Optimal Truncation in Matching Markets*

Peter Coles†

Harvard Business School

July 2009

Abstract

Since no stable matching mechanism can induce truth-telling as a dominant strategy for all participants, there is often room in matching markets for strategic misrepresentation. In this paper we study a natural form of strategic misrepresentation: reporting a truncation of one’s true preference list. Roth and Rothblum [14] prove an important but abstract result: in certain symmetric settings, agents optimally submit some truncation of their true preference lists. In this paper we constructively examine truncation, both in symmetric and general settings, when agents must submit preference lists to the Deferred Acceptance Algorithm.

We first characterize truncation payoffs in a general setting in terms of the distribution of an agent’s achievable mates. Compared to truthful reporting, in any equilibrium in truncation strategies, welfare diverges for men and women: women prefer the truncation equilibrium, while men would prefer that participants truthfully report. In a uniform setting, we obtain a formula for truncation payoffs and demonstrate that both the optimal degree of truncation and the benefits from truncation can be substantial: agents might wish to truncate 90% of their lists, generating large gains in utility. While several recent papers have focused on the limits of strategic manipulation, this result serves as a reminder that even in settings where agents have little information, gains to manipulation can be significant.

Returning to the general environment, we show that the more risk averse a player, the lower the degree of her optimal truncation. Finally, when correlation in preferences increases, players should truncate less, implying that correlation too suggests less room for strategic misrepresentation.

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

---

*I am grateful to Jonathan Levin for his many suggestions throughout the writing of this paper. Doug Bernheim, Michael Grubb, John Hatfield, Paul Milgrom, Muriel Niederle, Mikko Packalen, and Al Roth offered helpful comments and discussions. I thank the Stanford Institute for Economic Policy Research for generous support.

†Harvard Business School, pcoles@hbs.edu
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 pair of agents would mutually prefer each other to their
assigned match, nor would any participant prefer to have remained unmatched. A second advantage is
that eliminating a decentralized offer process may save time and other resources. Finally, as Roth and
Xing [16] have shown, a centralized mechanism can successfully halt the unraveling of a market.\(^2\) Other
markets that successfully employ centralized matching mechanisms, apart from the NRMP, are the public
school systems in New York, Boston and Singapore, where students are matched with schools, and the
market for specialized fellowships.

These centralized markets all employ versions of an algorithm proposed by Gale and Shapley [6]. The
algorithm, often referred to as the Deferred Acceptance Algorithm (DAA), takes as its inputs preferences
reported by agents, and outputs a stable matching. When agents are asked to report preferences for
submission to the DAA, this begs the question: Do all agents have an incentive to report preferences
truthfully? Dubins and Freedman [4] and Roth [12] 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 the DAA, for all participants
on one side of the market, truth-telling is a dominant strategy.\(^3\) But this leaves open the question of
how participants on the other side of the market 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 [13] observed that in the data from the NRMP, very few participants
could have improved their outcomes by reporting different preferences. They showed 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 [7] 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 [9] generalize this result, showing that in many-

\(^1\)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. (c.f. Roth and Peranson [13]).

\(^2\)Before the NRMP was introduced in the 1950s, offers and interviews were made as early as the fall of their junior 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 25,000 students participating in the March 2004 match.

\(^3\)This is true in ‘marriage markets,’ markets where each agent has the capacity to match with at most one other agent.
In many-to-one settings, truth-telling is no longer a dominant strategy when in the DAA, the ‘many’ side makes the offers
(c.f. Roth and Sotomayor [15]).
to-one markets, preference list manipulation, as well as other modes of strategic manipulation such as non-truthful reporting of capacities (c.f. Sönmez [17]), are again limited.

Our approach takes the opposite tack. In markets that do not satisfy non-manipulability conditions, how should players optimally misrepresent preferences? We wish to study the optimal form of manipulation, payoffs, and market-wide welfare effects, and ask how strategic behavior and outcomes change as we vary certain 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 as unacceptable low-ranked partners, 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 (or at the very least, not go down). When agents are submitting preference lists to the DAA, this intuition is confirmed: Submitting a truncated preference list weakly increases the likelihood of being matched to some agent on the list, regardless of beliefs about what lists other agents will report. While always a method for weakly improving likelihood of matches with high-ranked opponents, in some uncertain settings, truncation is optimal: Roth and Rothblum [14] show that when agents beliefs satisfy a certain type of symmetry (denoted $M$-symmetry), agents can do no better than to truncate.

In this paper we constructively examine truncation, both in symmetric and general settings. We ask to what degree players should truncate, if it 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? What is the welfare of players in a truncation equilibrium?

Submitting a truncated preference list is a risky strategy. By limiting her acceptable choices, a participant is effectively rejecting offers in the hope of receiving a better offer. By doing so, the chances of receiving a top choice mate will increase. However, the likelihood of ending up with no match also increases. Analysis of this tradeoff is the crux of the results in this paper.

To evaluate the consequences of truncation, we first characterize the payoff from truncation for a player with general beliefs in terms of the distributions of her highest and lowest achievable mates. Since good approximations of these distributions are known when beliefs are uniform, we can use these approximations to solve for payoffs as a function of degree of truncation. Hence, we can graph the payoff function and find the optimal degree of truncation. We illustrate payoffs for the case of linear utility, that is, when agents care about the rank of their partner. In this setting, we find the degree of truncation to be significant: when others report truthfully, agents truncate 90% of their lists, generating large payoff gains.

Under uniform preferences, we demonstrate the existence of a symmetric equilibrium where all players use truncation strategies, and illustrate the equilibrium in the case when agents have linear utility. In
the unique pure strategy truncation equilibrium, agents continue to truncate at high levels: each agent lists only 50% of her full preference list. In general settings, the two sides of the market diverge in their tastes for truncation equilibria. Compared to the outcome from truthful reporting, women prefer any truncation equilibrium; for men the opposite is true.

Relaxing the symmetry requirement and returning to the environment where players have arbitrary beliefs, we examine comparative statics. We find that optimal truncation levels correspond to risk preference: regardless of beliefs over reported opponent preferences, the more risk-averse a player, the less she should truncate. We then turn to correlation in players’ preferences. The correlated preferences we consider are meant to capture the notion that in most settings, agents largely agree in their preferences over partners on the other side of the market, but that an individual’s preferences may idiosyncratically depart from common opinion. We find that the higher the likelihood a participant places on opponents having preferences similar to her own, the less she should truncate. Roth and Peranson [13] found via simulations that when preferences are correlated, the set of stable matchings is small, and therefore the set of submitted preference lists that could lead to gains is small. The findings in this paper provide theoretical evidence to corroborate their result.

To place this analysis in context, several comments are in order. While, for the reasons stated earlier, we believe truncating the bottom of one’s list is an intuitive manipulation in the preference list submission problem, in different environments eliminating higher-ranked members from one’s preference list might be a reasonable strategy. For example, in the job market for economists (c.f. Coles et al [2]), departments may choose not to interview certain highly-accomplished candidates, reasoning that these candidates will receive offers they prefer more. In efforts to best use costly and scarce interview slots, departments may effectively ‘top-truncate’ their preferences lists, focusing instead on candidates more likely to ultimately accept an offer. At a general level, it is when frictions in the market generate costs that this behavior arises. Lee and Schwarz [10] consider a setting where information acquisition is costly, so that firms prefer to interview workers who have a high likelihood of accepting (and likelihood is based on the number and identity of other firms interviewing a worker). Coles, Kushnir and Niederle [3] consider a setting where workers can signal their preferences to firms, so that firms may choose not to make offers to higher-ranked candidates, and instead make offers to candidates who have indicated likeliness to accept. In our paper, the analysis is performed after any costly information gathering has taken place, so these considerations do not arise.

The rest of the paper is organized as follows: Section 2 lays out the stable marriage setting and illustrates the fundamental tradeoff associated with truncation. In Section 3 we characterize the return to truncation, first for general beliefs, then in a uniform setting. In Section 4, we consider agent welfare in a truncation equilibrium. Section 5 and Section 6 examine how truncation behavior relates to risk preferences and correlation of agent preferences, respectively. Section 7 concludes.
2 Matching Markets Background

We begin by setting out the basic model of matching. In contrast to some of the well-known matching papers, we approach the notion of preferences of participants from a cardinal rather than an ordinal perspective, which allows us to discuss choice under uncertainty. As we shall see, the results involving ordinal preferences can be discussed using the preference orderings induced from cardinal utilities.

2.1 Marriage Markets and Stability

A marriage market of size \( N \) consists of a triplet \((M, W, u)\), where \( M \) is the set of men, \( W \) is the set of women, \(|M| = |W| = N\), and \( u \) is the profile of preferences for men and women.

Preferences for man \( m \in M \) are given by a utility function \( u_m : W \cup \{m\} \rightarrow \mathbb{R} \). Man \( m \) derives utility \( u_m(w) \) from woman \( w \) and \( u_m(m) \) from remaining single. For simplicity, we assume that \( u_m \) is one-to-one, so that \( m \) is never indifferent between any two options. Preferences \( u_w \) for women are defined similarly on the set \( M \cup \{w\} \).

As \( u_m \) is one-to-one, \( m \)'s preferences \( u_m \) induce a strict preference relation \( P_m \) over \( W \cup \{m\} \). We refer to \( P_m \) as \( m \)'s preference list. For example, if \( N = 3 \), \( u_m(w_1) > u_m(w_3) > u_m(w_2) > u_m(m) \) yields preference list \((w_1, w_3, w_2, m)\), meaning he prefers woman \( w_1 \) to \( w_3 \) to \( w_2 \) to being single. Note that man \( m \) may prefer bachelorhood over some of the women. For example, \((w_1, w_3, m, w_2)\) indicates that \( m \) prefers \( w_1 \) to \( w_3 \) to remaining single to \( w_2 \). We say that man \( m \) finds \( w \) acceptable if \( m \) prefers \( w \) to remaining single. Preference lists \( \{P_w\} \) are defined similarly for women, and we define \( P \) to be the profile of preference lists.

A matching is a pairing of men and women, so that each woman is assigned at most one man and each man at most one woman. Formally, a matching \( \mu \) is a mapping from \( M \cup W \) to \( M \cup W \) such that for every \( m \in M, \mu(m) \in W \cup \{m\} \), and for every \( w \in W, \mu(w) \in M \cup \{w\} \), and also for every \( m, w \in M \cup W, \mu(m) = w \) if and only if \( \mu(w) = m \). We refer to \( \mu(m) \) as \( m \)'s wife and \( \mu(w) \) as \( w \)'s husband. When \( \mu(x) = x \), agent \( x \) is single or unmatched under matching \( \mu \). The terms partner and mate are also used.

As is standard, we assume that each agent cares only about his or her partner in a matching. Therefore, \( i \) prefers matching \( \mu_1 \) to \( \mu_2 \) if and only if \( u_i(\mu_1(i)) > u_i(\mu_2(i)) \).

A matching is stable if no agent desires to leave his or her mate to remain single, and no pair of agents mutually desire to leave their mates and pair with each other. Given a matching \( \mu \), we say that it is blocked by \((m, w)\) if \( m \) prefers \( w \) to \( \mu(m) \) and \( w \) prefers \( m \) to \( \mu(w) \). A matching \( \mu \) is individually rational if for each \( x \in M \cup W \) with \( \mu(x) \neq x \), \( x \) finds \( \mu(x) \) acceptable. A matching \( \mu \) is stable if it is individually rational and is not blocked. In general, more than one stable matching may exist for given preferences.

