

Determinants of core size in matching markets: going beyond short rank lists

By TYLER HOPPENFELD

I examine two-sided matching in large markets with a novel preference structure that allows me to understand the effects of asymmetry in preference formation. I demonstrate conditions in which the core of a matching market remains large when the market has a large number of participants, and show a sharp effect of market asymmetry on core size and the relative rank with which participants are matched.

I. Introduction

Centralized matching markets encompass only a small fraction of the labor market, however the study of centralized matching markets sheds light on the broader matching of employees to jobs. There are also a variety of markets, including academic job markets and some entry-level job markets for new college graduates, that are decentralized and sometimes chaotic despite having synchronized start dates for a large number of applicants to very similar jobs. A better understanding of the strengths and limits of centralized markets could help us understand what inefficiencies exist in these markets, whether more efficient solutions are possible, and even allow for the design of more efficient markets.

Theory tells us that there is no matching mechanism that reliably produces a stable outcome wherein both sides have the incentive to accurately report their preferences (Roth, 1985), however in practice many matching markets persist with both sides (apparently) reporting their preferences honestly (Roth and Peranson, 1991). This result can be explained by the fact that in many of these markets the set of stable matchings is quite small, and under reasonable assumptions only agents with multiple potential stable matches have an incentive to lie about their preferences. For example, in the National Resident Matching Program (NRMP) matching of recently graduated physicians to resident positions, all but approximately 10 of 20,000 participants have a unique partner in the set of stable matchings (Roth and Peranson, 1999), and so in all but a very small number of cases any misrepresentation of preferences would lead to an identical or less-preferred outcome for the agent misrepresenting their preferences.

Roth and Peranson posit that the small set of stable matchings (SSM) of these markets is driven by the correlated nature of preferences, and by the fact that for highly correlated preferences, the SSM is very small. In a parallel argument, they note that finite preference lists, as is the case with the NRMP, lead to a very small SSM as well. They use simulation to illustrate their claims with randomly drawn short preference lists, and (Immorlica and Mahdian (2005) and then Kojima and

Pathak (2009) build on this with the theoretical result that as the number of market participants grows, the SSM becomes arbitrarily small as a fraction of market size. This theoretical work takes short preference lists to be a feature of preferences rather than a feature of the market mechanism, and uses a highly restrictive preference structure to accomplish that.

In a related strain of work Ashlagi, Kanoria and Leshno (2017) look at matching outcomes in small unbalanced two-sided matching markets with uniform random preferences. They find that the market outcome sharply depends on the imbalance, and the core is large only if the market is balanced, with the same number of participants on each side.

This paper builds a bridge between these two strains of work, in that it relaxes the restrictive preferences found in Immorlica and Mahdian (2005) and Kojima and Pathak (2009), and finds that, like in Ashlagi, Kanoria and Leshno (2017), a large core depends on a symmetrical two-sided market. In prior literature where hospitals form k length preference lists over n doctors and preference lists of length k , for any ϵ_a , no more than k/ϵ_a doctors can be acceptable to $n\epsilon_a$ hospitals. That is to say that as the market grows, nearly all doctors are unacceptable to nearly all hospitals. (See appendix A.A2 for more). In this paper I examine short rank lists in a setting that allows for more realistic preference formation. I construct preferences where all participants on each side in a matching market broadly share preferences over their potential matches, requiring short rank-lists does not introduce strategic behavior, and the fraction of agents with multiple stable matches has a positive limit as the market becomes infinitely large.

II. Model

A. Preliminary Definitions

A two sided matching market consists of a set of doctors D and hospitals H , each of which might hire at most one doctor. Doctors and hospitals together form the universe of participants $U = D \cup H$. Each doctor d has a strict preference relation over the set of hospitals he could be matched with and being unmatched, which we denote as being matched with oneself. Therefore we say that the doctor has preferences \succ_d over the set $H \cup \{d\}$. Each hospital h likewise has preferences \succ_h over $D \cup \{h\}$.

A matching ϕ is a mapping from U onto itself such that

- for every d , $|\phi(d)| = 1$, $\phi(d) \in H \cup \{d\}$ and for every h , $|\phi(h)| = 1$, $\phi(h) \in D \cup \{h\}$
- $\phi(d) = h$ iff $\phi(h) = d$. That is, a matching is simply a one-to-one assignment of doctors to hospitals with the possibility of being unmatched. We also denote the set of all participants not matched to themselves in ϕ as $\{\phi\}$.

We say a matching ϕ is blocked by a pair $\{d, h\}$ if $h \succ_d \phi(h)$ and $d \succ_h \phi(h)$. It

is individually rational if $\forall a \in U, \phi(a) \succ_a a | \phi(a) = a$. A matching is stable if it is individually rational and not blocked.

