program quality are available before the scheduling of interviews, and that doctors use this information to guide their applications. One might wonder how the conclusions in this paper are affected by this sorting behavior.

Solving for an equilibrium of the game with known qualities on both sides is a daunting challenge; here, I merely posit plausible-sounding behavior on the part of doctors. Suppose that doctors choose to apply only to hospitals of a caliber similar to their own, meaning that a doctor who knows himself to be of quality q applies uniformly at random among the set of hospitals whose quality is in the interval $[q-\epsilon, q+\epsilon]$ for some $\epsilon > 0$. Suppose furthermore that hospitals have sufficiently idiosyncratic preferences (or sufficiently noisy signals of doctor quality) that their ranking of doctors *who applied to them* is effectively uniformly random (recall that all applicants to a given hospital differ in quality by at most 2ϵ). In this case, Theorems 3 and 4 by Arnosti et al. (2014) imply that from the perspective of any individual agent, the market "looks" like a market with uniformly random preferences.[21] As we now know, this implies that a smaller (though still non-trivial) proportion of agents on each side will go unmatched (relative to the case where doctors do not strategically target their interviews). Thus, if we believe that doctors primarily schedule interviews only with hospitals of comparable quality, the work of Arnosti et al. (2014) provides rigorous justification for the "independent preferences" analysis in this paper.

Another extension that could be considered is one in which the number of interviews is unlimited, but each one is costly. If preferences are idiosyncratic, an equilibrium of this model should look very similar to the outcomes discussed in this paper. However, if hospital preferences are correlated and doctors observe their own quality, then doctors of different qualities may choose to send different numbers of applications. Analyzing such a model is one possible route for future work.

[21]Formally, each doctor perceives that hospitals they apply to will extend an offer independently with probability p, where p is the unique solution to a consistency equation. Arnosti et al. (2014) frame their result quite differently, as in their model, the vertical component of preferences is assumed to be time (in other words, applicants apply to positions which were recently posted, and accept the first offer that they receive), but mathematically, the two formulations are effectively equivalent. One difference is that in their model, applicants apply to each position independently with equal probability (whereas this paper assumes that doctors send exactly m applications). I study this case in Appendix 9.2.

References

Abdulkadiroglu, A., Pathak, P. A., and Roth, A. E. (2009). Strategy-proofness versus efficiency in matching with indifferences: Redesigning the new york city high school match. Working Paper 14864, National Bureau of Economic Research.

Arnosti, N., Johari, R., and Kanoria, Y. (2014). Managing congestion in decentralized matching markets. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 451–451. ACM.

Ashlagi, I., Kanoria, Y., and Leshno, J. (2013). Unbalanced random matching markets. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, EC '13, pages 27–28, New York, NY, USA. ACM.

Avery, C., Lee, S., and Roth, A. E. (2014). College admissions as non-price competition: The case of south korea. Working Paper 20774, National Bureau of Economic Research.

Balinski, M. and Sönmez, T. (1999). A tale of two mechanisms: Student placement. *Journal of Economic Theory*, 84(1):73 – 94.

Biró, P. (2007). Higher education admission in hungary by a score-limit algorithm. In *The 18th International Conference on Game Theory at Stony Brook University*.

Boudreau, J. W. and Knoblauch, V. (2010). Marriage matching and intercorrelation of preferences. *Journal of Public Economic Theory*, 12(3):587–602.

Braun, S., Dwenger, N., and Kübler, D. (2010). Telling the truth may not pay off: An empirical study of centralized university admissions in germany. *The BE Journal of Economic Analysis & Policy*, 10(1).

Dietzfelbinger, M., Goerdt, A., Mitzenmacher, M., Montanari, A., Pagh, R., and Rink, M. (2010). Tight thresholds for cuckoo hashing via xorsat. In Abramsky, S., Gavoille, C., Kirchner, C., Meyer auf der Heide, F., and Spirakis, P., editors, *Automata, Languages and Programming*, volume 6198 of *Lecture Notes in Computer Science*, pages 213–225. Springer Berlin Heidelberg.

Fountoulakis, N. and Panagiotou, K. (2012). Sharp load thresholds for cuckoo hashing. *Random Structures & Algorithms*, 41(3):306–333.

Frieze, A. and Melsted, P. (2012). Maximum matchings in random bipartite graphs and the space utilization of cuckoo hash tables. *Random Structures & Algorithms*, 41(3):334–364.

Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):pp. 9–15.

Immorlica, N. and Mahdian, M. (2005). Marriage, honesty, and stability. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '05, pages 53–62, Philadelphia, PA, USA. Society for Industrial and Applied Mathematics.

Kojima, F., Pathak, P. A., and Roth, A. E. (2013). Matching with couples: Stability and incentives in large markets. *The Quarterly Journal of Economics*, 1(48):48.

Lee, R. S. and Schwarz, M. (2012). Interviewing in two-sided matching markets.

Lee, S. and Yariv, L. (2014). On the efficiency of stable matchings in large markets. Technical report, mimeo.

Mastin, A. and Jaillet, P. (2013). Greedy online bipartite matching on random graphs. *arXiv preprint arXiv:1307.2536*.

Pathak, P. A. (2011). The Mechanism Design Approach to Student Assignment. *Annual Review of Economics*, 3(1):513–536.

Pittel, B. (1989). The average number of stable matchings. *SIAM Journal on Discrete Mathematics*, 2(4):530–549.

Rastegari, B., Condon, A., Immorlica, N., and Leyton-Brown, K. (2013). Two-sided matching with partial information. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, EC '13, pages 733–750, New York, NY, USA. ACM.

Roth, A. E. (1984). The evolution of the labor market for medical interns and residents: A case study in game theory. *Journal of Political Economy*, 92(6):pp. 991–1016.

Roth, A. E. and Peranson, E. (1999). The redesign of the matching market for american physicians: Some engineering aspects of economic design. Working Paper 6963, National Bureau of Economic Research.

Wormald, N. C. (1995). Differential equations for random processes and random graphs. *The Annals of Applied Probability*, 5(4):pp. 1217–1235.

8. Appendix: Proofs

8.1. **Proofs for Section 4.**

I begin with two technical lemmas, which are very closely related to one another.

Lemma 1. *For every $r > 0, m \in \mathbb{N}, \gamma \in (0,1]$, there exists a unique $x \in (0,1)$ satisfying*

$$x = r\left(1 - \left(1 + \frac{\gamma x}{\log(1-x)}\right)^m\right), \tag{8}$$

Proof. Note that the expressions on both sides of (8) are continuous on the interval $(0,1)$. The left side starts at zero, ends at one, and is increasing. The right side starts at $r(1-(1-\gamma)^m)$, ends at zero, and is decreasing. Thus, they must cross at a single interior point x. □

Lemma 2. *For every $r > 0, m \in \mathbb{N}$, there exists a unique $p \in (0,1]$ satisfying*

$$r(1-(1-p)^m) = 1 - e^{-r(1-(1-p)^m)/p}. \tag{9}$$

Proof. Define the function $x_1(p) = r(1-(1-p)^m)$, and define $x_2(p)$ to be the solution to $p = -x_2/\log(1-x_2)$. Note that x_2 is well-defined and continuous, as $-x/\log(1-x)$ decreases continuously from one to zero on the interval $[0,1]$.

Rearranging (9), note that we are seeking a solution to $p = -x_1(p)/\log(1-x_1(p))$, or equivalently a value p for which $x_1(p) = x_2(p)$. Such a solution exists because x_1 begins below x_2 ($x_1(0) = 0, x_2(0) = 1$) and ends above it ($x_1(1) = r, x_2(1) = 0$), and both functions are continuous. Uniqueness of the solution follows because x_1 is increasing and x_2 is decreasing. □

Proof of Theorem 1.

This theorem is an immediate consequence of Theorem 3, as

$$\lim_{n\to\infty} \frac{1}{n} E[|\mu_n^I|] = 1 - G_h^I(0)$$
$$\lim_{n\to\infty} \frac{1}{n} E[|\mu_n^C|] = 1 - G_h^C(0)$$
$$\lim_{n\to\infty} \frac{1}{n} E[|\mu_n^A|] = 1 - G_h^A(0).$$

□

Proof of Theorem 2.