Given preferences \( u \), woman \( w \) is achievable for \( m \) if there is some stable matching \( \mu \) in which \( w = \mu(m) \).
Achievable mates of women are defined similarly.

2.2 The (Men-Proposing) Deferred Acceptance Algorithm

In their seminal 1962 paper, Gale and Shapley prove that in any marriage market there exists a stable matching. To demonstrate this result, they propose an algorithm - the Deferred Acceptance Algorithm (DAA) - to generate a stable matching given any profile of preferences lists. The DAA has a number of desirable properties, and variants of it are employed in many real world centralized matching markets today, including the market for medical residents.

The DAA takes as its input a set $M$ of men, a set $W$ of women, and preference list profile $P$. The output is a matching $g(P) \equiv \mu^M$ over $M \cup W$. The algorithm works iteratively as follows.

- **Step 1.** Each man proposes to the first woman on his preference list. Each woman then considers her offers, rejects all men deemed unacceptable, and if any others remain, rejects all but her most preferred mate.

- **Step $k$.** Each man who was rejected in step $k-1$ makes an offer to the next woman on his preference list. If his preference list is exhausted, or if he prefers bachelorhood to the next woman on his list, he makes no offer. Each woman behaves as in step 1, considering offers in hand (including any man she has retained from the previous step) and rejects all but her most preferred acceptable suitor.

- **Termination.** If in any step $k$, no man makes an offer (i.e. men all either have offers in hand or have exhausted their lists of acceptable women,) the algorithm terminates. Each woman is paired with her current mate and this matching is final.

This algorithm must terminate in finite time. Gale and Shapley provide a remarkable characteristic of the resulting outcome.

**Theorem 1.** (Gale-Shapley) The matching $\mu^M$ resulting from the DAA is stable. Furthermore, for any other stable matching $\mu$, every man (weakly) prefers $\mu^M$ to $\mu$ and every woman weakly prefers $\mu$ to $\mu^M$.

The matching produced by the DAA is stable, which is a positive result for all participants. But men are particularly satisfied with the outcome. For the men, the algorithm produces the optimal stable matching, based on reported preferences. For the women, however, this is not the case. As we will see, this feature may mean some women prefer to strategically misreport preferences, causing the algorithm to produce a different matching.\(^4\)

\(^4\)A women-proposing version of the DAA will naturally produce an analogous outcome: a stable matching that, among the set of stable matchings, is optimal for women. We refer to the outcomes from the men-proposing DAA and the women-proposing DAA as the men-optimal and women-optimal stable matchings, respectively.
2.3 The Preference List Submission Problem for Men

We now turn to the incentive properties of the DAA. That is, in a setting where agents are asked to announce their preferences for submission to the algorithm, we ask if they have an incentive to report something other than the truth. We will see that women may, while men do not.

Consider a set of agents $I = M \cup W$. Agent $i \in I$ with preferences $u_i$ is asked to submit a preference list to the DAA. The agent’s beliefs about what preference lists others will report are given by the random variable $\tilde{P}_{-i}$ which takes as its range $\mathcal{P}_{-i}$, the set of all possible preference list profiles for others.

Agent $i$ solves the Preference List Submission Problem:

$$\max_{\hat{P}_i \in \mathcal{P}_i} \mathbb{E}[u_i(g(\hat{P}_i, \tilde{P}_{-i}))(i)]$$

Dubins and Freedman\[4\], and Roth\[12\] have shown that for any man $m$ with preferences $u_m$ and beliefs $\hat{P}_m$, it is optimal for $m$ to submit his true preference list $P_m$ (which corresponds to $u_m$).

**Theorem 2.** (Dubins and Freedman; Roth) In the Preference List Submission Problem,

$$P_m \in \arg \max_{P_m \in \mathcal{P}_m} \mathbb{E}[u_m(g(\hat{P}_m, \tilde{P}_{-m}))(m)].$$

2.4 The Preference List Submission Problem for Women

For women submitting preferences to the DAA, truth-telling may not be optimal. One way a woman $w$ might misrepresent preferences is by submitting a truncation of her true preference list; that is, listing in order the first several men from her true preference list and declaring all other men unacceptable. When $w$ is uncertain about what preference lists others will submit, truncation generates a tradeoff: $w$ may match with the high-ranked men she leaves listed with increased probability, but may also increase the likelihood of $w$ remaining unmatched. In this section we demonstrate this tradeoff, pose the problem of optimal truncation, and describe conditions so that in the Preference List Submission Problem, among all possible preference list submission strategies, truncation is optimal.

2.4.1 Strategic Truncation

By truncating her preference list, a woman strategically rejects certain offer and may wind up obtaining a better offer as a result. But over-truncation can leave the woman ending up unmatched. The following example demonstrates this principle when preference lists are known.
Suppose men and women have the following preferences:

<table>
<thead>
<tr>
<th></th>
<th>$m_1$</th>
<th>$m_2$</th>
<th>$m_3$</th>
<th>$w_1$</th>
<th>$w_2$</th>
<th>$w_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w_3$</td>
<td>$w_2$</td>
<td>$w_1$</td>
<td>$m_1$</td>
<td>$m_1$</td>
<td>$m_1$</td>
<td></td>
</tr>
<tr>
<td>$w_1$</td>
<td>$w_1$</td>
<td>$w_2$</td>
<td>$m_2$</td>
<td>$m_3$</td>
<td>$m_2$</td>
<td></td>
</tr>
<tr>
<td>$w_2$</td>
<td>$w_3$</td>
<td>$w_3$</td>
<td>$m_3$</td>
<td>$m_2$</td>
<td>$m_3$</td>
<td></td>
</tr>
</tbody>
</table>

We consider the strategic incentives of woman 1. Assume all other agents report truthfully. First, suppose $w_1$ submits her true preferences. In this case, the DAA stops after one step and $w_1$ is matched to $m_3$, her least preferred mate. The stable matching is indicated in bold.

<table>
<thead>
<tr>
<th></th>
<th>$m_1$</th>
<th>$m_2$</th>
<th>$m_3$</th>
<th>$w_1$</th>
<th>$w_2$</th>
<th>$w_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w_3$</td>
<td>$w_2$</td>
<td>$w_1$</td>
<td>$m_1$</td>
<td>$m_1$</td>
<td>$m_1$</td>
<td></td>
</tr>
<tr>
<td>$w_1$</td>
<td>$w_1$</td>
<td>$w_2$</td>
<td>$m_2$</td>
<td>$m_3$</td>
<td>$m_2$</td>
<td></td>
</tr>
<tr>
<td>$w_2$</td>
<td>$w_3$</td>
<td>$w_3$</td>
<td>$m_3$</td>
<td>$m_2$</td>
<td>$m_3$</td>
<td></td>
</tr>
</tbody>
</table>

Now suppose that $w_1$ misrepresents her preferences and submits the truncated list $(m_1, m_2)$. In this case, she will reject man $m_3$’s first round offer. Man $m_3$ must then make an offer to $w_2$ in the next round. Woman $w_2$ will accept $m_3$ over $m_2$, who made her an offer in the previous round. Man $m_2$ then finds himself single, and must make an offer to $w_1$. Woman $w_1$ accepts $m_2$’s offer, and the algorithm terminates yielding the matching in bold below.

<table>
<thead>
<tr>
<th></th>
<th>$m_1$</th>
<th>$m_2$</th>
<th>$m_3$</th>
<th>$w_1$</th>
<th>$w_2$</th>
<th>$w_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w_3$</td>
<td>$w_2$</td>
<td>$w_1$</td>
<td>$m_1$</td>
<td>$m_1$</td>
<td>$m_1$</td>
<td></td>
</tr>
<tr>
<td>$w_1$</td>
<td>$w_1$</td>
<td>$w_2$</td>
<td>$m_2$</td>
<td>$m_3$</td>
<td>$m_2$</td>
<td></td>
</tr>
<tr>
<td>$w_2$</td>
<td>$w_3$</td>
<td>$w_3$</td>
<td>$m_3$</td>
<td>$m_2$</td>
<td>$m_3$</td>
<td></td>
</tr>
</tbody>
</table>

Therefore, by truncating her list, $w_1$ improves her outcome.

To see how truncation can be problematic, suppose $w_1$ truncates her list even more and submits $(m_1)$ only. In this case, the algorithm will leave her unmatched, as shown in the final matching below.

<table>
<thead>
<tr>
<th></th>
<th>$m_1$</th>
<th>$m_2$</th>
<th>$m_3$</th>
<th>$w_1$</th>
<th>$w_2$</th>
<th>$w_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w_3$</td>
<td>$w_2$</td>
<td>$w_1$</td>
<td>$m_1$</td>
<td>$m_1$</td>
<td>$m_1$</td>
<td></td>
</tr>
<tr>
<td>$w_1$</td>
<td>$w_1$</td>
<td>$w_2$</td>
<td>$m_3$</td>
<td>$m_2$</td>
<td>$m_2$</td>
<td></td>
</tr>
<tr>
<td>$w_2$</td>
<td>$w_3$</td>
<td>$w_3$</td>
<td>$m_3$</td>
<td>$m_2$</td>
<td>$m_3$</td>
<td></td>
</tr>
</tbody>
</table>

Of course if woman $w_1$ is certain about the preferences others will submit to the algorithm, she can easily identify the optimal truncation. Proposition 1 characterizes woman $w$’s match when she submits a truncated version of her preference list, demonstrating generally how truncation can lead to the three outcomes in the above example. Define $P^k_w$, the preference list which includes in order only $w$’s $k$ most-
preferred men, as the $k$-truncation of her true preference list $P_w$. If fewer than $k$ men are acceptable to $w$, then $P^k_w \equiv P_w$.

**Proposition 1.** Let $P$ be the preference list profile of all agents $M \cup W$. Let $P^k$ be the preference profile in which woman $w$ submits a $k$-truncation of her preference list and everyone else submits their true preference lists to the DAA. Let $\mu$ and $\mu^k$ be the matchings generated from submitted profiles $P$ and $P^k$ respectively. Then $\mu^k(w)$ is $w$’s lowest ranked achievable mate (under $P$) with rank $\leq k$. Should no such mate exist, $\mu^k(w) = w$.

The example illustrates a general principle: Given the preference lists submitted by others, truncation by woman $w$ can have one of three consequences:

1. No effect. Woman $w$ has truncated below her lowest achievable mate
2. Improvement. Woman $w$ truncates above her lowest ranked mate, and is matched with the lowest ranked achievable mate above the point of truncation.
3. Unmatched. Woman $w$ has over-truncated, truncating above her highest ranked achievable mate.

If woman $w$ is certain of the preference lists $P_{-w}$ others are submitting, her truncation decision is simple: she calculates her highest achievable mate under $P = (P_w, P_{-w})$ and truncates her list just below him. If instead $w$ believes her opponents will submit preference lists according to some probability distribution, then truncating her list at $k$ generates a lottery over outcomes in which her partner will be her $j$th ranked choice, $j \in \{1, \ldots, k\}$, or else she will be unmatched. It is the tradeoff between improvement and becoming unmatched that is the crux of analysis in this paper.