THEOREM 1: *In a 1:1 matching, the core and set of stable matchings are equivalent.*

PROOF:

An unstable matching contains at least one couple who can defect to their mutual benefit, or an individual who can defect to become unmatched, and so is not part of the core.

The defecting coalition of a non-core assignment consists of individuals for whom the matching is not individually rational and blocking pairs, so any non-core outcome contains is not stable.

B. Market Participants

There are n doctors indexed $\{d_1, d_2, \dots, d_n\}$ and n hospitals indexed $\{h_1, h_2, \dots, h_n\}$. The degree of preference variation among doctors is indexed by the variable k_d , while the degree of preference variation among hospitals is indexed by k_h .

Each doctor forms their preference list by first ranking all hospitals according to their index number, so h_1 is the doctor's favorite and h_n their least favorite (the doctor's least favorite option is to be unmatched). Doctor each doctor d_a randomizes their preferences over hospitals $\{h_{a-k_d}, \dots, h_{a+k_d}\}$, leaving their other preferences unchanged. Doctors $d_a, a \in [1, k_d + 1]$ randomize preferences over hospitals $\{h_1, \dots, h_{a+k_d}\}$, and doctors $d_a, a \in [n - k_d, n]$ randomize preferences over hospitals $\{h_{a-k_d}, \dots, h_n\}$. The hospitals form their preferences in the same way, using k_h in place of k_d .

In this way, we arrive at a preference structure where agents all broadly agree on which matches are desirable, but have idiosyncratic preference variation.

III. Analysis

A. Rank List Lengths

LEMMA 1: *For any pair in a stable matching (d_a, h_b) , $|a - b| < \max(k_d, k_h) + 2$*

PROOF:

Call $k = \max(k_d, k_h)$. First consider the match $(d_1, h_a), a > k + 1$. (d_1, h_{k+1}) form a blocking pair, because $h_{k+1} \succ_{d_1} h_{k+2} \succ_{d_1} h_{k+3} \succ_{d_1} h_{k+4} \dots$, and d_1 is the first choice of h_{k+1} . Thus $(d_1, h_a), a > k + 1$ are not matched in a stable matching.

Likewise, $(d_a, h_1), a > k + 1$ are not matched in a stable matching.

Now consider $(d_2, h_a), a > k + 2$. (d_2, h_{k+2}) form a blocking pair, because $h_{k+2} \succ_{d_2} h_{k+3} \succ_{d_2} h_{k+4} \succ_{d_2} h_{k+5} \dots$, and since h_{k+2} is not matched to d_1 , h_{k+2} prefers d_2 to their current match. Thus $(d_2, h_a), a > k + 2$ are not matched in a stable matching.

Likewise, $(d_a, h_2), a > k + 2$ are not matched in a stable matching. continuing by induction, we see that no pair $(d_a, h_b), |a - b| > k + 1$ is included in a stable matching.

IV. Computational Experiments

This section presents simulation results that complement my theoretical results and provide evidence for a key theoretical conjecture. in a variety of market sizes size markets demonstrating:

A. Numerical Results for Block Preferences

My first computational experiment illustrates the sharp effect that imbalances in block size have on the size of the core, and the minor impact that market size has on the size of the core. I consider markets with between 100 and 10,000 participants (evenly divided between doctors and hospitals) where the doctors randomize their preferences over a block of 21 hospitals, and hospitals randomize their preferences over a block of a varying number of doctors ranging between 1 and 21. For each market configuration I simulate 200 realizations by drawing uniformly from the possible preferences. For each realization I compute the Doctor-optimal and Hospital-optimal stable matchings.

I present three major findings in two graphs. The first finding, that market outcome depends sharply on is presented in figure 1 which shows the average relative rank of doctors' in the hospital and doctor optimal matchings over a range of market imbalances. This graph shows that even a small imbalance in market conditions leads to a large improvement in the match outcome for the favored side. Most importantly, this effect also dramatically shrinks the difference between the optimal and pessimal matches for each side. In a balanced market there is approximately a 5-rank difference in the expected outcome between the optimal and pessimal match, but with with even a single increment of imbalance in their favor, the difference in expected match quality between optimal and pessimal matches is much smaller.

The final finding is that the size of the core does not depend on the size of the market. Figure 2 shows that as the market grows in size, the fraction of agents with multiple stable matches is insensitive to market size, remaining above 40% as the market grows to 5,000 Doctors and 5,000 Hospitals.

REFERENCES

- Ashlagi, Itai, Yash Kanoria, and Jacob D. Leshno.** 2017. "Unbalanced random matching markets: The stark effect of competition." *Journal of Political Economy*, 125(1): 69–98.
- Immorlica, Nicole, and Mohammad Mahdian.** 2005. "Marriage, Honesty, and Stability."