Proof that $M^I \geq M^C$:

For fixed m, define $X(t)$ by

$$X(0) = 0, \qquad X'(t) = 1 - X(t)^m,$$

and define $Y(r)$ to be the unique solution (see Lemma 1) to

$$r = \frac{Y(r)}{1 - \left(1 + \frac{Y(r)}{\log(1-Y(r))}\right)^m}.$$

Theorem 1 states that $M^C = X(r)$ and $M^I = Y(r)$. Because $X(\cdot)$ is increasing, to show that $X(r) \leq Y(r)$ for al r, it suffices to show that for $y \in [0,1)$,

$$X\left(\frac{y}{1 - \left(1 + \frac{y}{\log(1-y)}\right)^m}\right) \leq y. \tag{10}$$

For $y \in [0,1)$, define $u = 1 + \frac{y}{\log(1-y)} < y$. Note that the inequality in (10) is tight for $y = 0$. Furthermore, I claim that whenever (10) is tight, the left side of (10) grows slower than the right side. To show this, note that

$$\begin{aligned}\frac{d}{dy} X\left(\frac{y}{1-u^m}\right) &= X'\left(\frac{y}{1-u^m}\right)\left(\frac{1}{1-u^m} + \frac{myu^{m-1}}{(1-u^m)^2}\frac{du}{dy}\right). \\ &= \left(1 - X\left(\frac{y}{1-u^m}\right)^m\right)\left(\frac{1}{1-u^m} + \frac{myu^{m-1}}{(1-u^m)^2}\frac{du}{dy}\right) \\ &= (1-y^m)\left(\frac{1}{1-u^m} + \frac{myu^{m-1}}{(1-u^m)^2}\frac{du}{dy}\right). \qquad (11)\end{aligned}$$

I must show that the expression in (11) is less than one. Multiply by $\frac{1-u^m}{1-y^m}$ and subtract one to see that this is equivalent to

$$\frac{y^m - u^m}{1-y^m} > \frac{mu^{m-1}}{1-u^m}\, y\frac{du}{dy}. \qquad (12)$$

Note furthermore that

$$y\frac{du}{dy} = \frac{y}{\log(1-y)}\left(1 + \frac{1}{1-y}\frac{y}{\log(1-y)}\right) = (u-1)\left(1 + \frac{u-1}{1-y}\right) = \frac{(1-u)(y-u)}{1-y}.$$

Substituting into (12), we see that we must show that

$$\frac{y^m - u^m}{1-y^m} > \frac{mu^{m-1}}{1-u^m}\frac{(1-u)(y-u)}{1-y}.$$

Rearranging, we must show that

$$\frac{1}{u^{(m-1)}}\frac{y^m-u^m}{y-u}\frac{1-u^m}{1-u} > m\frac{1-y^m}{1-y},$$

or equivalently, that

$$\sum_{k=0}^{m-1}\left(\frac{y}{u}\right)^k \sum_{k=0}^{m-1} u^k > m\sum_{k=0}^{m-1} y^k.$$

This is a version of Chebyshev's inequality; to see that it holds, subtract the expression on the left from both sides and rearrange to get that the above inequality is equivalent to

$$\sum_{0\le j\le k<m} (u^j - u^k)\left(\left(\frac{y}{u}\right)^j - \left(\frac{y}{u}\right)^k\right) < 0.$$

This holds because $u < 1$ and $y/u > 1$, so each term in the sum is negative.

Proof that $M^C > M^A$:

The proof leverages the fact that under both correlated and aligned preferences, there is a unique stable matching, which can be constructed through a simple greedy procedure.

- P1 Correlated preferences: the top-ranked doctor must receive their most preferred option. Form this match, and apply this principle to the remaining subgraph.
- P2 Aligned preferences: the match of highest quality must be in any stable matching. Form this match, and apply this principle to the remaining subgraph.

Thus, the result $M^C \geq M^A$ is formally a statement about the sizes of maximal matchings in large bipartite random graphs. M^C is the expected size of the matching formed by an algorithm which at each step selects a vertex (doctor) uniformly at random, and matches it to a random neighbor (interpreted as their most preferred remaining hospital). M^A is the expected size of the matching formed when instead an *edge* is selected uniformly at random at each stage.

To argue that the first of these produces more matches in expectation, I construct two Markov chains, (N^A, Y) and (N^C, Y), on $\{N \in \mathbb{N}^m : ||N||_1 \leq \lfloor rn \rfloor\} \times \{0, \ldots, n\}$. Intuitively, N tracks the number of doctors with k outstanding applications to the Y unmatched hospitals, for $k = 1, \ldots, m$. Both chains start in the state $N = (0, \ldots, 0, \lfloor rn \rfloor), Y = n$, indicating that initially, all $\lfloor rn \rfloor$ doctors have scheduled m interviews with the n unmatched hospitals. I will write down transition dynamics for these chains such that the number of matches formed corresponds to the hitting time of the absorbing set $\{(N, Y) : N = (0, \ldots, 0)\}$. To prove the claim that $E[M^C] \geq E[M^A]$, I provide a coupling between N^A and N^C such that N^C is (weakly) larger, component-wise, than N^A.

The transition dynamics for the chain corresponding to perfectly correlated preferences are as follows:

C1. Sample $k^C \in \{1, \ldots, m\}$, with $P(k^C = i) = N_i^C / \sum_j N_j^C$.
C2. For $i \in \{1, \ldots, m\}$, independently sample $B_i^C \sim \text{Binom}(N_i^C - \mathbf{1}_{(k^C=i)}\ ,\ i/Y)$.
C3. Transition to state $(\tilde{N}^C, Y - 1)$, where

$$\tilde{N}_m^C = N_m^C - B_m^C - \mathbf{1}_{(k^C=m)}, \text{ and } \tilde{N}_i^C = N_i^C - B_i^C + B_{i+1}^C - \mathbf{1}_{(k^C=i)} \text{ for } i = 1, \ldots, m - 1.$$

The first step corresponds to selecting a random doctor d to match; k^C represents the number of remaining applications that d had. The second step corresponds to determining the number of doctors who also had applied to the hospital that d selects; if d' had i outstanding applications among the Y remaining hospitals, then the chance that d' had scheduled an interview with the hospital selected by d is i/Y. The final step updates the number of doctors who have interviewed with exactly i of the still-unmatched hospitals.

The transition dynamics for the chain corresponding to aligned preferences are as follows:

A1. Sample $k^A \in \{1, \ldots, m\}$, with $P(k^A = i) = iN_i^A / \sum_j jN_j^A$.
A2. For $i \in \{1, \ldots, m\}$, independently sample $B_i^A \sim \text{Binom}(N_i^A - \mathbf{1}_{(k^A=i)}\ ,\ i/Y)$.
A3. Transition to state $(\tilde{N}^A, Y - 1)$, where

$$\tilde{N}_m^A = N_m^A - B_m^A - \mathbf{1}_{(k^A=m)}, \text{ and } \tilde{N}_i^A = N_i^A - B_i^A + B_{i+1}^A - \mathbf{1}_{(k^A=i)} \text{ for } i = 1, \ldots, m - 1.$$

Note that the chains differ only in the first step: the selection of k. The chain for correlated preferences selects a doctor uniformly at random (among those with at least one interview remaining), whereas the chain for aligned preferences selects each doctor in proportion to the number of interviews that they have remaining.

Say that vector N^C *dominates* vector N^A if and only if for all $k \in \{1, \ldots, m\}$, we have $\sum_{j \geq k} N_j^C \geq \sum_{j \geq k} N_j^A$. I claim that it is possible to couple (N^C, Y) and (N^A, Y) such that for all fixed Y, N^C dominates N^A. This immediately implies that the hitting time to $\{(N, Y) : N = (0, \ldots, 0)\}$ is smaller for N^A than for N^C.

Because the initial states of the two chains are identical, it is enough to inductively argue that if N^C dominates N^A at time t, then the chains can be coupled such that dominance continues to hold at time $t + 1$. The coupling is as follows. First, correlate k^C and k^A such that $P(\{\sum_{j \geq k^C} N_j^C = \sum_{j \geq k^C} N_j^A\} \cap \{k^A < k^C\}) = 0$. In other words, whenever the constraint that N^C dominates N^A is "tight" at k^C, it must be that $k^A \geq k^C$. This is possible because if $\sum_{j \geq k} N_j^C = \sum_{j \geq k} N_j^A$, then

$$P(k^A \geq k) = \frac{\sum_{j \geq k} jN_j^A}{\sum_j jN_j^A} \geq \frac{\sum_{j \geq k} N_j^A}{\sum_j N_j^A} = \frac{\sum_{j \geq k} N_j^C}{\sum_j N_j^A} \geq \frac{\sum_{j \geq k} N_j^C}{\sum_j N_j^C} = P(k^C \geq k).$$

Note that the second inequality follows because dominance implies that $\sum_j N_j^C \geq \sum_j N_j^A$. It follows that under this coupling, for $k \in \{1, \ldots, m\}$, we have

$$\sum_{j \geq k} N_j^C - \mathbf{1}_{(k^C \geq k)} \geq \sum_{j \geq k} N_j^A - \mathbf{1}_{(k^A \geq k)}.$$

In other words, if we let e_k be the m-dimensional vector with a one in position k and zeros elsewhere, $N^C - e_{k^C}$ dominates $N^A - e_{k^A}$.

In the next stage of the coupling, we correlate B^C and B^A. For each i, let $M_i = (N_i^A - \mathbf{1}_{(k^A=i)}) - (N_i^C - \mathbf{1}_{(k^C=i)})$. Then generate B_i^C and B_i^A as follows:

- If $M_i \geq 0$, generate B_i^C and let $B_i^A = B_i^C + \text{Binom}(M_i, i/Y)$
- Otherwise, generate B_i^A and let $B_i^C = B_i^A + \text{Binom}(|M_i|, i/Y)$

It is clear that the marginal distributions of B_i^C and B_i^A are correct. Furthermore, I claim that under this coupling, $\tilde{N}^C$ dominates $\tilde{N}^A$ (recall that these are the next state of the Markov chains, and are defined in steps C3. and A3. above). That is, I claim that for each k, $\sum_{j \geq k} \tilde{N}_j^C \geq \sum_{j \geq k} \tilde{N}_j^A$. Note that by the definition of $\tilde{N}^C$, for every k we have that

$$\sum_{j \geq k} \tilde{N}_j^C = \sum_{j \geq k} N_j^C - \mathbf{1}_{(k^C \geq k)} - B_k^C, \qquad \sum_{j \geq k} \tilde{N}_j^A = \sum_{j \geq k} N_j^A - \mathbf{1}_{(k^A \geq k)} - B_k^A.$$

Thus, we must show that for $k \in \{1, \ldots, m\}$, we have

$$\sum_{j \geq k} N_j^C - \mathbf{1}_{(k^C \geq k)} - B_k^C \geq \sum_{j \geq k} N_j^A - \mathbf{1}_{(k^A \geq k)} - B_k^A. \tag{13}$$

First consider the case when $M_k \geq 0$. In this case, we have $B_k^C \leq B_k^A$, and thus (13) is immediate from the fact that $N^C - e_{k^C}$ dominates $N^A - e_{k^A}$. When $M_k < 0$, the coupling of B_k^C and B_k^A implies that $B_k^C + M_k \leq B_k^A$. Furthermore, the fact that $N^C - e_{k^C}$ dominates $N^A - e_{k^A}$ implies that $\sum_{j > k} N_j^C - \mathbf{1}_{(k^C > k)} + M_k \geq \sum_{j > k} N_j^A - \mathbf{1}_{(k^A > k)}$. Combining these inequalities shows that (13) continues to hold. Thus, the chains can be coupled such that N^C dominates N^A; the claim that $E[|\mu_n^C|] \geq E[|\mu_n^A|]$ follows immediately. ☐

8.2. **Proofs for Section 5.**

Given random variables X and Y with common domain $\mathcal{D}$, define the *total variation distance* between X and Y to be $d(X, Y) = \sup_{A \subseteq \mathcal{D}} |P(X \in A) - P(Y \in A)|$.

