Games and Economic Behavior 87 (2014) 591–615

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Optimal truncation in matching markets ☆

Peter Coles [a,*], Ran Shorrer [b,c]

[a] eBay Data Labs, United States
[b] Harvard University, United States
[c] Harvard Business School, United States

ARTICLE INFO

Article history:
Received 23 June 2013
Available online 24 February 2014

JEL classification:
C78
D47

Keywords:
Matching
Deferred acceptance
Truncation
Market design
Manipulation

ABSTRACT

Although no stable matching mechanism can induce truth-telling as a dominant strategy for all participants (Roth, 1982), recent studies have presented conditions under which truthful reporting by all agents is close to optimal (Immorlica and Mahdian, 2005; Kojima and Pathak, 2009; Lee, 2011). Our results demonstrate that in large, uniform markets using the Men-Proposing Deferred Acceptance Algorithm, each woman's best response to truthful behavior by all other agents is to truncate her list substantially. In fact, the optimal degree of truncation for such a woman goes to 100% of her list as the market size grows large. In general one-to-one markets we provide comparative statics for optimal truncation strategies: reduction in risk aversion and reduced correlation across preferences each lead agents to truncate more. So while several recent papers focused on the limits of strategic manipulation, our results serve as a reminder that without pre-conditions ensuring truthful reporting, there exists a potential for significant manipulation even in settings where agents have little information.

1. Introduction

One of the great success stories in economic theory is the application of matching theory to two-sided markets. A classic example is the National Resident Matching Program (NRMP), in which medical school students are matched to residency positions in hospitals. Rather than hospitals pursuing students via a decentralized series of offers, refusals and acceptances, matching occurs via a centralized mechanism. In this mechanism, each student ranks the hospital programs, and each hospital ranks the students. They submit these lists to an algorithm, which determines which students will be matched to which programs.

Such a centralized process has a number of advantages. First and foremost, the algorithm on which this and many similar centralized processes are based produces an outcome that is *stable* with respect to reported preferences.[1] In a stable matching, no two agents mutually prefer each other to their assigned match, nor does any matched participant prefer to be unmatched. A second advantage is that eliminating a decentralized offer process may save time and other resources.

☆ We thank Yakov Babichenko, B. Douglas Bernheim, William Diamond, Michael Grubb, Avinatan Hassidim, John Hatfield, Guido Imbens, Jacob Leshno, Jonathan Levin, Paul Milgrom, Muriel Niederle, Mikko Packalen, and Al Roth for helpful discussions, and two anonymous referees for providing detailed comments. We thank the Stanford Institute for Economic Policy Research for their generous support.

* Corresponding author.
E-mail addresses: pacoles@gmail.com (P. Coles), rshorrer@hbs.edu (R. Shorrer).

[1] In 1998, the algorithm used in the NRMP was altered to accommodate student couples and allow for specialized hospital positions, so that the outcome is "close to" a stable matching (see Roth and Peranson, 1999).

http://dx.doi.org/10.1016/j.geb.2014.01.005
0899-8256/

Finally, as Roth and Xing (1994) have shown, a centralized mechanism can successfully halt the unraveling of a market.[2] Centralized matching mechanisms also power a variety of other markets, including the public school systems in New York, Boston, Singapore and other cities, as well as numerous specialized medical fellowships.

These centralized markets all employ versions of an algorithm proposed by Gale and Shapley (1962). The algorithm, which in one-to-one markets is often referred to as the *Men-Proposing Deferred Acceptance Algorithm* (*MP-DA*), takes as its inputs preference lists reported by agents, and outputs a stable matching. When agents are asked to report preference lists for submission to *MP-DA*, this begs the question: Do all agents have an incentive to report truthfully? Dubins and Freedman (1981) and Roth (1982) provide the answer: they do not. In fact, Roth showed that no mechanism that produces stable matchings will induce truth-telling as a dominant strategy for all agents. However, in the preference list submission game induced by *MP-DA*, for all participants on *one side* of the market, the "men," truth-telling is a dominant strategy.[3] But this leaves open the question of how participants on the other side of the market, the "women," might benefit by strategically misrepresenting their preferences.