### 2.4.2 Optimality of Truncation

Truncation is not the only possible misrepresentation of preferences. A woman could reverse two men in her preference list, list men as acceptable who are in fact unacceptable, or use some combination of these. However, under certain conditions we can be sure truncation is optimal.

The next proposition states that under certainty, women can do no better than to truncate.

**Proposition 2.** Suppose woman $w$ has preferences $u_w$ and believes others will report preference lists $P_{-w}$ to the DAA. Let $\mu^W(w)$ be $w$’s partner when $(P_w, P_{-w})$ are reported to the women-proposing version of the algorithm. Then $w$’s optimal strategy is to truncate her list so that $\mu^W(w)$ is the lowest ranked acceptable partner.

Perhaps surprisingly, when a woman has very little information about the preference lists others might report, she again can do no better than to truncate. In order to gain from swapping positions of two
men in her report preference list, a woman must have very specific knowledge about preferences of others over these men. Without this knowledge, it is best to leave the men in their correct order. Roth and Rothblum [14] demonstrate this principle using the following framework.

Let woman \( w \)'s beliefs about reported preference lists of others be given by \( \tilde{P}_{-w} \), a random variable that can take on a finite number of values in \( \mathcal{P}_{-w} \). If \( P_{-w} \) is a preference list profile for agents \(-w\), define \( P_{m\rightarrow m'}^{-w} \) to be the preference list profile in which \( m \) and \( m' \) swap preference lists, and all women swap the positions of \( m \) and \( m' \) in their lists. We say that woman \( w \)'s beliefs are \((m, m')\)-symmetric if

\[
\Pr( \tilde{P}_{-w} = P_{-w} ) = \Pr( \tilde{P}_{-w} = P_{m\rightarrow m'}^{-w} )
\]

for all \( P_{-w} \in \mathcal{P}_{-w} \). For a subset \( \mathcal{M}' \subseteq \mathcal{M} \), beliefs \( \tilde{P}_{-w} \) are \( \mathcal{M}' \)-symmetric if they are \((m, m')\)-symmetric for all \( m, m' \in \mathcal{M}' \).

**Theorem 3.** (Roth and Rothblum) Suppose \( w \)'s beliefs about reported preference lists of others are \( \mathcal{M} \)-symmetric. Then any preference list \( \hat{P}_w \) she might submit to the DAA is (weakly) stochastically dominated by a truncation of her true preference list.

Hence, when \( w \) is certain about reported preference lists of her opponents, or when she has extreme (but symmetric) uncertainty, truncation is optimal.

### 2.4.3 The Truncation Problem

Even when truncation is not optimal, we may sometimes wish to restrict the choice set for women to truncations of her true preference list. We define the **Truncation Problem** for woman \( w \) with preferences \( u_w \) and beliefs \( \tilde{P}_{-w} \) as

\[
\max_{k \in \{1, \ldots, N\}} \mathbb{E}[u_w(g(P^k_w, \tilde{P}_{-w})(w))]
\]

where \( P^k_w \) is a \( k \)-truncation of her true preference list \( P_w \), and \( g(\cdot) \) gives the outcome from the DAA. For convenience, whenever we consider the Truncation Problem for a woman \( w \), we will relabel men so that \( w \) has \( m_1 \succ m_2 \succ \ldots \succ m_N \).

### 3 Characterizing Truncation Payoffs

In this section we characterize woman \( w \)'s payoff from submitting a truncated version of her true preference list, first in terms of the distribution of her highest and lowest achievable mates. When a woman believes that reported preference lists of her opponents are distributed uniformly over the set of all preference lists, good approximations for these distributions are known. Furthermore, under uniform beliefs (in fact, under any \( \mathcal{M} \)-symmetric beliefs), we can pin down the exact returns to truncation, conditional on truncation yielding an improvement. Using the approximations and the “conditional-on-improvement” result, we can graph \( w \)'s payoff as a function of her truncation point. We find that in the uniform beliefs setting, the optimal degree of truncation can be significant: for linear utility, \( w \) prefers to truncate 90% of her preference list. Simulations in the next section show that the unique symmetric truncation equilibrium...
(where each woman may truncate) can also involve a high degree of truncation. These results demonstrate that even in a setting where agents possess very little information about opponent preferences, there is still room for significant strategic misrepresentation.

3.1 Truncation under General Beliefs

In this subsection, we characterize woman $w$’s payoff from submitting a truncated version of her true preference list in terms of the distribution functions for her highest and lowest achievable mates. The results build on Proposition 1, which illustrates how in settings of certainty, a woman may gain, lose or see no change from truncation.

We return to the Truncation Problem for woman $w$ with preferences $u_w$ and beliefs $\tilde{P}_w$ about reports of others. Throughout the section, $u_w$ (and hence, $P_w$) is fixed, so we can define $w$’s payoff from $k$-truncation when others submit preference lists $P_{-w}$ as

$$v(k, P_{-w}) \equiv u_w(g(P^h_w, P_{-w})(w)).$$

Notice that $v(N, P_{-w})$ gives $w$’s payoff if she reports truthfully, and $v(k, P_{-w}) = u_w(w)$ if $k$-truncation leaves $w$ unmatched. The Truncation Problem then becomes

$$\max_{k \in \{1, \ldots, N\}} \mathbb{E}[v(k, \tilde{P}_{-w})].$$

To evaluate $\mathbb{E}[v(k, \tilde{P}_{-w})]$, we condition on the three possible effects of truncation: no effect, improvement, or leaving $w$ unmatched.

Define $k_l(P_{-w})$ to be the rank of $w$’s mate under the DAA outcome $g(P_w, P_{-w})$. That is, $k_l(P_{-w})$ gives the rank of $w$’s lowest achievable mate when $-w$ report preference lists $P_{-w}$. Let $f(\cdot)$ be the probability mass function of the random variable $k_l(\tilde{P}_{-w})$ so that

$$f(x) = \operatorname{pr}(k_l(\tilde{P}_{-w}) = x)$$

for $x \in \{1, \ldots, N\}$. Let $F(\cdot)$ be the associated distribution function.

Similarly define $k_h(P_{-w})$ to be the rank of $w$’s highest achievable mate when $-w$ report $P_{-w}$. Let $g(\cdot)$ be the probability mass function of the random variable $k_h(\tilde{P}_{-w})$ so that

$$g(x) = \operatorname{pr}(k_h(\tilde{P}_{-w}) = x)$$

for $x \in \{1, \ldots, N\}$. Let $G(\cdot)$ be the associated distribution function.

Using $F(\cdot)$, $G(\cdot)$, and Proposition 1, we can express $w$’s payoff from $k$-truncating her list using the law

\footnote{Found, for example, via the women-proposing version of the DAA.}
of conditional expectations.\footnote{A complete derivation is given in the Appendix.}
\begin{equation}
E[v(k, \hat{P}_w)] = F(k) \cdot \sum_{i=1}^{k} \frac{f(i)}{F(k)} u_w(m_i) \\
+ [G(k) - F(k)] \cdot E[v(k, P_{\hat{w}}) \mid \hat{P}_w \in \mathcal{P}_2(k)] \\
+ [1 - G(k)] \cdot u_w(w),
\end{equation}
where the set \( \mathcal{P}_2(k) = \{P_{\hat{w}} \mid v(k, P_{\hat{w}}) > v(N, P_{\hat{w}})\} \) gives the cases when truncation yields an improvement, compared to truthful reporting. When truncation leaves \( w \) unmatched, her payoff is clearly \( u_w(w) \), and when truncation has no effect, the likelihood of being matched with \( x \) is \( f(x)/F(x) \).

This characterization of truncation payoffs is useful because in certain settings, we can compute \( F(\cdot) \), \( G(\cdot) \) and \( E[v(k, \hat{P}_w) \mid \hat{P}_w \in \mathcal{P}_2(k)] \). This will allow us to examine optimal truncation points and the payoff benefits from optimal truncation.

In the next section two sections, which will focus on the middle term of the sum in 3.1.1; that is, we will put structure on the returns to truncation conditional on truncation yielding an improvement.

3.2 Truncation as a Two Stage Process

In this section we show that woman \( w \) submitting a truncated version of her preference list to the DAA is equivalent to a two stage process: In the first stage, agents submit true preferences. In the second stage, \( w \) ‘divorces’ her mate, which sets off a chain of new offers and proposals. This equivalence will help us analyze the consequences of truncation by considering outcomes conditional on what happens in the first stage.

We will demonstrate the equivalence of two algorithms by proving that for a given input, the algorithms generate the same output, which in each case is the DAA outcome.

Both algorithms take as input a profile \( P \) of preference lists, an arbitrary woman \( w \), and a truncation point \( k \in \{1, \ldots, N\} \) and output a matching. The first algorithm plainly produces the DAA outcome\footnote{This algorithm, outcome equivalent to the DAA, is due to McVitie and Wilson \cite{McVitie1961}. This algorithm differs from the DAA in that men make offers one at a time, instead of in rounds.} \( g(P_k w, P_{\hat{w}}) \), the DAA outcome when \( w \) \( k \)-truncates her preference list. The second algorithm takes \( g(P_w, P_{\hat{w}}) \) as a starting point. If the rank of \( w \)’s partner is greater than \( k \), then we divorce \( w \) from her mate and all men with rank \( \geq k \) are deemed unacceptable. Following a chain of new offers, another matching results.

Let \( \mu_1(P, k, w) \) be the output of Algorithm 1 and let \( \mu_2(P, k, w) \) be the output of Algorithm 2.

\textbf{Algorithm 1}

\begin{itemize}
\item **Step 0.** Initialization. Identify the lowest achievable mate \( m_l \) of woman \( w \).
\end{itemize}
Iteration over steps 1 and 2:

- **Step 1.** Choosing a man. Pick any single man other than \( m_{l} \) who has not exhausted his preference list. If no such man exists, pick \( m_{l} \). If we have picked \( m_{l} \), and \( m_{l} \) is not single, or if \( m_{l} \) has exhausted his preference list, terminate.

- **Step 2.** The man chosen in the previous step makes an offer to the next woman on his preference list. If the woman finds the man acceptable and prefers him to her current mate, she holds his offer and divorces her previous mate (if any). Return to step 1.

**Algorithm 2**

- **Step 0.** Initialization. Run the DAA to find the men-optimal matching \( g(P) \). If \( w \)'s mate \( m \) has rank \( \leq k \) in \( P_{w} \), terminate. Otherwise, divorce \( w \) from \( m \). Declare candidates ranked lower than \( k \) unacceptable for \( w \).

Iteration over steps 1 and 2:

- **Step 1.** Pick an arbitrary single man who has not exhausted his preference list. If no such man exists, terminate.

- **Step 2.** The man chosen in the previous step makes an offer to the next woman on his preference list. If the woman finds the man acceptable and prefers him to her current mate, she holds his offer and divorces her previous mate (if any). Return to step 1.

The next proposition states that the two algorithms yield identical outcomes.

**Proposition 3.** For all \( k \in \{1, \ldots, N\} \), \( P \in \mathcal{P} \) and \( w \in \mathcal{W} \), we have \( \mu_1(P, k, w) = \mu_2(P, k, w) \).