Given a sequence of random variables X_n, we say that X_n *converges in distribution* to X (written $X_n \to X$) if for each x such that $P(X = x) = 0$, it holds that $P(X_n \leq x) \to P(X \leq x)$. When $X = x$ deterministically, this is equivalent to saying that for all $\epsilon > 0$, $P(|X_n - x| > \epsilon) \to 0$.

Lemma 3. *Suppose that preferences are independent. Given r, m, let P_n denote the* total *number of applications submitted by doctors during the run of the deferred acceptance algorithm when there are n hospitals. The value P_n is concentrated about its mean; that is, $P_n / E[P_n] \to 1$ as $n \to \infty$.*

The proof of this Lemma follows a standard path. Consider the martingale which results from sequentially revealing random individual edges (interviews); formally, define M_k to be the expected value of P_n, conditioned on the realization of the first k interviews. Because with high probability, at least $\frac{n}{2}e^{-rm}$ hospitals receive no interview requests, with high probability the revelation of a single edge can only change the expected number of interviews by a constant; that is, the differences $M_{k+1} - M_k$ are bounded by a constant depending only on r and m (and not on n). From this point, the result follows from standard concentration results.

My proof of Theorem 3 uses the following result, which describes a large market with independent preferences from the perspective of representative individuals on each side:

- From the perspective of each doctor d, each proposal that d makes will be independently accepted with probability p given by (9).
- From the perspective of each hospital h, the number of applications received during the deferred acceptance procedure is Poisson with parameter $\lambda = r(1-(1-p)^m)/p$.

Theorem 7. *Suppose that preferences are independent. Given $r > 0, m \in \mathbb{N}$, let p be defined by (9), and define $\lambda = r(1-(1-p)^m)/p$ and $R^I = \min(Geom(p) - 1, m)$. Let R^I_n be a random variable representing the number of times that doctor d is rejected in the n^{th} market, and let N^I_n be a random variable representing the number of proposals received by hospital h. Then $R^I_n \to R^I$ and $N^I_n \to Pois(\lambda)$.*

The proof of Theorem 3 also makes use of an analogous result for the case of correlated preferences.

Theorem 8. *Given $r > 0, m \in \mathbb{N}$, define $X(t)$ by*

$$X(0) = 0, \qquad X'(t) = 1 - X(t)^m.$$

Let $Z_n(k)$ denote the random variable representing the number of hospitals that match to one of the k doctors of the highest quality in the n^{th} market. Then for any $t \in (0,1)$

$$\frac{1}{n} Z_n(\lfloor rtn \rfloor) \to X(rt).$$

Define R^C to be a random variable satisfying $R^C \in \{0, \ldots, m\}$, with

$$P(R^C \geq k) = \int_0^1 X(rt)^k dt \quad \text{for } k \in \{0, \ldots, m\}.$$

In the case where hospital preferences are correlated (and doctor preferences independent), let R^C_n be a random variable representing the number of times that doctor d is rejected in the n^{th} market. Then $R^C_n \to R^C$.

The proofs of these theorems appear at the end of the section; for now, I proceed taking them as given.

Proof of Theorem 3.
Consider the case of independent preferences; by symmetry and linearity of expectation, in order to compute G^I_d and G^I_h, it suffices to consider the match probabilities for a single doctor and/or hospital.

Fix a doctor d; by Theorem 7, the distribution of the number of hospitals who would accept d's proposal converges to Binom(m, p), with p the unique solution to (9). Furthermore, u_{dh} is independent from whether h would accept d's proposal, implying that conditioned on their being k hospitals willing to accept a proposal from d, the probability that d gets a match of quality below $F^{-1}(s)$ is s^k. Thus,

$$P(u_{d\mu^I_n(d)} \leq F^{-1}s) \to \sum_{k=0}^{m} \binom{m}{k} p^k (1-p)^{m-k} s^k = (1 - p(1-s))^m. \tag{14}$$

Similarly, Theorem 7 states that distribution of the number of applications received by a given hospital h converges to Pois(λ), where $\lambda = r(1-(1-p)^m)/p$. If h receives N_h proposals, then the distribution of $v_{\mu(h)h}$ is distributed as the maximum of N_h draws from F, so $P(v_{\mu(h)h} \leq F^{-1}(t)|N_h) = t^{N_h}$. It follows that

$$P(v_{\mu^I_n(h)h} \leq F^{-1}(t)) \to \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} t^k = e^{-\lambda(1-t)}. \tag{15}$$

When considering correlated preferences, we may again compute G_d^I and G_h^I by considering the market from the perspective of a single doctor and/or hospital.

For a given doctor d, the number of hospitals with whom d interviewed and $u_{dh} \geq F^{-1}(s)$ is distributed as Binom$(m, 1-s)$. In the case that there are k such hospitals, d fails to obtain a match of quality above $F^{-1}(s)$ if and only if all k of these hospitals rejects d. By Theorem 8, the probability of this event converges to $P(R^C \geq k)$ as n grows (note that we are using the fact that u_{dh} is independent from the availability of h to d). Thus,

$$\begin{aligned} P(u_{d\mu_n^C(d)} \leq F^{-1}(s)) &\to \sum_{k=0}^{m} \binom{m}{k} (1-s)^k s^{m-k} P(R^C \geq k) \\ &= \int_0^1 \sum_{k=0}^{m} \binom{m}{k} (1-s)^k s^{m-k} X(rt)^k dt \\ &= \int_0^1 (s + (1-s)X(rt))^m dt. \end{aligned}$$

As for computing G_h^C, it is an immediate consequence of the model that for any $t \in [0,1]$, if $Q_n(t)$ denotes the number of doctors with $F(q_d) \geq t$ in the n^{th} market, then $\frac{1}{rn} Q_n(t) \to 1-t$. If we define $Z_n(k)$ to be the (random) number of hospitals that match to one of the k doctors of the highest quality in the n^{th} market, then by the definition of G_d^C and by Theorem 8, we have

$$\left| X(r(1-t)) - (1 - G_d^C(t)) \right| \leq r \lim_{n\to\infty} d\left(\frac{1}{n} Z_n(\lfloor rn(1-t) \rfloor), \frac{1}{n} Z_n(Q_n(t)) \right) = 0.$$

When considering the case of aligned preferences, I derive the expressions for G_d^A and G_h^A below, while omitting a formal proof.

First, note that for any finite market and any matching μ, the number of hospitals and doctors receiving a match above any given threshold must be equal; thus, $\lfloor rn \rfloor - N_d^A(s) = n - N_h^A(s)$ for all s. It follows that

$$r(1 - G_d^A(s)) = 1 - G_h^A(s) \qquad \forall s \in [0,1]. \tag{16}$$

Furthermore, it is possible to derive $G_d^A(s)$ as a function of G_h^A. In particular, if d and h interviewed and $F(u_{dh}) = t$, then the probability that hospital h would accept an offer from d is $G_h^A(t)$. It follows that the probability that d matches to a hospital h for which $F(u_{dh}) \geq s$ is given by

$$1 - G_d^A(s) = 1 - \left(1 - \int_s^1 G_h^A(t) dt\right)^m. \tag{17}$$

Combining this with (16) yields the expression

$$1 - G_h^A(s) = r(1 - G_d^A(s)) = r\left(1 - \left(1 - \int_s^1 G_h^A(t)dt\right)^m\right).$$

Differentiating the above with respect to s yields (7). □

Some of the following proofs make use of a version of Chebyshev's inequality, given below.

Lemma 4. *For any $f : \mathbb{R} \to \mathbb{R}_+$, any $A \subseteq \mathbb{R}$ and any $\gamma \geq 1$,*

$$\int_A f(t)^\gamma dt \geq \left(\int_A f(t)dt\right)^\gamma.$$

I now provide the discrete analog of Theorem 4, which considers the rank that students gave their matches. This Theorem requires some additional notation. Given values u_{dh}, v_{dh} and a matching μ, define $\hat{N}_d^I(k)$ to be the number of doctors who fail to match to a hospital ranked among their top k when preferences are independent, and let $\hat{G}^I_d(k) = \lim_{n\to\infty} \frac{1}{rn} E[\hat{N}^I_d(k)]$. Define $\hat{N}^C_d(k)$ and $\hat{G}^C_d(k)$ analogously.