- Kojima, Fuhito, and Parag A. Pathak.** 2009. “Incentives and Stability in Large Two-Sided Matching Markets Author (s): Fuhito Kojima and Parag A . Pathak Source : The American Economic Review , Vol . 99 , No . 3 (Jun ., 2009), pp . 608-627 Published by : American Eco.” *American Economic Review*, 99(3): 608–627.
- Roth, Alvin E.** 1985. “The college admissions problem is not equivalent to the marriage problem.” *Journal of Economic Theory*, 36(2): 277–288.
- Roth, Alvin E, and Elliott Peranson.** 1991. “A natural experiment in the organization of entry-level labor markets: regional markets for new physicians and surgeons in the United Kingdom.” *The American economic review*, 415–440.
- Roth, Alvin E., and Elliott Peranson.** 1999. “The redesign of the matching market for american physicians: Some engineering aspects of economic design.” *American Economic Review*, 89(4): 748–780.

APPENDIX

A1. *Deferred Acceptance*

The deferred acceptance mechanism is a mainstay of this paper. In the doctor-proposing version of this mechanism, the algorithm iteratively selects a doctor d who proposes to their most preferred hospital that has not yet rejected them, unless the doctor has been rejected by all hospitals who they prefer to being unmatched. If the hospital prefers the proposing doctor to the current tentative assignment, the hospital holds the doctor’s proposal and rejects their current assignment (if one exists). Otherwise, the hospital rejects d . This process repeats until all doctors are either tentatively matched or have run out of proposals to make, and arrives at a stable matching that is weakly preferred by all doctors to any other stable matching. We call this the doctor-optimal stable match, because it is weakly preferred by all doctors to any other stable match. A parallel result exists for the hospital-proposing algorithm.

A2. *Review of Preferences From Immorlica and Mahdian (2005)*

Immorlica and Mahdian (2005) construct preferences as follows (substituting men and women for doctors and hospitals):

Let \mathcal{D} be an *arbitrary* fixed distribution over the set of women such that the probability of each woman in \mathcal{D} is nonzero. Intuitively, having a high probability in \mathcal{D} indicates that a woman is popular. The preference lists are constructed by picking each entry of the list according to \mathcal{D} , and removing the repetitions. More precisely, we construct a

random list (l_1, \dots, l_k) of k women as follows. At step i , repeatedly select women w independently according to \mathcal{D} until $w \notin (l_1, \dots, l_{i-1})$ and then set $l_i = w$.

LEMMA 2: *For any ε_a , no more than k/ε_a women can be acceptable ε_a fraction of men.*

PROOF:

Proceeding by contradiction, suppose the set of women who are acceptable to ε_a fraction of men is $J = \{j_1, \dots, j_p\}$, with $p > k/\varepsilon_a$.

Each woman's probability in \mathcal{D} is at least ε_a/k

The on any draw the likelihood of choosing some $j \in J$ from \mathcal{D} is at least

$$p \frac{\varepsilon_a}{k} > \frac{k}{\varepsilon_a} \frac{\varepsilon_a}{k} = 1$$

This is a contradiction.