To show the equivalence, we demonstrate that the algorithms reach a point after which they will coincide. The proof relies on the fact that in Algorithm 1, since \( m_{l} \) is chosen only when he is the sole single man available, from that point forward, there will never be more than one single man. The same is true in Algorithm 2 following \( w \)'s divorce from her initial mate. We show that the temporary match and the remaining preference lists are identical for the two algorithms at these two stages, so thereafter the algorithms coincide. A detailed proof is in the Appendix.

With this equivalence in hand, when we consider the submission of a truncated preference list, we can think of it as a two stage process, focusing on the chain of offers in Algorithm 2. We will be interested in whether the chain will end with the last single man exhausting his list, or with a man proposing to woman \( w \). These outcomes correspond to truncation leaving \( w \) unmatched, and truncation yielding an improvement over truthful reporting, respectively. As we shall see in the next section, knowing that following a ‘divorce’ \( w \) will receive at most one more offer will allow us to calculate the returns to truncation, conditional on truncation yielding an improvement.
3.3 Truncation under $M$-Symmetric Beliefs

In this section, we examine the Truncation Problem when woman $w$ has $M$-symmetric beliefs. We show that if $w$ is convinced that truncation will yield an improvement compared to truthful reporting, then her mate is equally likely to be any man she lists as acceptable. This is somewhat surprising, because when $w$ has unconditional $M$-symmetric beliefs and submits preferences in the Truncation Problem, we would certainly not expect woman $w$’s mate to be equally likely to be any man on her list: because of the stability requirement, she is far more likely to be matched with her higher-ranked mates.

Proposition 4. Suppose woman $w$’s beliefs $\tilde{P}_{-w}$ about the reported preferences of her opponents are $M$-symmetric. Then

$$\Pr\left(g(P^k_w, \tilde{P}_{-w})(w) = m_i \mid \tilde{P}_{-w} \in \mathcal{P}_2(k)\right) = \Pr\left(g(P^k_w, \tilde{P}_{-w})(w) = m_j \mid \tilde{P}_{-w} \in \mathcal{P}_2(k)\right)$$

for all $k \in \{1, \ldots, r - 1\}$, $i, j \in \{1, \ldots, k\}$. Hence,

$$\mathbb{E}[v(k, \tilde{P}_{-w}) \mid \tilde{P}_{-w} \in \mathcal{P}_2(k)] = \frac{\sum_{i=1}^{k} u_w(m_i)}{k}.$$  

The intuition in this result comes from Proposition 3, which show that truncation can be treated as a two stage process. Consider the settings where truncation will yield an improvement ($\tilde{P}_{-w} \in \mathcal{P}_2(k)$). By first reporting true preferences, and then divorcing her mate, we know that in the ensuing chain of offers, exactly when one man will make an offer to $w$, at which point the chain ends. By the symmetry of $w$’s beliefs, this is equally likely to be any of the men in her list.

Crucial to the reasoning is that since we know that truncation will yield an improvement, this corresponds to an algorithmic outcome where exactly one new offer is made to $w$. In the DAA generally, multiple offers may be made to $w$, making it difficult to pinpoint the distribution of her mate.

With Proposition 4 in hand, we can now express $w$’s payoff from truncation at $k$ as

$$\mathbb{E}[v(k, \tilde{P}_{-w})] = F(k) \cdot \frac{\sum_{i=1}^{k} f(i) u_w(m_i)}{k} + [G(k) - F(k)] \cdot \frac{\sum_{i=1}^{k} u_w(m_i)}{k} + [1 - G(k)] \cdot u_w(w).$$

(3.3.1)

3.4 Truncation under Uniform Beliefs

In this section we consider the special case of uniform beliefs. That is, when facing the Truncation Problem, we suppose that woman $w$ believes her opponents’ reported preference lists $P_{-w}$ to be chosen uniformly and randomly from the set of all possible full preference list profiles $\mathcal{P}_{-w}$, where a full preference
list profile is a profile in which each agent prefers any possible mate to being unmatched. Uniform beliefs are a special case of the $\mathcal{M}$-symmetric beliefs.\footnote{But uniformity is not equivalent to $\mathcal{M}$-symmetry. Under $\mathcal{M}$-symmetric beliefs, a woman may have specific knowledge about how the men rank her. For example, she may know that all the men prefer her to $w_2$. With uniform beliefs, such knowledge is ruled out.} Hence, under uniform beliefs, truncation is again optimal for women.

The stable marriage problem under uniform beliefs has received attention (particularly in the computer science literature) primarily because it offers a tractable setting where agents have very little information about preferences of others. See, for example, Knuth [8], Dzierzawa and Oméro [5], and Immorlica and Mahdian [7].

When the size of the market $N$ is large, $F(\cdot)$ and $G(\cdot)$, the distribution of $w$’s lowest and highest achievable mate, given beliefs $\tilde{P}$, take on a simple form: Dzierzawa and Oméro [5] show that for large $N$, $F(\cdot)$ and $G(\cdot)$ are both approximately exponential:

$$F(x) \approx 1 - e^{-x/k_m} \quad G(x) \approx 1 - e^{-x/k_w}$$

where $k_m = N \ln(N)$ and $k_w = \ln(N)$.

Intuition for this exponential decay in the rank of the lowest and highest achievable mates can be gleaned by analyzing the DAA. Under uniform beliefs, the average number of offers received by a woman $w$ is approximately $\ln(N)$ (see Knuth [8]). Woman $w$ holds the maximum of these offers, so that the rank of her lowest achievable mate is the first-order statistic. For uniformly random offers, the distribution of the first order statistic is the exponential $F(\cdot)$ given above. A similar (but slightly more complicated) argument holds for $G(\cdot)$.

Substituting these expressions for $F(\cdot)$ and $G(\cdot)$, we can approximate the payoff from truncation for any utility function $u_w$. To illustrate, we consider the case of linear utility; that is, when $w$ is interested in (minimizing) expected partner rank, and the utility from being unmatched is just worse than being matched to her least-preferred man.\footnote{Utility to $w$ from matching with her $i$th ranked man is $-i$, and $u_w(w) = -(N + 1)$.} Under uniform beliefs and linear utility, $w$’s expected payoff from truncating at $k$ becomes

$$E[v(k, \tilde{P}_w)] \approx -\left[ k_m - (k_m + k)e^{-k/k_m} + u_w(w)e^{-k/k_w} + (e^{-k/k_m} - e^{-k/k_w}) \cdot \frac{k + 1}{2} \right],$$

where the initial minus sign emerges because $w$ wishes to minimize partner rank. Truncation payoffs, as expressed by the RHS of equation 3.4.1 are strictly quasi-concave in $k$, with a unique optimum $k^*$ in $[1, N]$. Details of the derivation and of quasi-concavity can be found in the appendix.

Hence, when $w$ is interested in expected partner rank and believes opponents’ reported preferences to be uniform, we are able to calculate the expected payoff from truncation for any $k$, provided $N$ is large. We graph payoff as a function of truncation point $k$ when $N = 200$, using the approximations for $F(\cdot)$
and $G(\cdot)$.$^{10}$

Figure 1: Expected truncation payoff $N + \mathbb{E}[v(k, \tilde{P}_{-w})]$ as a function of $w$'s truncation point $k$. Beliefs for $w$ are uniform, and utility $u_w(\cdot)$ is linear.

Notice that in Figure 1 we plot $N + \mathbb{E}[v(k, \tilde{P}_{-w})]$, where the constant $N$ is added so that the graph lies above the $x$-axis.

The optimal truncation point is approximately $k = 20$, roughly a 90% truncation of $w$’s full list. We find this high degree of truncation to be surprising: despite the extreme uncertainty about her opponents’ reported preferences, it is to woman $w$’s advantage to significantly misrepresent her preferences.

Starting from the point of no truncation ($k = 200$), additional truncation yields small but steady gains. However, the gains begin to tail off as $w$ submits shorter and shorter lists as additional truncation poses the risk of leaving $w$ unmatched with increasing likelihood.

Intuition for this payoff structure can be gleaned from the exponential nature of the distributions. When $k$ is large, truncating an additional man from (say from $k + 1$ to $k$) is unlikely to be pivotal. We can see this from the exponential decay distribution of her lowest achievable mate $G(\cdot)$, which indicates that her mate under the DAA is likely to be at the top of her list and is unlikely to be $k + 1$. But when truncation is pivotal, the potential gains are large, since if she improves her match, she is equally likely to be matched with any mate ranked $1 \ldots k$. Furthermore, the chances that incremental truncation leaves her unmatched are extremely small ($w$ would have to truncate even above her highest achievable mate), and the losses from being unmatched compared to being matched to a very low-ranked mate are also small.

The result brings to mind settings where agents must expend resources to learn their preferences (e.g.

$^{10}$Simulations of truncation under uniform beliefs result in nearly identical graphs, even for small $N$.}

16
by interviewing candidates, or by flying out to visit hospitals). For men that a woman expects to rank low on her list but is unsure of their order, costly acquisition of information is fruitless because when behaving optimally, she should not list them anyway.

4 Truncation Equilibrium and Welfare

In this section, we consider the setting where agents must submit preference lists to the DAA. We demonstrate that in equilibria in truncation strategies, compared to outcomes from truthful preference list reporting, welfare for men is greater, welfare for women is lower, and the expected number of matches is lower. Under uniform beliefs, we demonstrate the existence of a symmetric equilibrium in truncation strategies. We use simulations to show that in a setting with linear payoffs, there is a unique pure strategy truncation equilibrium, and in this equilibrium, all women truncate roughly 50% of their preference lists.

4.1 Welfare in Truncation Equilibria under General Beliefs

Define the Preference List Submission Game as follows: Let $U$ be the set of all possible utility profiles for $I = M \cup W$. Let $\phi(\cdot)$ be a distribution over $U$. For any agent $i \in I$, let the message space $S_i = \mathcal{P}_i$, the set of all possible preference lists for player $i$. Recall that $g(P)$ gives the Gale-Shapley matching for reported preference lists $P$. The Preference List Submission Game is the Bayesian game defined by

$$\Gamma = \langle I, \mathcal{P}, g(\cdot), U, \phi(\cdot) \rangle.$$ 

A strategy for agent $i$ is a mapping $s_i : U_i \rightarrow \mathcal{P}_i$. A mixed strategy for $i$ is a mapping $\sigma_i : U_i \rightarrow \Delta(\mathcal{P}_i)$ which describes a randomization over submitted preference lists for each possible type. Define a truncation strategy for woman $w$ as a strategy in which $w$ always reports a truncated version of her true preferences. We will restrict attention to equilibria where men report preferences truthfully, an assumption motivated by the dominant-strategy result in Theorem 2.

Define a Bayesian Nash equilibrium $\sigma = (\sigma_m, \ldots, \sigma_m, \sigma_w, \ldots, \sigma_w)$ in which men report truthfully and women mix over truncation strategies as an equilibrium in truncation strategies. The following theorem characterizes welfare in such equilibria.