Theorem 9 (Ranks of STB vs MTB)**.** *For $m \geq 2$, and all r, the following holds: $\hat{G}^C_d(1) < \hat{G}^I_d(1)$. Furthermore, there exists $1 \leq k' < m$ such that $\hat{G}^C_d(k) \leq \hat{G}^I_d(k)$ for $k \leq k'$ and $\hat{G}^C_d(k) > \hat{G}^I_d(k)$ for $k > k'$.*

Proof of Theorem 9.

From Theorems 7 and 8, we have that

$$\hat{G}^I_d(k) = (1-p)^k, \qquad \hat{G}^C_d(k) = \int_0^1 X(rt)^k dt.$$

Suppose that for some k, $\hat{G}^C_d(k) > \hat{G}^I_d(k)$, i.e.

$$\int_0^1 X(rt)^k dt > (1-p)^k.$$

It follows that

$$\begin{aligned}\int_0^1 X(rt)^{k+1} dt &= \int_0^1 X(rt)^{k\cdot\frac{k+1}{k}} dt \\ &\geq \left(\int_0^1 X(rt)^k dt\right)^{\frac{k+1}{k}} \\ &> (1-p)^{k+1}.\end{aligned}$$

It immediately follows that $\hat{G}^I_d$ and $\hat{G}^C_d$ either cross once or never.

We already know from Theorem 2 that for $m \geq 2$,

$$1 - M^I = \hat{G}^I_d(m) < 1 - X(r) = \hat{G}^C_d(m) = 1 - M^C.$$

That is, more doctors get one of their top m choices when hospital preferences are independent. To complete the proof, I show that more doctors get their first choices when hospital preferences are correlated. Seeking a contradiction, suppose that $1 - p \leq \int_0^1 X(rt)dt$. By Theorem 3, we have $1 - M^I = e^{-r(1-(1-p)^m)/p}$, so we conclude that

$$\begin{aligned}-\log(1-M^I)/r = (1-(1-p)^m)/p &= \sum_{k=0}^{m-1} (1-p)^k \\ &\leq \sum_{k=0}^{m-1}\left(\int_0^1 X(rt)dt\right)^k \leq \int_0^1 \sum_{k=0}^{m-1} X(rt)^k dt \\ &= \int_0^1 \frac{1-X(rt)^m}{1-X(rt)} dt = \int_0^1 \frac{X'(rt)}{1-X(rt)} dt \\ &= \frac{1}{r}\int_0^r \frac{X'(u)}{1-X(u)} du = -\log(1-X(r))/r. \\ &= -\log(1-M^C)/r.\end{aligned}$$

The first inequality on the second line follows by assumption, and the second by Lemma (4). The third line follows from the identity $X' = 1 - X^m$. Jointly, the above chain of inequalities implies that $M^I \leq M^C$, contradicting Theorem 2. Thus, it must be that $\hat{G}^I_d(1) = 1 - p > \int_0^1 X(rt)dt = \hat{G}^C_d(1)$. ☐

Proof of Theorem 4.
It is clear that $G_d^I(1) = G_d^C(1) = 1$. Differentiating expressions from Theorem 3 reveals that

$$\frac{d}{ds}G_d^I(s) = mp(1-(1-s)p)^{m-1}, \tag{18}$$

$$\frac{d}{ds}G_d^C(s) = m\int_0^1 (s+(1-s)X(rt))^{m-1}(1-X(rt))dt. \tag{19}$$

Evaluating each of these at $s = 1$ reveals that

$$\frac{d}{ds}G_d^I(s)\bigg|_1 = mp = m(1-\hat{G}_d^I(1)) < m(1-\hat{G}_d^C) = m(1-\int_0^1 X(rt)dt) = \frac{d}{ds}G_d^C(s)\bigg|_1,$$

where the inequality follows from Theorem 9. Thus, $G_d^I(s) > G_d^C(s)$ for all sufficiently large $s < 1$. Furthermore, Theorem 2 states that $G_d^I(0) < G_d^C(0)$; by continuity, G_d^I and G_d^C cross. Let $\hat{s} < 1$ satisfy $G_d^I(\hat{s}) = G_d^C(\hat{s})$. To prove the Theorem, it suffices to show that $\frac{d}{ds}G_d^C(\hat{s}) < \frac{d}{ds}G_d^I(\hat{s})$; that is, at any interior point where G_d^I and G_d^C cross, G_d^I is steeper.

By (5) we have that for any s,

$$\begin{aligned}(1-(1-s)p)^{m-1} - G_d^I(s) &= (1-(1-s)p)^{m-1} + (1-(1-s)p)^m \\ &= \frac{1-s}{m}\frac{d}{ds}G_d^I(s). \end{aligned}\tag{20}$$

Similarly, by (6) we have that for any s,

$$\begin{aligned}&\int_0^1 (s-(1-s)X(rt))^{m-1}dt - G_d^C(s) \\ &\qquad = \int_0^1 (s-(1-s)X(rt))^{m-1}dt - \int_0^1 (s-(1-s)X(rt))^m dt \\ &\qquad = \frac{1-s}{m}\frac{d}{ds}G_d^C(s). \end{aligned}\tag{21}$$

Examining (20) and (21), we see that to prove that $\frac{d}{ds}G_d^C(\hat{s}) < \frac{d}{ds}G_d^I(\hat{s})$, it suffices to show that

$$\int_0^1 (\hat{s}-(1-\hat{s})X(rt))^{m-1}dt < (1-(1-\hat{s})p)^{m-1}.$$

This fact follows from Lemma 4:

$$\begin{aligned}\int_0^1 (\hat{s}-(1-\hat{s})X(rt))^{m-1}dt &< \left(\int_0^1 (\hat{s}-(1-\hat{s})X(rt))^m dt\right)^{\frac{m-1}{m}} \\ &= G_d^C(\hat{s})^{\frac{m-1}{m}} = G_d^I(\hat{s})^{\frac{m-1}{m}} \\ &= (1-(1-\hat{s})p)^{m-1}.\end{aligned}$$

□

I now provide expressions giving the number of students assigned a top k choice by the Boston algorithm. Define $\hat{N}_d^I(k)$ to be the number of students who fail to match to a school ranked among their top k when using the Boston mechanism, and define $\hat{G}_d^I(k) = \lim_{n\to\infty} \frac{1}{rn}E[\hat{N}_d^I(k)]$. Let $N_d^B(s)$ be the random variable representing the number of students who receive utility at most $F^{-1}(s)$ under the Boston mechanism, and define $G_d^B(s) = \lim\limits_{n\to\infty} \frac{1}{rn}N_d^B(s)$. Theorem 10 provides a method for computing $\hat{G}_d^B$ which was used for Figure 2.

Theorem 10. *The values $\hat{G}_d^B(k)$ are given by the recursion*

$$\hat{G}_d^B(0) = 1, \qquad r(1 - \hat{G}_d^B(k+1)) = 1 - e^{-r\sum_{j=0}^{k}\hat{G}_d^B(j)}. \tag{22}$$

Furthermore, $G_d^B(s) = \sum_{k=0}^{m}\binom{m}{k}(1-s)^k s^{m-k}\hat{G}_d^B(k)$.

For the intuition behind this recursion, note that n times the left side of (22) gives the number of students matched to one of their top $k+1$ choices, while n times the right side gives the number of schools assigned to a student in the first $k+1$ rounds (since $rn\sum_{j=0}^{k}\hat{G}_d^B(j)$ is the total number of applications sent by students to this point).

I omit a formal proof of Theorem 10; I am happy to provide one upon request.

Proof of Theorem 7.

My proof follows in the footsteps of Pittel (1989), Immorlica and Mahdian (2005) and Ashlagi et al. (2013). In particular, I use the principle of deferred decisions to go from studying a deterministic algorithm (doctor-proposing deferred acceptance) on random input (agent preference lists) to studying a randomized algorithm which only makes comparisons between candidates as needed. In particular, in the analysis that follows I will take the perspective of an outside observer who witnesses the sequence of proposals and rejections, but not the set of interviews or agents' cardinal utilities. Thus, if hospital h has received N_h proposals and is holding a proposal from doctor d', then from the observer's perspective, $v_{d'h}$ is distributed as the maximum of N_h independent draws from F, and if doctor d proposes to h, then the probability that this proposal is accepted is $(N_h+1)^{-1}$.

I also use the fact for fixed preference lists, the order of proposals does not affect the outcome of the deferred acceptance algorithm. Fix r, m, n, fix $d \in D$, and let $H_d \subset H$ be the set of m hospitals with which d interviewed. Imagine running deferred acceptance among all agents excluding d, and let N_h be the number of proposals received by hospital h at the conclusion of this process. The key step in the proof is to show that in a large market, there exists a constant λ (depending on r and m, but not n) such that the vector $(N_h)_{h\in H_d}$ is well-approximated by a vector of m independent Pois(λ) random variables.

I show this by considering a run of the deferred acceptance algorithm, and tracking the set of doctors (other than d) who at some point apply to some hospital $h \in H_d$. The algorithm for doing this is given below.

Algorithm 1 (Algorithm)**.** *Fix $d \in D$, and a set $H_d \subset H$ such that $|H_d| = m$ (these represent hospitals that interviewed d).*

Define $D' = D\backslash\{d\}$ and $H' = H\backslash H_d$ (we will carefully track all interviews with hospitals in H_d)

For $d' \in D'$, initialize $H_{d'} = \emptyset$ ($H_{d'}$ represents the hospitals to which d' has so far applied).