□

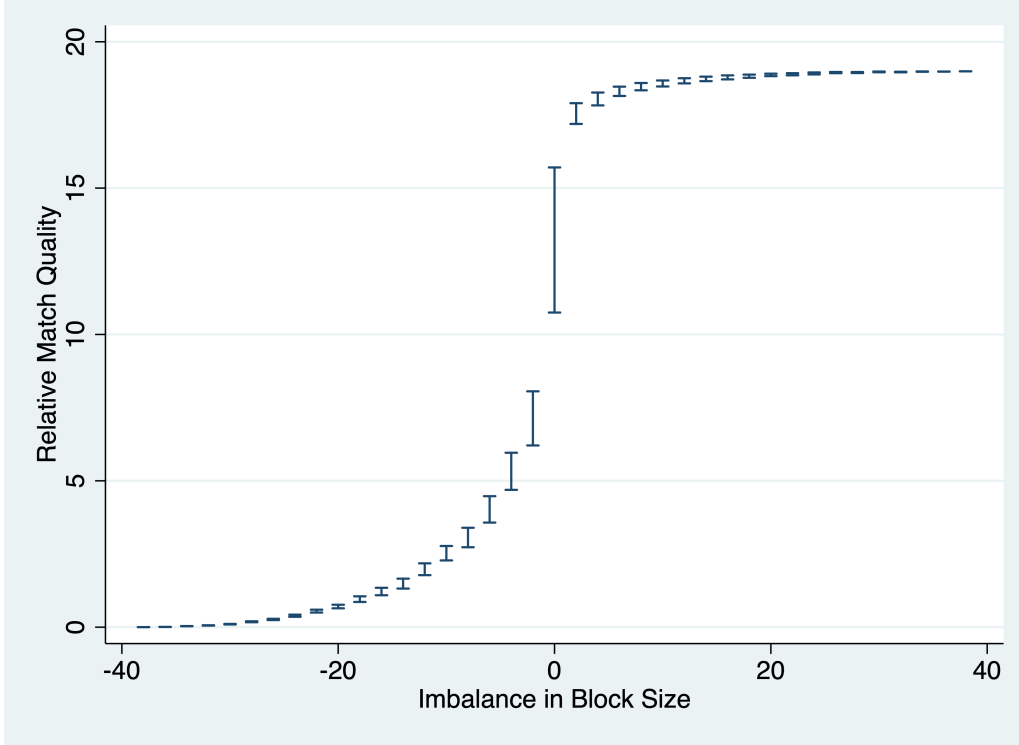


FIGURE 1. MATCH QUALITY AS A FUNCTION OF BLOCK IMBALANCE: THIS CHART SHOWS THE AVERAGE MATCH QUALITY DOCTORS EXPERIENCE. BARS INDICATE DOCTORS' MATCH QUALITY IN THE DOCTOR-OPTIMAL AND DOCTOR-PESSIMAL MATCH. A RELATIVE MATCH QUALITY OF 0 INDICATES THAT DOCTOR d_i MATCHED WITH THEIR i 'TH CHOICE HOSPITAL, AND A RELATIVE MATCH QUALITY OF 20 INDICATES THAT d_i MATCHED WITH THEIR $(i - 20)$ TH CHOICE HOSPITAL. BLOCK SIZE IMBALANCE IS THE DIFFERENCE BETWEEN THE SIZE OF THE INDICATES THE DIFFERENCE BETWEEN THE SIZE OF THE DOCTORS' RANDOMIZED BLOCK AND THE HOSPITALS'. IN ALL CASES THE LARGER RANDOMIZED BLOCK IS 41 ENTRIES LONG, AND THERE ARE 500 HOSPITALS AND 500 DOCTORS.

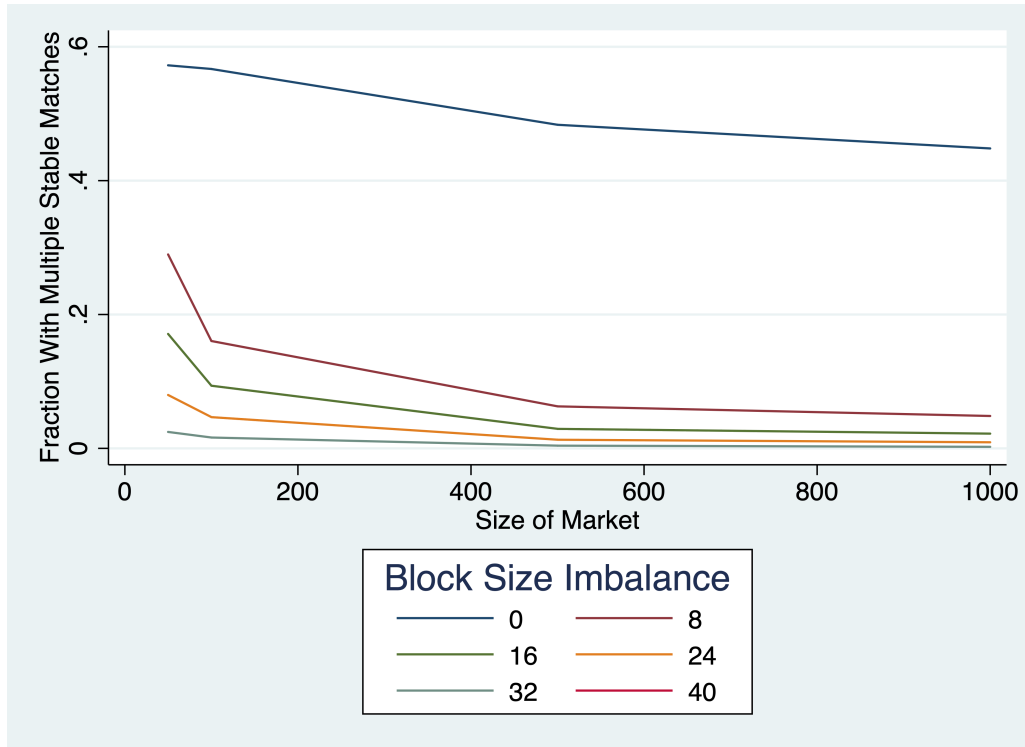


FIGURE 2. SIZE OF CORE AS A FUNCTION OF MARKET SIZE AND IMBALANCE: THIS CHART SHOWS THE FRACTION OF MARKET PARTICIPANTS WHO HAVE MORE THAN ONE STABLE OUTCOME AS A FUNCTION OF MARKET SIZE FOR A VARIETY OF MARKET BALANCE CONDITIONS. IN ALL CASES THE LARGER RANDOMIZED BLOCK IS 41 ENTRIES LONG.