Theorem 4. Let $\sigma$ be an equilibrium in truncation strategies. Then compared to outcomes from truthfully reporting preferences to the DAA,

1. welfare for women is (weakly) greater
2. welfare for men is (weakly) smaller
3. the expected number of matches is (weakly) smaller.
The results of Theorem 4 can be obtained by considering the effect of incremental truncations in light of Proposition 3. An incremental truncation for any woman $w$ has a (weakly) positive spillover on the welfare of other women: rejection of a man can only lead to more offers for other women. Furthermore, such a truncation can only negatively affect the welfare of other men: the ‘divorced’ man will receive a worse mate, and the chain of offers that follows will can only lead to worse mates for the other men as well. Finally, since the chain of offers will end in either a better offer for the ‘divorcing’ woman, or else in no match, the incremental truncation weakly decreases the number of matches. Since all incremental truncations affect welfare in the directions stated in the theorem, the same welfare effects hold in any truncation equilibrium.

A corollary of Theorem 4 is that we can make the same welfare calculations between two equilibria in truncation strategies when one equilibrium involves ‘more’ truncation (where ‘more’ can be defined in terms of stochastic dominance). This corollary brings to mind the welfare result in [3] in which signaling equilibria with varying cutoffs are compared. In each of these cases, actions by one side of the market serve to ‘shift the balance of power,’ so that when there are equilibria with varying degrees of action (truncation in this paper, response to signals in [3]), the sides of the market are at odds over which equilibrium is preferred.

### 4.2 Equilibria under Uniform Beliefs

Consider now the setting of uniform beliefs, so that each player is equally likely to have any full preference list. Additionally, players identically value a match with their $r$th ranked choice $\forall r \in \{1, \ldots, N\}$ and have identical value to being unmatched.

**Proposition 5.** There exists symmetric equilibrium $((\sigma_m)^N, (\sigma_w)^N)$ where men each use strategy $\sigma_m$ of truthful reporting and women each use strategy $\sigma_w$, which is a mixture over truncation strategies.

The proof follows from Kakutani’s fixed point theorem and Theorem 3 (Roth and Rothblum). Assume men report truthfully, and let us initially restrict attention to symmetric (mixed) truncation strategies for women $-w$. Consider the best response correspondence for $w$, where we again restrict best responses to being truncation strategies. By Kakutani’s theorem, there exists a fixed point. But by Theorem 3, this mixed truncation strategy for $w$ must also be a best response in the unrestricted domain of strategies for $w$. If it were not, then some non-mixed strategy would do strictly better. But since $w$’s beliefs in this setting are $\mathcal{M}$-symmetric, a truncation strategy weakly dominates this non-mixed strategy, which yields a contradiction.

### 4.3 Graphing the Symmetric Equilibrium

In this section we illustrate symmetric equilibria for the case of linear utility (the same functional form used in Figure 1, but with $N = 20$ instead of $N = 200$). We find that there is a unique pure strategy,
symmetric equilibrium, and that the common truncation point involves a non-trivial degree of truncation: women truncate roughly 50% of their lists. Strategies in this equilibrium are fragile, however, in that truncation yields payoffs only slightly higher than truthful reporting.

We first examine how returns to truncation for $w$ change when other women also truncate their lists.

![Figure 2: Woman $w$'s payoff as a function of her truncation point when women $-w$ truncate at $j$. $N=20$. Iterations = 1,000,000.](image)

In Figure 2 we examine the effect on $w$'s payoffs when women $-w$ all truncate at a common point $j$, where $j$ takes on various values. First for lower $j$, curves for $w$ are higher. This follows from the positive spillover of truncation: when $-w$ truncate their lists, this helps woman $w$.

In Figure 2 it is also apparent that as $w$'s opponents truncate more, from $j = 20$ to $j = 15$ to $j = 10$, $w$ should truncate less. This is interesting because it stands in contrast to the complete information result: under complete information, when another woman truncates, $w$'s highest achievable mate can only improve, so that $w$'s optimal degree of truncation is weakly higher.

But observe also that when $-w$ continue to truncate their lists from $j = 5$ to $j = 3$ to $j = 1$, the $w$'s optimal truncation is hard to observe because her payoff curve becomes very flat. The reasons for this flatness are two-fold: When others submit very short lists, the expected number of stable matchings is known to be very small (see, for example [7].) Hence, there should be minimal gains to truncation. However, since when others truncate, the expected partner rank for $w$ is very high so there is little danger that truncation will leave $w$ unmatched. The minimal risks and rewards to truncation lead to the flatness of payoff curves.

By running a very large number of iterations, it is possible to identify the peaks of all the curves in Figure 2. We observe that truncation under uniform beliefs is a case of strategic substitutes: as illustrated
in Figure 3, the optimal truncation point for \( w \) is inversely related to \( j \), the common truncation point of women \(-w\). Of course, due to the flatness of the expected payoff graph, optimal truncation points for small \( j \) are ‘just barely’ optimal.

![Figure 3: Optimal truncation point of woman \( w \) as a function of the common truncation point \( j \) of women \(-w\). \( N = 20 \). Iterations = 1,000,000.](image)

By overlaying the 45° line, we identify the point of truncation in the pure strategy symmetric equilibrium to be approximately 11, roughly a 50 percent truncation of the entire list.

When all women truncate at this common point, no single woman sees significant gains from truncation compared to truthful reporting. However, since truncation has a positive externality on other women, the equilibrium payoff is non-trivially greater than the payoff should all women report truthfully. Hence, it would be in the interest any woman to convince others that she is not truncating, thereby inducing them to truncate more.

This equilibrium produces two types of ex-post instability with respect to true preferences. The first involves women that have over-truncated and are left unmatched. The second is more subtle. When other women truncate, woman \( w \) may be paired with a mate that is not achievable under the true preferences. Both types of instability could lead to a potential post-market ‘scramble’ for positions.

Analysis of a post-market scramble could be an interesting application of the framework in this paper. If after the match a significant fraction of players remain single, a second, organized match might be helpful to find these players partners.\(^{11}\)

\(^{11}\)The American Economic Association organizes a Scramble in which candidates seeking jobs and employers with positions open late in the job market can announce their availability on the AEA website. See: http://www.aeaweb.org/joe/scramble
At first observation, the scramble would likely reduce the downside to remaining unmatched in the primary match. But this might induce additional risk-taking behavior (more truncation) by participants. Such behavior would increase the pool size in the second match, raising the value of being unmatched, inducing even more truncation. A secondary match might ultimately enjoy high participation levels, but only because it has drawn participants away from the primary match, complicating welfare analysis.

While gains to truncation can be significant, truncation is nevertheless a risky strategy. In an equilibrium where w’s opponents truncate, truncation for w is particularly risky: compared to truthful reporting, optimal truncation offers minimal benefit, and over-truncating can lead to large losses. One might expect agents with more conservative attitudes toward risk to shy away from this proposition. We address this question in the following section.

5 Truncation and Risk Aversion

In this section, we ask how a woman’s truncation behavior varies as we vary her attitude towards risk. We take as a starting point arbitrary preferences for woman w and arbitrary beliefs about reported preferences of others. We will show that for any beliefs about reports, woman w with preferences \( u_w(\cdot) \) will truncate more than a woman who has identical beliefs, but preferences given by \( \psi(u_w(\cdot)) \), where \( \psi(\cdot) \) is a concave transformation.

Recall that for a woman w with preferences \( u_w(\cdot) \), we defined

\[
v(k, P_{-w}) = u_w(g(P^k_w, P_{-w})(w)),
\]

the payoff from submitting truncated preference list \( P^k_w \). Now define

\[
v(\psi, k, P_{-w}) = \psi(u_w(g(P^k_w, P_{-w})(w))),
\]

the payoff from submitting truncated preference list \( P^k_w \) for a woman \( w_\psi \) who has preferences given by \( \psi(u_w(\cdot)) \), where \( \psi(\cdot) \) is a concave function.

The proposition states that if w prefers truncating less to more, then \( w_\psi \) definitely prefers truncating less to more.

**Theorem 5.** Let \( \tilde{P}_{-w} \) be any random variable over \( \mathcal{P}_{-w} \). Then

\[
\mathbb{E}
\left[
v(k, \tilde{P}_{-w})
\right] \leq \mathbb{E}
\left[
v(k + 1, \tilde{P}_{-w})
\right] \Rightarrow \\
\mathbb{E}
\left[
v(\psi, k, \tilde{P}_{-w})
\right] \leq \mathbb{E}
\left[
v(\psi, k + 1, \tilde{P}_{-w})
\right].
\]

The constructive proof nicely illustrates incremental truncation analysis, so we provide it in-text:
Proof. We focus on two lotteries over outcomes. Let $Q^{k+1}$ be the lottery over mates for $w$ when she truncates at $k+1$, and let $Q^k$ be the lottery when she truncates at $k$. We will show that $Q^k$ is a spread of lottery $Q^{k+1}$. If $Q^k$ is mean-decreasing from $w$’s perspective, then it will be mean-decreasing from $w_\psi$’s perspective as well.

$Q_{k+1}$ and $Q_k$ are shown in Figure 4. Recalling Proposition 3, $k$-truncating is equivalent to $(k+1)$-truncating followed by $k$-truncating. That is, lottery $Q^k$ is equivalent to starting with lottery $Q^{k+1}$, then rolling the die again if $w$ receives her $(k+1)$st ranked choice. Hence,

$$q_i^k \geq q_i^{k+1}$$

for $i \in \{1, \ldots, k\} \cup \{w\}$; that is, $Q^k$ is a spread of lottery $Q^{k+1}$.

Suppose first that $E[v(k, \tilde{P}_-w)] = E[v(k+1, \tilde{P}_-w)]$, so that from $w$’s perspective, $Q^k$ is a mean-preserving spread of $Q^{k+1}$. Then by Jensen’s inequality, $E[v_\psi(k, \tilde{P}_-w)] \leq E[v_\psi(k+1, \tilde{P}_-w)]$, and if $\phi(\cdot)$ is strictly concave, $E[v_\psi(k, \tilde{P}_-w)] < E[v_\psi(k+1, \tilde{P}_-w)]$.

Now suppose that $E[v(k, \tilde{P}_-w)] < E[v(k+1, \tilde{P}_-w)]$, so that from $w$’s perspective, $Q^k$ is a mean-decreasing spread of $Q^{k+1}$. We will now construct an intermediate lottery $Q'$ such that

1. $Q^{k+1}$ first order stochastic dominates $Q'$ and
2. From $w$’s perspective, $Q^k$ is a mean preserving spread of $Q'$.

Define lottery $Q'$ so that $Q'$ is identical to $Q^{k+1}$, except that we replace outcome $k+1$ ($w$’s $(k+1)$st ranked choice) with lottery $\alpha(k+1) + (1-\alpha)w$. Choose $\alpha$ such that $w$ has $u_w(Q^k) = u_w(Q')$. Such
an $\alpha$ exists since because $q_i^k \geq q_i^{k+1}$ for $i \in \{1,\ldots,k\} \cup \{w\}$, choosing $\alpha = 1$ gives $u_w(Q') < u_w(Q^k)$. By construction, $Q^{k+1}$ first-order stochastic dominates $Q'$. But we also have that $Q^k$ is a spread of $Q'$, and by construction, is mean-preserving.