For $h \in H'$, initialize $D_h = \emptyset$ and $\mu_h = \emptyset$ (D_h represents the set of doctors who have applied to h, and μ_h represents the doctor whose application is currently held by h).

Initialize $\tilde{D} = \emptyset$ (this represents the set of doctors who have applied to some hospital in H_d).

At all times, define $D_{matched} = \bigcup_{h\in H'}\mu_h$, and $D_{rejected} = \{d' \in D'\backslash D_{matched} : |H_{d'}| = m\}$ (these represent the set of currently matched doctors, and the set of doctors who have been rejected by m hospitals).

While$\left(\left|D_{matched} \cup D_{rejected} \cup \tilde{D}\right| < |D'|\right)$

- *Select $d' \in \{D'\backslash\left(D_{matched} \cup D_{rejected} \cup \tilde{D}\right)\}$ arbitrarily.*
- *Select h uniformly at random from $H\backslash H_{d'}$.*
- *If $h \in H_d$, add d' to $\tilde{D}$.*
- *Otherwise, add h to $H_{d'}$ and d' to D_h. With probability $1/|D_h|$, set $\mu_h = d'$.*

Call the set of proposals made during a run of the above algorithm "Phase 1" of the deferred acceptance algorithm, and call all other proposals needed to conclude the deferred acceptance procedure "Phase 2." Note that at the conclusion of Phase 1, we have a set $\tilde{D}$ of doctors other than d' who applied to (at least) one hospital in H_d, a set $\mathcal{D}_{rejected}$ of doctors who did not interview with any hospital in H_d and have been rejected by every hospital with which they interviewed, and a set $\mathcal{D}_{matched}$ of doctors who have yet to apply to any hospital in H_d and have been tentatively matched. Given $|\tilde{D}|$, let $N = (N_h)_{h \in H_d}$ be distributed according to a multinomial distribution with $|\tilde{D}|$ draws and m equally likely outcomes. Define $|\hat{D}|$ to be a Poisson random variable with mean $E[|\tilde{D}|]$. Given $|\hat{D}|$, let $\hat{N} = (\hat{N}_1, \ldots, \hat{N}_m)$ be distributed according to a multinomial distribution with $|\hat{D}|$ draws and m equally likely outcomes.

The remainder of the proof proceeds conversationally; I am happy to provide more formal notation upon request to any reviewer with concerns about how to rigorously fill in the steps outlined below. I claim the following:

I $\hat{N}$ is distributed as m independent Pois($E[|\tilde{D}|]/m$) random variables.
II $d(N, \hat{N}) = d(|\tilde{D}|, |\hat{D}|)$.
III The expected number of proposals in Phase 2 is bounded by a constant (i.e. a term that does not depend on n).
IV The probability that, during Phase 2, no $d \in D_{matched}$ applies to H_d and no $d' \in \tilde{D}$ applies to multiple hospitals in H_d, is $1 - \mathcal{O}(1/n)$.
V $|\tilde{D}|$ is approximately Poisson distributed; that is, $d(|\tilde{D}|, |\hat{D}|)$ is $o(1)$.

Claim I is a straightforward calculation. Fix non-negative integers $n_1, \ldots, n_m$, let $n = \sum_{i=1}^m n_i$. Define $|\hat{N}| = \sum_{h \in H_d} \hat{N}_h$. Then if $\lambda = E[|\tilde{D}|]/m = E[|\hat{D}|]/m$, we have

$$\begin{aligned} P(\hat{N}_1 = n_1, \ldots, \hat{N}_m = n_m) &= P(|\hat{N}| = n) P(\hat{N}_1 = n_1, \ldots, \hat{N}_m = n_m |\ |\hat{N}| = n) \\ &= \frac{e^{-m\lambda}(m\lambda)^n}{n!} \frac{n!}{\prod_{i=1}^m (n_i!)} \frac{1}{m^n} \\ &= \prod_{i=1}^m \frac{e^{-\lambda}\lambda^{n_i}}{n_i!}, \end{aligned}$$

which is the distribution of m independent Pois(λ) random variables.

To see II, couple N and $\hat{N}$ as follows:

- For all $n \in \mathbb{N}$, let $P(|N| = |\hat{N}| = n) = \min(P(|N| = n), P(|\hat{N}| = n))$.
- Whenever $|N| = |\hat{N}|$, let $N = \hat{N}$.

I now turn my attention to Claim III. Note that the number of proposals in Phase 1 is (trivially) at most rmn. Because each Phase 1 proposal is sent to a hospital in H_d (i.e. doctor d' is added to $\tilde{D}$) with probability somewhere between m/n and $m/(n-m+1)$, this implies that $|\tilde{D}|$ is stochastically dominated by a Poisson random variable with mean $rm^2 \frac{n}{n-m+1}$. Furthermore, if we define $H_{unmatched}$ to be the set of hospitals with $\mu_h = \emptyset$ at the conclusion of Phase 1, it follows that the expected size of $H_{unmatched}$ is at least $n(1 - m/\lfloor rn \rfloor)^{\lfloor rn \rfloor}$; standard techniques show that the probability that fewer than $ne^{-rm}/2$ such hospitals exist is exponentially small in n.

Each doctor in $\tilde{D}$ sends at least one application to a hospital in H_d; thus, in any run of the algorithm, they each trigger at most $m-1$ additional rejection chains. Because $\tilde{D}$ is $\mathcal{O}(rm^2)$ with high probability, the expected number of rejection chains in Phase 2 is bounded by a constant that does not depend on n. Furthermore, because rejection chain are generated using the principle of deferred decisions, each new proposal in a rejection chain goes to an unmatched hospital with probability at least the number of unmatched hospitals divided by n; with high probability this is at least $e^{-rm}/2$, and thus the expected length of each rejection chain again can be bounded by a constant that does not depend on n. This establishes claim III.

From here, IV follows easily, as the number of proposals in Phase 2 is bounded by a constant with high probability, and each proposal from $d' \in D'$ is directed to a hospital in H_d with probability at most $m/(n-m+1)$. Thus, the expected number of additional proposals to H_d (beyond the minimum of one proposal made by each doctor in $\tilde{D}$) from doctors in D' is $\mathcal{O}(1/n)$, and therefore the probability that any such proposal occurs is $\mathcal{O}(1/n)$. In other words, with high probability, no rejection chain returns to H_d, and the number of proposals that each $h \in H_d$ receives from doctors $d' \in D'$ is equal to N_h with high probability.

To see V, note that each time through the main loop of the algorithm, d' is added to $\tilde{D}$ with probability somewhere between m/n and $m/(n-m+1)$. Furthermore, by Lemma 3, P_n (the total number of proposals across Phases 1 and 2) is concentrated about its mean. Because $\lfloor rn \rfloor \le P_n \le m\lfloor rn \rfloor$, and because III states that the expected number of proposals in Phase 2 is bounded by a constant, it follows that the number of proposals in Phase 1 is concentrated around $E[P_n]$ (formally, the ratio of these two converges in probability to one as n grows). Thus, with high probability, Phase 1 consists of $E[P_n](1+o(1))$ proposals, each of which is directed at a hospital in H_d with probability between m/n and $m/(n-m+1)$. Standard results stating that Binom$(n, \lambda/n)$ converges in distribution to Pois(λ) then imply that $|\tilde{D}|$ is approximately Poisson.

All that remains is to solve for λ, the large-market average number of applications received by each hospital. The above analysis implies that d is rejected from each hospital with probability approaching $p = E[(1+N_h)^{-1}] = g(\lambda)$. Thus, the probability that d matches approaches $1-(1-p)^m$, and the expected number of applications sent by d approaches $(1-(1-p)^m)/p$ as n grows. By symmetry among doctors, we must have that

$$\frac{1}{rn}E[P_n] \to (1-(1-p)^m)/p.$$

By symmetry among hospitals we must have that

$$\lambda = \lim_{n\to\infty} \frac{1}{n}E[P_n] = r(1-(1-p)^m)/p.$$

Putting it all together, we see that p and λ must solve

$$p = g(\lambda), \qquad \lambda = r(1-(1-p)^m)/p,$$

which is equivalent to (9). □

Proof of Theorem 8.

Note that when hospital preferences are correlated, there is a unique stable match. We can construct it through a doctor serial dictatorship: order the doctors according to q_d, and have them sequentially pick their favorite remaining hospital (among those with whom they interviewed).

Given n and m, define the random sequences Z_n, B_n as follows:

$Z_n(0) = 0$, and for $k = 1, 2, \ldots$, let

$$B_n(k) = \begin{cases} 0 & w.p. \ \binom{Z_n(k-1)}{m}/\binom{n}{m} \\ 1 & \text{otherwise} \end{cases}, \qquad Z_n(k) = \sum_{j=1}^{k} B_n(j).$$

Furthermore, let $B_n(k)$ be conditionally independent of $\{B_n(j)\}_{j=1}^{k-1}$, given $Z_n(k-1)$.[22] Think of $B_n(k)$ as the indicator that the k^{th}-ranked doctor matches; then $Z_n(k)$ is a count of the number of hospitals that are filled after doctor k has selected. Given $Z_n(k-1)$, the k^{th}-ranked doctor fails to match if and only if all of their interviews were with hospitals who have already matched. Because interview scheduling is independent from doctor preferences, this occurs with probability $\binom{Z_n(k-1)}{m}/\binom{n}{m}$. Thus, we have that $\left|\mu_n^C\right| \stackrel{D}{=} Z_n(\lfloor rn \rfloor)$. I will show that for all $r > 0$,

$$\frac{1}{n}E[Z_n(\lfloor rn \rfloor)] \to X(r). \tag{23}$$

[22]Unless specifically stated otherwise, when defining a random variable I intend that this variable is independent of other random variables defined on the same probability space.