Recent work has examined conditions under which gains to strategic manipulation are limited for all participants in the market, not just those on one side. One approach in the literature concerns *large markets*. Roth and Peranson (1999) observe that in the data from the NRMP, very few participants could have improved their outcomes by reporting different preferences. They show via simulations that when the length of preference lists is held fixed and the number of participants grows, the size of the set of stable matching shrinks (a property they term "core convergence"), so that opportunities for manipulation are reduced. Immorlica and Mahdian (2005) demonstrate this result theoretically, finding that in large marriage markets where preference list length is bounded, nearly all players have an incentive to truthfully report preferences. Kojima and Pathak (2009) generalize this result, showing that in many-to-one markets, preference list manipulation, as well as other modes of strategic manipulation such as non-truthful reporting of capacities (see also Sönmez, 1997), are again limited. Lee (2011) considers one-to-one matching markets where agent utilities are drawn from distributions with bounded support that have both a common and an independent component. He shows that when all agents report truthfully, the proportion of participants who can achieve a significant utility gain from manipulation vanishes as the market grows large.[4]

Our approach takes a different tack. We do not require preference lists to be short, and ask: how should players optimally misrepresent preferences in markets that do not satisfy non-manipulability conditions? How "far" could optimal behavior be from truthfulness? We wish to study optimal manipulation, along with payoffs and market-wide welfare effects, and ask how strategic behavior and outcomes change as we vary market conditions.

The particular form of strategic misrepresentation we focus on is preference list *truncation*; that is, listing in order the first several partners from one's true preference list, and identifying all other partners as unacceptable. Truncation has an intuitive logic: by listing less-preferred partners as unacceptable, the probability of being matched with these partners drops to zero. Agents using this strategy might hope that correspondingly, the likelihood of being matched to a partner who remains on the truncated list will go up. In the context of *MP-DA*, this intuition is confirmed: submitting a truncated preference list weakly increases the likelihood of being matched to some agent on the truncated list, regardless of beliefs about the lists other agents submit. But submitting a truncated preference list is a risky strategy. Limiting acceptable partners also increases the likelihood of ending up with no match. Analysis of this tradeoff is the crux of the results in this paper.

While always a method for weakly increasing the likelihood of matching with better-ranked opponents, in some uncertain settings, truncation is optimal: Roth and Rothblum (1999) show that when agents' beliefs satisfy a form of symmetry termed "$\mathcal{M}$-symmetry," they can do no better than to truncate. Ehlers (2004) demonstrated that this result holds under somewhat more general conditions.

Whether optimal or not, we analyze truncation, both in symmetric and general settings. We ask: to what degree should players truncate, if at all? (Note that submitting one's true preference list is also a form of truncation.) Can a participant realistically gain from truncation when she is extremely uncertain about what opponents might report? If players anticipate that others may be truncating, how does this affect their behavior? Do participants truncate in equilibrium? What are the welfare implications in a truncation equilibrium?

To evaluate the consequences of truncation, we first characterize the payoff from truncation for a woman with general beliefs over the preference lists other agents will submit in terms of the distributions of her most and least preferred achievable mates. In a market with N men and N women, when a woman believes submitted preferences of others are uniformly chosen from the set of all full-length preference lists, she may safely truncate a large fraction of her list with low risk of becoming unmatched. Further, as there is a large gap between the expected rank of the mate she receives from truthful revelation and her most preferred achievable mate (Pittel (1989) shows these asymptote to $N/\log N$ and $\log N$

[2] Before the NRMP was introduced in the 1950s, offers and interviews were made as early as the fall of students' third year in medical school, which was undesirable for a number of reasons. The willingness of both hospitals and students to participate argues strongly in favor of the program's effectiveness. The NRMP enjoys participation rates of close to 100% of eligible students, with over 38 000 students participating in the March 2012 match.

[3] This is true in one-to-one, or "marriage" markets, where each agent has the capacity to match with at most one other agent. In many-to-one settings, e.g. students matching to hospitals, truth-telling is no longer a dominant strategy when, in the Deferred Acceptance Algorithm, the "hospitals" side makes the offers (see Roth, 1985).

[4] In a very different approach, Featherstone and Mayefsky (2010) run lab experiments in 5×5 marriage markets, and find that participants have trouble learning to find beneficial deviations under *MP-DA*, even if there are potential gains (though participants have more success in finding successful manipulation when facing "priority" mechanisms).