By first order stochastic dominance, $w_\psi$'s strictly prefers $Q^{k+1}$ to $Q'$. By Jensen, $w_\psi$ prefers $Q'$ to $Q^k$. Hence,

$$E[v_\psi(k, \tilde{P}_w)] < E[v_\psi(k + 1, \tilde{P}_w)].$$

□

We can now use Theorem 5 to sort optimal truncation points based on degree of concavity.

**Proposition 6.** Let $k_i$ be the minimum optimal truncation point and let $k^i$ be the maximum optimal truncation point for woman $i \in \{w, w_\psi\}$. Then $k_w \leq k_{w_\psi}$. Furthermore, if $\psi(\cdot)$ is strictly concave, then $k^w \leq k^w_\psi$.

**Proof.** If $k_w$ is $w$’s minimum optimal truncation point, then $w$ strictly prefers truncation at $k_w$ to truncation at any $k < k_w$. By Theorem 5, $w_\psi$ must then prefer truncation at $k_w$ to truncation at any $k < k_w$. Hence, $k_w \leq k_{w_\psi}$.

If $k^w$ is $w$’s maximum optimal truncation point, then $w$ prefers truncation at $k_w$ to truncation at any $k < k_w$. By Theorem 5, $w_\psi$ must then strictly prefer truncation at $k^w$ to truncation at any $k < k^w$. Hence, $k^w \leq k_{w_\psi}$. □

Thus, players who are more risk averse truncate less, with the set of optimal truncation points overlapping at at most one point.

The key insight in the analysis is the interpretation of truncation as a risky lottery, and the mapping of the additional risk associated with incremental truncation to an extra lottery a woman must face. If a woman doesn’t like to face the extra lottery, then certainly a woman with more concave preference will not want to face it.

Note that despite pertaining to risk aversion, the results in this section do not restrict us to the consideration of concave utility functions $u_w(\cdot)$. Rather, it is the relative risk aversion that is crucial. For example, if we restrict ourselves to a class of $s$-shaped utility functions, we know that within the class of $s$-shaped utility functions, concave transformations induce less truncation.

In a very general sense, this result can be taken as advice to participants. Players can observe the patterns of behavior of others, size up their own attitudes toward risk, and truncate more or less accordingly.
6 Correlated Preferences

In this section, we let woman \( w \) believe that with some probability, other women in the market share her preferences. We consider how woman \( w \) should vary her degree of truncation as this probability varies. We find that the greater the likelihood that others share preferences, the less woman \( w \) should truncate.

6.1 Perfectly Correlated Preferences

Consider first the case of perfectly correlated preferences on the women’s side of the market.

Remark 1. When women have identical preferences, there is a unique stable matching.

To see this, note that the top-ranked man, as agreed upon by all women, must be matched with his top choice in any stable matching, or else these two would constitute a blocking pair. The second-ranked man must then be matched to his top choice of the remaining women, and so on. The DAA reduces to a serial dictatorship, determined by the common ranking of the men.

Since there is a unique stable matching in this setting, misrepresentation of preference lists via truncation can never improve a woman’s match. In fact, if a woman is certain that other women share her preferences, but is uncertain about what men will submit to the algorithm, truncation can very well lower her outcome by leaving her unmatched.

6.2 Partially Correlated Preferences

In this section, we introduce a notion of partial correlation of preferences indexed by a single parameter \( \alpha \). We will show that the greater the degree of correlation, the less a woman should truncate.

We return to the Preference List Submission Problem, where woman \( w \) with preferences \( u_w \) has beliefs about reported preferences of opponents given by \( \tilde{P}_{-w} \). Let \( p(\cdot, \cdot) \) be the probability mass function for \( w \)’s beliefs. That is,

\[
p(p_M, p_{-\omega})
\]
gives the likelihood that the men will report preference lists \( P_M \) and other women \( -\omega = W \setminus \{w\} \) will report preference lists \( P_{-\omega} \). Define the marginal probability over mens’ preference profiles by \( p^M(\cdot) \).

Given \( p(\cdot, \cdot) \), define beliefs \( p^C(\cdot, \cdot) \) by

\[
p^C(P_M, P_{-\omega}) = \begin{cases} p^M(P_M) & \text{if } P_w = P_{\tilde{w}} \quad \forall \tilde{w} \in \{-\omega\} \\ 0 & \text{otherwise} \end{cases}
\]

\( p^C(\cdot, \cdot) \) is the distribution that preserves the marginal distribution over men’s preferences \( p^M(\cdot) \), but where the other women share the preferences of \( w \).
Define beliefs $p^\alpha(\cdot)$ by

$$p^\alpha(P_{-w}) = (1 - \alpha)p(P_{-w}) + \alpha p^C(P_{-w}).$$

Hence, as $\alpha$ varies from 0 to 1, $p^\alpha$ ranges from $p$ to $p^C$, keeping the marginal distribution over men’s preferences fixed, at the same time steadily increasing correlation of women’s preferences.

The set of optimal truncation points for woman $w$ with utility $u_w$ and beliefs indexed by $\alpha$ is given by

$$k^*(\alpha, p, u_w) = \arg\max_{k \in \{1, \ldots, N\}} E_{p^\alpha}[v(k, \tilde{P}_{-w})].$$

Notice that since the choice set is finite, $k^*(\cdot, \cdot)$ will be non-empty.

Let $k^h(\alpha, p, u_w) = \max[k^*(\alpha, p, u_w)]$ and $k^l(\alpha, p, u_w) = \min[k^*(\alpha, p, u_w)]$, the optimal choices involving the least and most truncation respectively.

The following proposition states that for any preferences $u_w$ and beliefs $p$, as we increase the degree of correlation $\alpha$, woman $w$ should truncate less.

**Proposition 7.** Let $\alpha, \alpha' \in [0, 1]$ with $\alpha' > \alpha$. Then

$$k^l(\alpha', p, u_w) \geq k^l(\alpha, p, u_w) \quad \text{and} \quad k^h(\alpha', p, u_w) \geq k^h(\alpha, p, u_w).$$

The proof relies on the fact that when there is a unique stable matching, it can never hurt to submit a full list. Using this fact, we can show that if under low correlation, $w$ prefers truncating less to more, then under high correlation $w$ definitely prefers truncating less to more. This is enough to sort optimal truncation points.

Intuition for this result is related to the size of the set of stable matchings. Truncation can yield improvement only when there are multiple stable matchings. The greater the degree of correlation, the smaller this set, and the lower the likelihood that a window for gain from truncation exists.

The anticipated level of correlation in the environment might influence the advice a market designer offers participants. If correlation is high, the designer can safely advise participants to report truthfully, and it is in their best interest to do so. With low correlation (sufficiently heterogenous preferences), players may anticipate gains from truncation, which if acted on, could lead to unstable matchings. In this case, a designer will wish to examine whether other market features might ensure truthful reporting to be optimal (for example, prevalence of short preference lists, as in [7] and [9]).

### 6.3 Noisy Preferences

In Section 6.2, a woman believes it is possible that opponents have preferences \textit{identical} to hers. In this section, woman $w$ believes women have preferences \textit{similar} to hers (but not necessarily identical).
We model such beliefs for women by generating noisy deviations from a common preference list. By performing simulations, we corroborate the theoretical results in Section 6.2: more correlation means a woman should truncate less.

We generate correlated preferences as follows. Each man \( m_i \) is assigned a random number \( r_i \sim U[0, 1] \). This represents the man’s “beauty,” to use the terminology of Caldarelli and Capocci [1]\(^{12} \), and this value is agreed upon by all women. Each woman \( w_j \) places weight \( \alpha \geq 0 \) on this value and adds idiosyncratic (noise) component \( q_{ij} \sim U[0, 1] \) to \( r_i \) for each man \( m_i \). Woman \( w_j \)’s rankings are determined by the sum \( \alpha \cdot r_i + q_{ij} \). From \( w \)’s perspective, the preferences of women are noisy versions of her own rankings. High \( \alpha \) means low noise, so \( \alpha \) measures the degree of correlation. Men are assumed to have uniformly random rankings over the women. The process just described is used only to determine rankings: we assume that utility is a linear function of partner rank, and remaining unmatched is just worse than being matched to one’s least-preferred man (so we can compare outcomes to the uniform case, as in Figure 1).

Figure 5 graphs the return to truncation for various levels of noise. For each level of noise, we used the procedure to randomly generate 1,000,000 preference lists and for each \( k \), we graph the average payoff from \( k \)-truncation.

\[ \begin{array}{ccc}
\alpha = 0 \text{ (no correlation)} & \alpha = .1 & \alpha = .5 \\
\alpha = .5 & \alpha = 1 & \alpha = 5 \text{ (most correlation)}
\end{array} \]

Figure 5: Woman \( w \)’s expected payoff as a function of her truncation point for various degrees of correlation. \( N = 200 \). Iterations = 1,000,000.

When \( \alpha = 0 \) (the top curve), this corresponds to the uniformly random preferences for the women discussed in Section 3.4. When \( \alpha \) is very large, all women rank men the same way and for any realization of preferences, the stable matching will be unique. Hence, the extremes for \( \alpha \) correspond to the extreme

\[ \text{In a model of college admissions, the starting point for preferences might be something like the US News and World Report ranking of universities.} \]
beliefs in Section 6.2.

From Figure 5, we make two key observations. First, woman \( w \) dislikes correlation. This fact is easy to explain. If all women agree on who the top men are, they ‘compete’ for them as mates. The lower the correlation, the less the competition, and the higher the expected mate for \( w \). Second, \( w \)’s optimal truncation point increases as correlation increases. This corroborates the result in Section 6.2: when there is more correlation, \( w \) should truncate less.

7 Conclusion

In this paper, we study optimal strategic behavior in matching markets under incomplete information. Among classes of strategies for preference list misrepresentation, truncation is an attractive option because it is guaranteed to raise (or at least, not decrease) the likelihood of matching with one’s highest ranked partners. By contrast, for more complicated strategies, such as swapping the order of agents in a preference list, detailed knowledge of the preferences reported by others may be necessary to ensure gains. As Roth and Rothblum [14] show, when agents have symmetric beliefs about reported preferences of others, no strategy can do better than truncation.

But in light of the recent work by Immorlica and Mahdian, and Kojima and Pathak, which demonstrates that in large markets where agents submit short preference lists, opportunities for manipulation are limited, one may ask whether agents – especially agents with very little detailed information – can ever substantially gain from manipulation. Our paper answers in the affirmative. Even when agents view reported preferences of others as being drawn uniformly from the set of all possible preferences, agents can truncate lists with little risk of being unmatched, but with the potential to see large gains. For many of the settings in which the DAA has been successfully applied, notably in the NRMP and in the Boston and New York school systems, the markets do reflect large numbers and short preference lists (and indeed, the costliness of information discovery may generally limit list length). Nevertheless, the results in the paper highlight that when these non-manipulability assumptions do not hold, agents may gain substantially from ‘cheating’ in a very simple way (both as individuals, and in equilibrium).

The analysis in this paper, and especially the discussion of risk preferences, offers a framework for predicting manipulation in matching markets. For example, one application concerns a topic that has recently received attention in labor markets: implementation of a second match for participants unmatched in the primary market. A proposal for a second match was considered and ultimately rejected by the NRMP. In the decentralized market for new economists, the ‘Scramble’ implemented by the AEA performs a similar function.