The first step is to replace the unwieldy combinatorial term above with the simpler $(Z_n(k-1)/n)^m$. To justify this, note that

$$\left(\frac{z-m+1}{n}\right)_+^m \leq \binom{z}{m} \Big/ \binom{n}{m} \leq \left(\frac{z}{n}\right)^m. \tag{24}$$

Define $\overline{Z}_n, \overline{B}_n$ by $\overline{Z}_n(0) = 0$, and for $k = 1, 2, \ldots$, let

$$\overline{B}_n(k) = \begin{cases} 0 & w.p. \ \left(\frac{Z_n(k-1)-m+1}{n}\right)_+^m \\ 1 & \text{otherwise} \end{cases}, \qquad \overline{Z}_n(k) = \sum_{j=1}^{k} \overline{B}_n(j).$$

Similarly, define $\underline{Z}_n, \underline{B}_n$ by $\underline{Z}_n(0) = 0$, and for $k = 1, 2, \ldots$, let

$$\underline{B}_n(k) = \begin{cases} 0 & w.p. \ \left(\frac{Z_n(k-1)}{n}\right)^m \\ 1 & \text{otherwise} \end{cases}, \qquad \underline{Z}_n(k) = \sum_{j=1}^{k} \underline{B}_n(j).$$

We can think of $\underline{Z}_n^k$ as representing the transition dynamics when doctors are "forgetful" and may accidentally schedule multiple interviews with the same hospital.

Note that we may couple $\underline{Z}_n$ and $\overline{Z}_n$ such that for $k \geq m$, $\overline{B}_n(k) = \underline{B}_n(k-m+1)$. This implies that

$$\overline{Z}_n(k) \stackrel{D}{=} m - 1 + \underline{Z}_n(k-(m-1)) \leq m - 1 + \underline{Z}_n(k) \tag{25}$$

Further, (24) implies that we can couple $\underline{Z}_n$ and $\overline{Z}_n$ with Z_n such that $\underline{Z}_n(k) = Z_n(k)$ implies $\underline{B}_n(k+1) \leq B_n(k+1)$, and $Z_n(k) = \overline{Z}_n(k)$ implies $B_n(k+1) \leq \underline{B}_n(k+1)$. It follows that we may couple $\underline{Z}_n$ and Z_n such that for all n and $k \geq m$,

$$\underline{Z}_n(k) \leq Z_n(k) \leq m - 1 + \underline{Z}_n(k).$$

Thus, in order to show (23), it suffices to show that $\frac{1}{n}E[\underline{Z}_n(\lfloor rn \rfloor)] \to X(r)$.

For $k \in \{1, \ldots, n\}$, define $Y_n(k)$ to be a geometric random variable with mean $\left(1 - \left(\frac{k-1}{n}\right)^m\right)^{-1}$. Define $N_n(k) = \sum_{j=1}^{k} Y_n(j)$. Then $N_n(k)$ represents the number of (forgetful) doctors that are needed in order for k of n hospitals to fill positions. Note that we can couple $\underline{Z}_n$ and N_n in the natural way, in which case we have that for all $k, z \in \mathbb{N}$,

$$\underline{Z}_n(k) < z \Leftrightarrow k < N_n(z). \tag{26}$$

Note that for $s \in [0, 1)$, as $n \to \infty$, we have

$$\frac{1}{n}E[N_n(\lfloor sn \rfloor)] = \frac{1}{n}\sum_{k=1}^{\lfloor sn \rfloor}\left(1 - \left(\frac{k-1}{n}\right)^m\right)^{-1} \to \int_0^s \frac{1}{1-t^m}dt \triangleq M(s). \tag{27}$$

$$\frac{1}{n}\mathrm{Var}[N_n(\lfloor sn \rfloor)] = \frac{1}{n}\sum_{k=1}^{\lfloor sn \rfloor}\left(\frac{k-1}{n}\right)^m\left(1 - \left(\frac{k-1}{n}\right)^m\right)^{-2} \to \int_0^s \frac{t^m}{(1-t^m)^2}dt \triangleq V(s). \tag{28}$$

From (27) and (28), it follows that for any $s \in (0, 1), \epsilon > 0$, as $n \to \infty$ we have

$$P\left(\frac{1}{n}N_n(\lfloor sn \rfloor) \in (M(s) - \epsilon, M(s) + \epsilon)\right) \to 1. \tag{29}$$

Additionally, we claim that $M(X(r)) = r$. To see this, note that $M(X(0)) = M(0) = 0$, and

$$\frac{d}{dr}M(X(r)) = \frac{d}{dr}\int_0^{X(r)} \frac{1}{1-t^m}dt = \frac{X'(r)}{1-X(r)^m} = 1.$$

Applying (26), we see that for any $r, \epsilon > 0$,

$$P\left(\lfloor n(X(r)-\epsilon)\rfloor < \underline{Z}_n(\lfloor rn \rfloor) < \lfloor n(X(r)+\epsilon)\rfloor\right) = P(N_n(\lfloor n(X(r)-\epsilon)\rfloor)) < \lfloor rn \rfloor < N_n(\lfloor n(X(r)+\epsilon)\rfloor)) \to 1,$$

with the final claim following from applying (29) twice. This is the definition of what it means for $\frac{1}{n}E[Z_n(\lfloor rn \rfloor)]$ to converge to $X(r)$ in distribution.

Now consider a given doctor d. If d ranks j^{th} among doctors, then the probability that d is rejected by his first k hospitals is $E[\binom{Z_n(j-1)}{k}]/\binom{n}{k}$; It follows that the ex-ante probability that d is rejected by his first k hospitals is

$$P(R_n^C \geq k) = \frac{1}{\lfloor rn \rfloor}\sum_{j=1}^{\lfloor rn \rfloor} E\left[\binom{Z_n(j-1)}{k}\right] / \binom{n}{k}.$$

By (24), it follows that

$$\frac{1}{\lfloor rn \rfloor}\sum_{j=1}^{\lfloor rn \rfloor} E\left[\left(\frac{Z_n(j-1)-m+1}{n}\right)^k\right] \leq P(R_n^C \geq k) \leq \frac{1}{\lfloor rn \rfloor}\sum_{j=1}^{\lfloor rn \rfloor} E\left[\left(\frac{Z_n(j-1)}{n}\right)^k\right].$$

By (23), as $n \to \infty$, we have that

$$\frac{1}{\lfloor rn \rfloor}\sum_{j=1}^{\lfloor rn \rfloor} E\left[\left(\frac{Z_n(j-1)}{n}\right)^k\right] \to \int_0^1 X(rt)^k dt,$$

proving the claim that $R_n^C \to R^C$. □

8.3. **Proofs from Section 6.** I begin by providing expressions for the functions G_h, G_d that result when doctors have an outside option $\overline{u}$ satisfying $F(\overline{u}) = \alpha$, and hospitals have an outside option $\overline{v}$ satisfying $F(\overline{v}) = \beta$. While studying this is interesting in its own right, this also forms an essential component of the analysis for computing an equilibrium when there is an aftermarket. Note that I define $G_d(s)$ to be the probability of getting a match worse than $F^{-1}(s)$ (or no match at all) *from the clearinghouse*, assuming that d lists all interview partners (i.e. G_d does not explicitly incorporate d's own outside option).

I provide the following expressions:

(1) Independent preferences: Let $x \in (0,1)$ be the unique solution (see Lemma 1) to

$$x = r\left(1 - \left(1 + \frac{(1-\alpha)(1-\beta)x}{\log(1-x)}\right)^m\right),$$

and define p by $r(1-(1-(1-\alpha)p)^m) = x$, and $\lambda = r(1-(1-(1-\alpha)p)^m)/p$. Then

$$G_d^I(s) = (1-(1-s)p)^m, \qquad G_h^I(t) = e^{-\lambda(1-t)}, \tag{30}$$

(2) Correlated preferences: Define $X(t)$ by the differential equation

$$X'(t) = 1 - (\alpha + (1-\alpha)X(t))^m, X(0) = 0.$$

Then we have that

$$G_d^C(s) = 1 - \int_0^{1-\beta}(s+(1-s)X(rt))^m dt, \qquad G_h^C(t) = \begin{cases} 1 - X(r(1-t)) & : t \geq \beta \\ (1-X(r(1-\beta)))e^{-rm(\beta-t)} & : t < \beta \end{cases}. \tag{31}$$

(3) Aligned preferences: Define $\tilde{G}$ to be the solution to

$$\tilde{G}'(t) = rm\tilde{G}(t)\left(1 - \int_t^1 \tilde{G}(u)du\right)^{m-1}, \qquad \tilde{G}(1) = 1.$$

Then all matches formed satisfy $F^{-1}(u_{dh}) \geq \max(\alpha, \beta)$; for $s, t \geq \max(\alpha, \beta)$, we have

$$G^A_d(s) = 1 - \frac{1 - \tilde{G}(s)}{r}, \qquad G^A_h(t) = \tilde{G}(t). \tag{32}$$

For the intrepid reviewer who has reached this stage of the appendix, I note that the proof of Theorem 6, while complete and accurate, is not fully formal: it uses somewhat imprecise English words rather than a formal definition of equilibrium. I trust that this will cause no confusion, however I do provide a formal equilibrium definition below.

Proof of Theorem 6. When $r = 1$, all agents who go unmatched in the clearinghouse will be matched in the aftermarket, so agents have a dominant strategy of listing only those partners whom they prefer to a random match.

For the remainder of the proof, let x_1 be the fraction of hospitals who match when both sides list only excellent matches, and let $x_2$be the fraction of hospitals who match when hospitals list only excellent matches, but doctors list each hospital with whom they interviewed. I make the following observation: [23]:

Observation 1. *a) For any r, m, γ, $0 < x_1 < x_2 < 1$.*
b) For fixed r, γ, the values x_1 and x_2 are increasing in m.
c) For fixed m, γ, the values x_1 and x_2 are increasing in r.

When $r > 1$, all hospitals who go unmatched in the clearinghouse will be matched in the aftermarket, so hospitals have a dominant strategy of listing only doctors whom they prefer to a random match. Suppose that doctors anticipate that a fraction x of hospitals will match through the clearinghouse. This implies that a fraction $\frac{1-x}{r-x}$ of unmatched doctors will find a match in the aftermarket. It follows that when deciding whether to list a marginal match, doctors compare the value of this match, s, to the value of the aftermarket, $\frac{1-x}{r-x}(\gamma + (1-\gamma)s)$. After rearrangement, we see that doctors prefer to list only excellent matches if and only if

$$s < \left(1 + \frac{1}{\gamma} \cdot \frac{r-1}{1-x}\right)^{-1}. \tag{33}$$

If we define $\overline{s}$ to be the value of the right side of (33) when $x = x_1$ and $\underline{s}$ to be its value when $x = x_2$, it follows from Observation 1 that:

(1) $0 < \underline{s} < \overline{s} < 1$;
(2) It is an equilibrium for doctors to list only excellent matches if and only if $s \leq \overline{s}$;
(3) It is an equilibrium for doctors to list all interview parteners if and only if $s \geq \underline{s}$;
(4) For fixed r, γ, as m increases, so do x_1, x_2, and thus $\underline{s}$ and $\overline{s}$ are decreasing in m;
(5) For fixed m, γ, as r increases, so do $\frac{r-1}{1-x_1}$, $\frac{r-1}{1-x_2}$, and thus $\underline{s}, \overline{s}$ are decreasing in r.

The case where $r < 1$ is exactly analogous, with the roles reversed. □

Formal Equilibrium Definition

Here, I formally define the equilibrium concept intended above. In English, an equilibrium is a pair of values $\overline{u}, \overline{v}$ such that when doctors list only hospitals with $u_{dh} > \overline{u}$ and hospitals list only doctors for which $v_{dh} > \underline{v}$, the "value of the aftermarket" to the two sides is precisely $\overline{u}, \overline{v}$. When preferences are independent or aligned, the "value of the aftermarket" is simply the probability of matching times the expected value from a random match.

For any α, β, define $M^I(\alpha, \beta)$ to be $1 - G^I_h(\beta)$, where G^I_d is computed according to (30); similarly, define $M^A(\alpha, \beta)$ to be $1 - G^A_h(\beta)$, where G^A_h is given by (32). Note that M^I (respectively, M^A) represents the fraction of hospitals who match when preferences are independent (respectively,

[23] These facts follow almost trivially from the model, but for completeness, note that I provides a procedure for computing x_1 and x_2 above.

aligned) and doctors reject an α fraction of all hospitals, while hospitals reject a β fraction of all doctors.

Given that the fraction of hospitals who match through the clearinghouse is M, the fraction of doctors, among those participating in the aftermarket, who go on to match is 1 if $r \leq 1$, and is $\frac{1-M}{r-M}$ if $r > 1$. Analogously, the fraction of hospitals in the aftermarket who go on to find a match is 1 if $r \geq 1$, and is $\frac{r-M}{1-M}$ if $r < 1$. Let $\hat{u} = E[u_{dh}]$ be the expected utility of a random match. Then the "value of the aftermarket" for given cutoff strategies $\overline{u}, \overline{v}$ is given by

$$V^I(\overline{u},\overline{v}) = \left(\min\left(\frac{1-M^I(F(\overline{u}),F(\overline{v}))}{r-M^I(F(\overline{u}),F(\overline{v}))},1\right)\hat{u}, \min\left(\frac{r-M^I(F(\overline{u}),F(\overline{v}))}{1-M^I(F(\overline{u}),F(\overline{v}))},1\right)\hat{u}\right).$$

$$V^A(\overline{u},\overline{v}) = \left(\min\left(\frac{1-M^A(F(\overline{u}),F(\overline{v}))}{r-M^A(F(\overline{u}),F(\overline{v}))},1\right)\hat{u}, \min\left(\frac{r-M^A(F(\overline{u}),F(\overline{v}))}{1-M^A(F(\overline{u}),F(\overline{v}))},1\right)\hat{u}\right).$$

Finally, define the pair $(\overline{u},\overline{v})$ to be an equilibrium of the game with independent preferences if $(\overline{u},\overline{v}) = V^I(\overline{u},\overline{v})$, and to be an equilibrium of the game with aligned preferences if $(\overline{u},\overline{v}) = V^A(\overline{u},\overline{v})$. Note that this equilibrium notion inherently assumes a large market: for small n, a hospital who expects value $\overline{v}$ from the aftermarket may have an incentive to truncate their list to include only partners of value $\overline{v}+\epsilon$. Furthermore, for small n, agents might account for the possibility that the aftermarket matches them to a partner whom they interviewed and declined to list.

9. Appendix: Extensions

9.1. **Many-to-One Matching.** Although the model and analysis in the body of the paper assumes a one-to-one matching market, in many real-world settings (including residency matching and school assignment), agents one side of the market have the capacity to match to multiple agents on the opposite side. In this appendix, I provide expressions for the number of matches that form in the general case where there are n hospitals, each with capacity $C \geq 1$, and $\lfloor rnC \rfloor$ doctors seeking a single match partner.

9.1.1. *Both Sides Have Independent Preferences.* In the independent case, the consistency equations defining the probability p with which each offer is accepted and the expected number of offers received by a program λ are:

$$\lambda p = rC(1-(1-p)^m) = E[\min(\text{Pois}(\lambda), C)].$$

9.1.2. *Hospitals Have Correlated Preferences.* Suppose that when doctor d picks, a fraction X of hospitals are filled. Then the expected number of applications sent by doctor d in a large market is $\frac{1-X^m}{1-X}$. Meanwhile, because each application is essentially sent uniformly at random, after a total of λn applications have been sent, the number of schools that have filled all positions is given by $X = P(Pois(\lambda) \geq C)$. This gives us the following differential equation:

$$\Lambda'(t) = \frac{1-X(t)^m}{1-X(t)}, X(t) = 1 - \sum_{k=0}^{C-1} \frac{e^{-\Lambda(t)}\Lambda(t)^k}{k!}, \Lambda(0) = 0.$$

Then $\Lambda(rC)$ gives the average number of applications received by each hospital, and the total number of doctors who match is given by $nE[\min(\Lambda(rC), C)]$.

9.1.3. *Comparing M^I, M^C, M^A.* The intuition for the result $M^I \geq M^C$ remains intact: the primary factor determining the number of matches that form is the likelihood that doctors applying to hospitals towards the bottom of their list will be accepted. This probability is higher under independent preferences than under correlated preferences, as in the latter case doctors who apply to many hospitals have been adversely selected. I do not currently have a formal proof establishing this fact.

By contrast, the coupling argument used to prove that $M^C \geq M^A$ carries through virtually unmodified; it is still the case that the greedy algorithm corresponding to aligned preferences is likely to match those doctors with many remaining options. However, note that unlike the case where $C = 1$, correlation in doctor preferences affects the number of matches that form. In particular, when doctor preferences are correlated, highly-ranked doctors tend to all pick highly-ranked hospitals, causing these hospitals to fill early and decreasing the total number of doctors who match.

While the coupling argument establishes that fewer doctors match when preferences are aligned than when doctor preferences are idiosyncratic and hospital preferences are perfectly correlated, it says no such thing about the case where both doctors and hospitals have correlated preferences. Indeed, simulations demonstrate that these two cases are generally incomparable: for some values of r, m, C, more matches form under aligned preferences, and for other parameter values, more matches form when both sides have perfectly correlated preferences.

9.2. **Varying Number of Applications.** What if, rather than each doctor having a hard capacity constraint, doctors apply to each hospital independently with probability m/n (so that they average number of hospitals with which a doctor interviews is m)? Note that in this case, it is without loss of generality to assume $r \geq 1$.

These results, which study the size of maximal matchings selected by three different procedures in bipartite Erdös-Renyi random graphs, are of independent interest (outside of the context of centralized matching considered in this paper). In addition, the closed-form expressions in the case of correlated and aligned preferences make it simple to compute the scaling behavior of the number of unmatched agents on the short side as m grows.

Theorem 11.

(1) Independent preferences:

$$\lim_{n\to\infty} \frac{1}{n} S_n(r, m, G_I) = x^*,$$

where x^* is the unique solution to

$$m = \frac{1}{x} \log(1 - x) \log(1 - x/r). \tag{34}$$

(2) Perfectly correlated preferences:

$$\lim_{n\to\infty} \frac{1}{n} S_n(r, m, G_C) = 1 - \frac{1}{m} \log(1 + (e^m - 1)e^{-mr}). \tag{35}$$

(3) Aligned preferences:

$$\text{For } r > 1, \lim_{n\to\infty} \frac{1}{n} S_n(r, m, G_A) = \frac{e^{rm} - e^m}{e^{rm} - e^m/r} \tag{36}$$

$$\text{For } r = 1, \lim_{n\to\infty} \frac{1}{n} S_n(r, m, G_A) = \frac{m}{m + 1} \tag{37}$$

The derivations in the case of independent and correlated preferences are similar to those for Theorem 1. For independent preferences, it remains the case that each hospital is available with probability p. The consistency equation for p becomes

$$1 - e^{-\frac{1}{p} r(1 - e^{-mp})} = r(1 - e^{-mp}),$$

which can be transformed via the change of variables $x = r(1 - e^{-mp})$ into (34). For a proof that there is a unique solution for all r, m, see Arnosti et al. (2014).