The NRMP envisioned that participants with appetite for risk might strategically truncate their preference lists:

“Of greatest concern: Students might shorten their ranking list and wait for the second match to list
their “safety” programs. That, says the NRMP, would mean ‘many positions may be filled in phase one by independent applicants, leaving more unmatched seniors at the conclusion of phase two.’

The second match would have the direct effect of improving outcomes for participants who remain unmatched after the primary market. But an indirect effect is that agents may be inclined to truncate more in the first period, as they become more comfortable with the risk of remaining unmatched. The quote suggests that this outcome is bad, although in actuality, the welfare effects are less clear. The results in this paper show that truncation has a positive spillover on the other women in the match. The framework in the paper, in combination with a structured model of a second match, could shed light on this issue.

8 Appendix: Proofs of Propositions

Proposition 1. Let $P$ be the preference profile of all players $M \cup W$. Let $P^k$ be the preference profile in which woman $w$ submits a list of only the top $k$ members on her preference list and everyone else submits their true preference lists to the DAA. Let $\mu$ and $\mu^k$ be the matchings generated from $P$ and $P^k$ respectively. Then $\mu^k(w)$ is $w$’s lowest ranked achievable mate (under $P$) with rank $\leq k$. Should no such mate exist, $\mu^k(w) = w$.

Proof. Observe first that any matching that is stable with respect to $P^k$ must be stable with respect to $P$, and that any matching $\tilde{\mu}$ that is stable with respect to $P$ with $\tilde{\mu}(w)$ ranked $\leq k$ must be stable with respect to $P^k$. Hence, setting

$$\mathcal{M}_1 = \{m \in M \mid m \text{ achievable for } w \text{ under } P^k\}$$

and

$$\mathcal{M}_2 = \{m \in M \mid m \text{ achievable for } w \text{ under } P \text{ and } m \text{ ranked } \leq k \text{ in } w\text{’s list}\}$$

we have $\mathcal{M}_1 = \mathcal{M}_2$. By the Gale-Shapley result, $\mu^k(w)$ is the lowest ranked element of $\mathcal{M}_1$, and hence of $\mathcal{M}_2$. Should both sets be empty, then $\mu^k(w) = w$. □

Proposition 2. Suppose woman $w$ has preferences $u_w$ and believes others will report preferences $P_{-w}$ to the DAA. Let $\mu^W(w)$ be $w$’s partner when $(P_w, P_{-w})$ are reported to the women-proposing version of the algorithm. Then $w$’s optimal strategy is to truncate her list so that $\mu^W(w)$ is the lowest ranked acceptable partner.

Proof. Suppose that $\tilde{P}_w$ resulted in matching $\tilde{\mu}$ with $\tilde{\mu}(w)$ ranked better than $\mu^W(w)$. Then truncating her list after $\tilde{\mu}(w)$ must also yield mate $\tilde{\mu}(w)$ for $w$, as $\tilde{\mu}$ will also be stable with respect to these


\(^{14}\)Symmetric to the men-proposing algorithm, with the roles of men and women reversed.
preferences. But then $\tilde{\mu}(w)$ must also be achievable with respect to $P$, contradicting the optimality of $\mu^W(w)$.

**Proposition 3.** For all $k \in \{1, \ldots, N\}, P \in \mathcal{P}$ and $w \in \mathcal{W}$, we have $\mu_1(P, k, w) = \mu_2(P, k, w)$.

**Proof.** Let $l$ be the rank of $w$’s lowest achievable mate $m_l$ under full preferences.

- If $k \geq l$, both algorithms output $\mu^M$, the men-optimal matching.
- If $k < l$, then the algorithms will reach a point where they coincide. That is there will be a point where the sequence of single men chose coincide, as do the temporary matchings and preference lists.

Consider Algorithm 2 directly after the initialization step.

In Algorithm 1, we claim that (1) at some point, $m_l$ will make an offer to $w$, which will be rejected. (2) From this point forward, the algorithm coincides with Algorithm 2.

(1) By Roth and Sotomayor [15], we know that $w$’s $k$-truncation of her list makes men (weakly) worse off. Hence, $m_l$ must be matched with a candidate worse than $w$, and his offer to $w$ must have been rejected.

(2) When $m_l$ makes his offer to $w$, no higher ranked man has done so. Otherwise, let $m'_l$ be the first man ranked higher than $m_l$ to make an offer to $w$ and backtrack to the point in the algorithm where he does so. Note that up to this point, the path of the algorithm is consistent with $w$ having $l$-truncated her preferences, since she has not faced any man ranked $k$ through $l$. But this implies that if $w$ $l$-truncated her list, she would receive a mate at least as good as $m'_l$, not $m_l$. This contradicts proposition 1.

By the choice-of-proposer rule in the algorithm, we know that when $m_l$ proposes to $w$, he must be the only single man who has not yet exhausted his list. If $w$ accepted $m_l$’s offer, the path of the algorithm would be consistent with $w$ having $l$-truncated her list, and the algorithm would terminate with matching $\mu^M$. Hence, by instead rejecting $m_l$, we arrive at exactly the position of Algorithm 2 following step 0.

Thereafter, the algorithms coincide, thus yielding identical outcomes.

**Derivation of Equation (3.3.1)**

Begin with the sets

- $\mathcal{P}_1(k) = \{ P_{-w} | v(k, P_{-w}) = v(N, P_{-w}) \}$
- $\mathcal{P}_2(k) = \{ P_{-w} | v(k, P_{-w}) > v(N, P_{-w}) \}$
- $\mathcal{P}_3(k) = \{ P_{-w} | v(k, P_{-w}) < v(N, P_{-w}) \}$

29
Further partition $\mathcal{P}_i(k)$. Define

$$
\mathcal{P}_{1A}(k) = \{ P_{-w} \in \mathcal{P}_1(k) \mid v(N, P_{-w}) = u_w(w) \} \\
\mathcal{P}_{1B}(k) = \{ P_{-w} \in \mathcal{P}_1(k) \mid v(N, P_{-w}) > u_w(w) \}.
$$

$\mathcal{P}_{1A}(k)$ is the set of opponent preference profiles where $k$-truncation yields no effect because even truthful reporting leaves $w$ unmatched, and $\mathcal{P}_{1B}(k)$ is the set of opponent preference profiles where $k$-truncation yields no effect on $w$’s partner in $M$.

Rewrite the expectation $\mathbb{E}[v(k, \tilde{P}_{-w})]$ as

$$
\mathbb{E}[v(k, \tilde{P}_{-w})] = \text{pr}(\tilde{P}_{-w} \in \mathcal{P}_{1A}(k)) \cdot \mathbb{E}[v(k, \tilde{P}_{-w}) \mid \tilde{P}_{-w} \in \mathcal{P}_{1A}(k)] \\
+ \text{pr}(\tilde{P}_{-w} \in \mathcal{P}_{1B}(k)) \cdot \mathbb{E}[v(k, \tilde{P}_{-w}) \mid \tilde{P}_{-w} \in \mathcal{P}_{1B}(k)] \\
+ \text{pr}(\tilde{P}_{-w} \in \mathcal{P}_{2}(k)) \cdot \mathbb{E}[v(k, \tilde{P}_{-w}) \mid \tilde{P}_{-w} \in \mathcal{P}_{2}(k)] \\
+ \text{pr}(\tilde{P}_{-w} \in \mathcal{P}_{3}(k)) \cdot \mathbb{E}[v(k, \tilde{P}_{-w}) \mid \tilde{P}_{-w} \in \mathcal{P}_{3}(k)],
$$

Since $w$ is unmatched when $\tilde{P}_{-w} \in \mathcal{P}_{1A}(k)$ or $\tilde{P}_{-w} \in \mathcal{P}_{3}(k)$, this becomes

$$
\mathbb{E}[v(k, \tilde{P}_{-w})] = \text{pr}(\tilde{P}_{-w} \in \mathcal{P}_{1A}(k)) \cdot u_w(w) \\
+ \text{pr}(\tilde{P}_{-w} \in \mathcal{P}_{1B}(k)) \cdot \mathbb{E}[v(k, \tilde{P}_{-w}) \mid \tilde{P}_{-w} \in \mathcal{P}_{1B}(k)] \\
+ \text{pr}(\tilde{P}_{-w} \in \mathcal{P}_{2}(k)) \cdot \mathbb{E}[v(k, \tilde{P}_{-w}) \mid \tilde{P}_{-w} \in \mathcal{P}_{2}(k)] \\
+ \text{pr}(\tilde{P}_{-w} \in \mathcal{P}_{3}(k)) \cdot u_w(w). \tag{8.0.1}
$$

We evaluate (8.0.1) component by component, beginning with the first component.

Define $k_l(P_{-w})$ to be the rank of $w$’s mate under the men-optimal outcome $g(P_{w}, P_{-w})$. That is, $k_l(P_{-w})$ gives the rank of $w$’s lowest achievable mate when preferences of others are $P_{-w}$. Let $f(\cdot)$ be the probability density of the random variable $k_l(\tilde{P}_{-w})$ so that

$$
f(x) = \text{pr}(k_l(\tilde{P}_{-w}) = x)
$$

for $x \in \{1, \ldots, r\}$. Let $F(\cdot)$ be the associated distribution function.

The first component of (8.0.1) then becomes

$$
f(r) \cdot u_w(w).
$$

Turning to the second component of (8.0.1), observe that $k$-truncation leaves $w$’s mate in $M$ unchanged exactly when she has truncated below the lowest achievable mate. Hence,

$$
\text{pr}(\tilde{P}_{-w} \in \mathcal{P}_{1B}(k)) = F(k).
$$
The second component of (8.0.1) then becomes

$$F(k) \cdot \sum_{i=1}^{k} \frac{f(i)}{F(k)} u_{w}(m_i),$$

where $\sum_{i=1}^{k} \frac{f(i)}{F(k)} u_{w}(m_i)$ is the expected value of $w$’s lowest achievable mate, given that this mate has rank $\leq k$.

We momentarily skip the third component and move to the fourth component of (8.0.1). Define $k_h(P_{-w})$ to be $w$’s highest ranked achievable mate when preferences of others are $P_{-w}$. Let $g(\cdot)$ be the probability density of the random variable $k_h(P_{-w})$ so that $f(x) = \Pr(k_h(P_{-w}) = x)$ for $x \in \{1, \ldots, r\}$. Let $G(\cdot)$ be the associated distribution function.