For correlated preferences, a doctor who selects at a time when a fraction X of hospitals have been filled matches with probability $1 - e^{-m(1-X)}$. It follows that the relevant differential equation is

$$X'(t) = 1 - e^{-m(1-X(t))}, \qquad X(0) = 0,$$

which has the closed-form solution given in (35). Computing G_d, G_h for both independent and correlated preferences also becomes straightforward. The expression in (35) was previously derived by Mastin and Jaillet (2013).

To the best of my knowledge, the expressions in (36) and (37) are novel. They come from considering all possible rn^2 edges, and ordering them according to the value u_{dh}. Let X track the fraction of hospitals who have matched so far. When considering an edge, it is added to the matching if and only if the corresponding interview occurred (probability m/n), the hospital has yet to match (probability $\approx 1 - X$), and the doctor has yet to match (probability $\approx 1 - X/r$). The interview structure implies independence of these events, implying that the fraction of hospitals who eventually match should be $X(r)$, where X is given by the differential equation

$$X'(t) = m(1 - X(t))(1 - X(t)/r), \qquad X(0) = 0,$$

which has the closed-form solution given in (36) and (37).

The above expressions provide over-estimates for the number of agents that remain unmatched in the case that each doctor schedules exactly m interviews, however for moderate to large values of m, these over-estimates should remain fairly tight.

Of course, as m grows, the fraction of hospitals who remain unmatched converges to zero. It is, however, illustrative to consider the rate at which the convergence occurs. When $r = 1$, we have that

$$\begin{aligned} 1 - S^I &= e^{-\sqrt{mS^I}} \sim e^{-\sqrt{m}} \\ 1 - S^C &= \log(2 - e^{-m})/m \sim \frac{\log(2)}{m} \\ 1 - S^A &= \frac{1}{m+1}. \end{aligned}$$

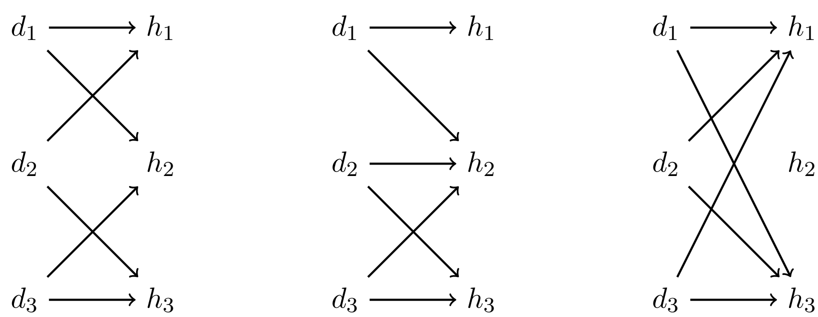
If $r > 1$, we have that

$$\begin{aligned} 1 - S^I &\sim e^{-m/\log(r/(r-1))} \\ 1 - S^C &\sim \frac{1}{m} e^{-m(r-1)} \\ 1 - S^A &\sim \frac{r-1}{r} e^{-m(r-1)}. \end{aligned}$$

Note that in this case all of the above expressions decrease exponentially in m, but the coefficient in the exponent for independent preferences is greater than for aligned or correlated preferences. For example, for $r = 1.1$, $1/\log(r/(r-1)) \approx 0.417$, whereas $r - 1 = 0.1$.

10. Appendix: An Example

Example 2. *There are three doctors and three hospitals, and $m = 2$. Up to isomorphism, there are three possible application graphs, depicted below:*



Intuition for independent better than perfectly correlated: look at case one, and suppose that doctors are ranked, with d_3 ranked last. Then d_3 goes unmatched only if h_1 is not selected by either of the top doctors; this occurs with probability $1/3$ when doctors have identical rankings, but with probability $1/4$ when doctors rank hospitals independently. Moving to the case where both sides have independent preferences, we observe the same effect, as the chance that both h_2 and h_3 decide that d_3 is their least preferred option drops from $1/3$ to $1/4$.

Intuition for the fact that aligned is the worst: look at case two. Note that the only hospital at risk of going unmatched is h_1; this hospital goes unmatched if and only if d_1 and h_2 match. Both d_1 and h_2 have outside options, and thus must both prefer each other to these options. This is most likely when preferences are aligned.

Interesting to note that in case two, "both sides perfectly correlated" is actually better than "u_{dh} independent, v_{dh} perfectly correlated." This is because if h_2 is not picked by the top-ranked doctor, under perfectly correlated preferences, h_2 is less likely to be selected by d_1.

Let us consider the first case, in which each hospital receives two applications. This case occurs with probability $2/9$.

- *Suppose that hospital preferences over doctors are identical, as are doctor preferences over hospitals. WLOG, let d_1 be the most preferred doctor and d_3 the least preferred. The only doctor that might go unmatched is d_3; this occurs if and only if h_1 is the least preferred hospital, which occurs with probability $1/3$.*
- *Suppose that hospital preferences over doctors are identical, but doctor preferences over hospitals are independent. WLOG, let d_1 be the most preferred doctor and d_3 the least preferred. The only doctor that might go unmatched is d_3; this occurs if and only if d_1 prefers h_2 to h_1 and d_2 prefers h_3 to h_1, which occurs with probability $1/4$.*
- *The case where hospital preferences over doctors are independent, but doctor preferences are perfectly correlated is identical to the previous case, by symmetry.*
- *Suppose that both sides have independent preferences. In this case, d_3 goes unmatched if and only if: h_2 prefers d_1 to d_3, h_3 prefers d_2 to d_3, d_1 prefers h_2 to h_1, and d_2 prefers h_3 to h_1. This occurs with probability $1/16$; by symmetry, the probability that only two matches form is $3/16$.*
- *If preferences are aligned, then WLOG the link with the highest value is d_1h_1. After matching these two, only d_2 can go unmatched; this occurs if and only if d_3h_3 is the highest of the three remaining links, which occurs with probability $1/3$.*

Now let us turn to the second case, in which h_1 receives one application, and h_2 receives three. This case occurs with probability $2/3$. Note that in this case, the only hospital at risk of going unmatched is h_1.

- *Suppose that hospital preferences over doctors are identical, as are doctor preferences over hospitals. Then h_1 goes unmatched if and only if one of the following holds:*
  (1) d_1 is ranked first, and prefers h_2 to h_1 (occurs with probability $1/6$), or
  (2) d_2 is ranked second, and h_3 is ranked first, followed by h_2 (occurs with probability $1/18$).

Alternatively, we can say that h_1 goes unmatched if and only if one of the following holds:

(1) h_2 is ranked first, and prefers d_1 to d_2 and d_3 (occurs with probability 1/9), or
(2) h_3 is ranked first, h_2 is ranked second, and d_1 is ranked first or second (occurs with probability 1/9).

- *Suppose that hospital preferences over doctors are identical, but doctor preferences over hospitals are independent. Then h_1 goes unmatched if and only if one of the following holds:*
  (1) d_1 is ranked first, and prefers h_2 to h_1 (occurs with probability 1/6), or
  (2) d_1 is ranked second, prefers h_2 to h_1, and the top-ranked doctor prefers h_3 to h_2 (occurs with probability 1/12).
- *Suppose that hospital preferences over doctors are independent, but doctor preferences over hospitals are identical. Then h_1 goes unmatched if and only if one of the following holds:*
  (1) h_2 is ranked first, and prefers d_1 to d_2 and d_3 (occurs with probability 1/9), or
  (2) h_3 is ranked first, h_2 is ranked second, and h_2 prefers d_1 to the doctor not selected by h_3 (occurs with probability 1/12).
- *Suppose that doctors and hospitals hold independent preferences. Then h_1 goes unmatched if and only if d_1 prefers h_2 to h_1 (probability 1/2), and one of the following holds:*
  (1) h_2 prefers d_1 to d_2 and d_3 (probability 1/3), or
  (2) h_2 ranks d_1 second, the doctor that h_2 ranks first prefers h_3, and h_3 prefers this doctor to the other doctor that they interviewed (probability 1/12).
- *Suppose that preferences are aligned. Then h_1 goes unmatched if and only if one of the following holds:*
  (1) d_1h_2 is the top link (occurs with probability 1/6), or
  (2) One of d_2h_3 or d_3h_3 is the top link, and d_1h_2 is the top link of the three remaining (occurs with probability 1/9).

The third case, in which one hospital receives no applications, occurs with probability 1/9 and is trivial to analyze: for all preference structures, exactly two matches form.

Probability of a Vacancy

Doctors Send 2 Apps Uniformly at Random		Hospital Preferences	
		Independent	Identical
Doctor Preferences	Independent	7/24	1/3
	Identical	8/27	1/3

Doctors Send 2 Apps Hospitals Receive 2 Apps		Hospital Preferences	
		Independent	Identical
Doctor Preferences	Independent	3/16	1/4
	Identical	1/4	1/3