When $P_{-w} \in \mathcal{P}_\beta(k)$, $k$-truncation leaves $w$ unmatched when truthful reporting would have given her an acceptable mate. This happens exactly when $w$ truncates above her highest achievable mate, provided the highest achievable mate is in $\mathcal{M}$. Hence,

$$\Pr(P_{-w} \in \mathcal{P}_\beta(k)) = 1 - G(k) - g(r).$$

Finally, we turn to the third component of (8.0.1). Because the set $\{\mathcal{P}_A(k), \mathcal{P}_B(k), \mathcal{P}_2(k), \mathcal{P}_3(k)\}$ is a partition of $\mathcal{P}_{-w}$, we have

$$\Pr(P_{-w} \in \mathcal{P}_2(k)) = 1 - f(r) - [F(k)] - [1 - G(k) - g(r)].$$

Notice that $f(r)$ and $g(r)$ both give the probability that $w$ is unmatched even if she reports truthfully, so we have

$$\Pr(P_{-w} \in \mathcal{P}_2(k)) = G(k) - F(k).$$

We are now able to express the return to truncation in terms of $F(\cdot)$ and $G(\cdot)$.

$$\mathbb{E}[v(k, \tilde{P}_{-w})] = \begin{cases} f(r) & u_{w}(w) \\ F(k) & \sum_{i=1}^{k} \frac{f(i)}{F(k)} u_{w}(m_i) \\ [G(k) - F(k)] & \mathbb{E}[v(k, \tilde{P}_{-w}) | \tilde{P}_{-w} \in \mathcal{P}_2(k)] \end{cases}$$

and since $f(r) = g(r)$, we have

$$\mathbb{E}[v(k, \tilde{P}_{-w})] = \begin{cases} F(k) & \sum_{i=1}^{k} \frac{f(i)}{F(k)} u_{w}(m_i) \\ [G(k) - F(k)] & \mathbb{E}[v(k, \tilde{P}_{-w}) | \tilde{P}_{-w} \in \mathcal{P}_2(k)] \end{cases} + [1 - G(k)] \cdot u_{w}(w),$$

(8.0.2)

**Proposition 4.** Suppose woman $w$ believes the reported preferences of her opponents to be distributed
according to $\tilde{P}_{-w} \sim U(\mathcal{P}_{-w})$. Then
\[
pr\left(g(P^k_w, \tilde{P}_{-w})(w) = m_i \middle| \tilde{P}_{-w} \in \mathcal{P}_2(k)\right) = pr\left(g(P^k_w, \tilde{P}_{-w})(w) = m_j \middle| \tilde{P}_{-w} \in \mathcal{P}_2(k)\right)
\]
for all $k \in \{1, \ldots, r-1\}$, $i, j \in \{1, \ldots, k\}$. Hence,
\[
E[v(k, \tilde{P}_{-w}) \mid \tilde{P}_{-w} \in \mathcal{P}_2(k)] = \frac{\sum_{i=1}^{k} u_w(m_i)}{k}.
\]

Proof. For each $i$ and $k$, define
\[
\mathcal{P}^i_2(k) = \left\{ P_{-w} \in \mathcal{P}_2(k) \mid g(P^k_w, P_{-w})(w) = m_i \right\}.
\]
Since $\tilde{P}_{-w} \sim U(\mathcal{P}_{-w})$, we wish to show that $|\mathcal{P}^i_2(k)| = |\mathcal{P}^j_2(k)|$ for all $i, j, k$. We proceed by finding a bijection from $\mathcal{P}^i_2(k)$ to $\mathcal{P}^j_2(k)$.

For $i, j \in \{1, \ldots, k\}, i \neq j$, define the mapping $f_{ijk} : \mathcal{P}_2(k) \rightarrow \mathcal{P}_2(k)$ as follows. Let $f_{ijk}(P_{-w}) \equiv P'_{-w}$ be given by the following

1. Switch $m_i$ and $m_j$ everywhere. Switch the positions of $m_i$ and $m_j$ in each woman’s list, and swap $m_i$ and $m_j$’s preference lists. (This is like relabeling.)
2. Now switch back $m_i$ and $m_j$ in $w$’s list.

Notice that this is equivalent to swapping $m_i$ and $m_j$ in $w$’s list only, and then relabeling $i$ and $j$.

Suppose $P_{-w} \in \mathcal{P}^i_2(k)$. We will show that $P'_{-w} \in \mathcal{P}^j_2(k)$. Note that it is not immediately clear that we even have $P'_{-w} \in \mathcal{P}_2(k)$, that is, that under $P'_{-w}$, $k$-truncation still yields an improvement for $w$.

We think of the matching as arising from the DAA. Since $P_{-w} \in \mathcal{P}_2(k)$, if $w$ does not truncate, she will be matched with a man worse than $m_k$. Hence, during the process of the algorithm, she will not receive an offer from any man $m_1, \ldots, m_k$. Hence, rearranging these men in $w$’s list will not affect the outcome, and in particular, swapping $m_i$ and $m_j$ will not affect the outcome (the stable matching). Furthermore, since $P_{-w} \in \mathcal{P}^i_2(k)$ we know that under $P_{-w}$, $k$-truncation leaves $w$ matched with $m_i$. Using proposition 3, we know that during the chain of proposals following an ‘ex-post’ $k$-truncation by $w$, the first man to make an offer to $w$ will be $m_i$. Hence, this will still be true if $w$ swaps the position of $m_i$ and $m_j$ in her list.

Thus, we have that if $w$ switches $m_i$ and $m_j$ in her list, $k$-truncation will yield an improvement and she will again be matched with $m_i$. But now relabeling $m_i$ and $m_j$ (so that $w$’s list is $(m_1, m_2, m_3, \ldots$), we have that $P'_{-w} \in \mathcal{P}^j_2(k)$.

Hence, $f_{ijk}(\cdot)$ is a bijection from $\mathcal{P}^i_2(k)$ to $\mathcal{P}^j_2(k)$, which is sufficient to prove the proposition. □
Derivation of Equation 3.4.1 and Quasi-Concavity of RHS:

\[ \mathbb{E}[v(k, \tilde{P}_{-w})] \approx - \left[ k_m - (k_m + k)e^{-k/k_m} + u_w(w)e^{-k/k_w} + (e^{-k/k_m} - e^{-k/k_w}) \cdot \frac{k + 1}{2} \right]. \]

Furthermore, \( \mathbb{E}[v(k, \tilde{P}_{-w})] \) is strictly quasi-concave in \( k \), with a unique optimum \( k^* \) in \( [1, N] \).

Proof. Define

\[ \pi(k) \equiv \mathbb{E}[v(k, \tilde{P}_{-w})] \]
\[ = - \left[ \int_0^k x \, dF(x) + (G(k) - F(k)) \cdot \frac{k + 1}{2} + (1 - G(k)) \cdot u_w(w) \right], \]

where

\[ F(x) = 1 - e^{-x/k_m} \quad G(x) = 1 - e^{-x/k_w}. \]

Note that \( \pi(\cdot) \) is a smooth function of \( k \).

Using the functional forms of \( F(x) \) and \( G(x) \), we have

\[ \int_0^k x \, dF(x) = \int_0^k x \left( \frac{e^{-x/k_m}}{k_m} \right) \, dx \]
\[ = k_m - (k_m + k)e^{-k/k_m} \]
\[ (G(k) - F(k)) \cdot \frac{k + 1}{2} = (e^{-k/k_m} - e^{-k/k_w}) \cdot \frac{k + 1}{2} \quad \text{and} \]
\[ (1 - G(k)) \cdot u_w(w) = u_w(w)e^{-k/k_w}. \]

Hence, we have

\[ \pi(k) = - \left[ k_m - (k_m + k)e^{-k/k_m} + u_w(w)e^{-k/k_w} + (e^{-k/k_m} - e^{-k/k_w}) \cdot \frac{k + 1}{2} \right]. \]

Setting \( u_w(w) = -(N + 1) \) and differentiating with respect to \( k \), we have

\[ \pi'(k) = e^{-k/k_m} - \frac{N}{k_w} e^{-k/k_w} - \left\{ \frac{1}{2} \left[ e^{-k/k_m} + e^{-k/k_w} \right] + \frac{k - 1}{2} \left[ -\frac{1}{k_m} e^{-k/k_m} - \frac{1}{k_w} e^{-k/k_w} \right] \right\} \]
\[ = \left\{ e^{k(1/k_w-1/k_m)} \left( \frac{1}{2} + \frac{k - 1}{2k_m} \right) - \frac{N}{k_w} - \frac{1}{2} + \frac{k - 1}{2k_w} \right\} \cdot \frac{1}{e^{k/k_w}}. \]
Observe that \( a(k) \) is increasing in \( k \). Since \( \frac{1}{e^{k/k_w}} \) is positive, \( \pi'(k) \) has at most one zero. If \( a(k^*) = 0 \) for some \( k^* \in (1, N) \), \( k^* \) is the optimal truncation point. If \( a(k) \) has no zero in the range \([1, N]\), then \( \pi(k) \) is monotonic on the range \([1, N]\) and the optimal truncation point is either 1 or \( N \). \( \square \)

**Proposition 7.** Let \( \alpha, \alpha' \in [0, 1] \) with \( \alpha' > \alpha \). Then

\[
k^l(\alpha', p, u_w) \geq k^l(\alpha, p, u_w) \quad \text{and} \quad k^h(\alpha', p, u_w) \geq k^h(\alpha, p, u_w).
\]

**Proof.** To prove the proposition, we first show that if under low correlation, we prefer truncating less to more, then under high correlation we definitely prefer truncating less to more.

First, observe that

\[
\mathbb{E}_{p^{\alpha'}}[v(k, \tilde{P}_{-w})] = \sum_{P_{-w}} p^{\alpha'}(P_{-w})v(k, P_{-w})
\]

\[
= \sum_{P_{-w}} [(1 - \alpha')p(P_{-w}) + \alpha'p^C(P_{-w})] v(k, P_{-w})
\]

\[
= \sum_{P_{-w}} [(1 - \alpha')p(P_{-w}) + \alpha \frac{1-\alpha'}{1-\alpha}p^C(P_{-w}) + \frac{\alpha'}{1-\alpha}p^C(P_{-w})] v(k, P_{-w})
\]

\[
= \left(1 - \frac{\alpha'}{1-\alpha}\right) \mathbb{E}_{p^\alpha}[v(k, \tilde{P}_{-w})] + \left(\frac{\alpha'}{1-\alpha}\right) \mathbb{E}_{p^C}[v(k, \tilde{P}_{-w})]
\]

Suppose that for \( k_2, k_1 \in \{1, \ldots, N\} \) with \( k_2 > k_1 \), we have

\[
\mathbb{E}_{p^{\alpha'}}[v(k_2, \tilde{P}_{-w})] \geq \mathbb{E}_{p^{\alpha'}}[v(k_1, \tilde{P}_{-w})]. \tag{8.0.3}
\]

Then since

\[
\mathbb{E}_{p^C}[v(k_2, \tilde{P}_{-w})] \geq \mathbb{E}_{p^C}[v(k_1, \tilde{P}_{-w})],
\]

we must have

\[
\mathbb{E}_{p^{\alpha'}}[v(k_2, \tilde{P}_{-w})] \geq \mathbb{E}_{p^{\alpha'}}[v(k_1, \tilde{P}_{-w})]. \tag{8.0.4}
\]

If the inequality in (8.0.3) is strict, then so too is the inequality in (8.0.4).

We can now use this payoff comparative static to sort optimal truncation points as follows.

By definition, \( k_l(\alpha, p, u_w) \) satisfies

\[
\mathbb{E}_{p^\alpha}[v(k_l(\alpha, p, u_w), \tilde{P}_{-w})] > \mathbb{E}_{p^\alpha}[v(k, \tilde{P}_{-w})] \quad \forall \ k < k_l(\alpha, p, u_w).
\]

From (8.0.4), we must then have

\[
\mathbb{E}_{p^{\alpha'}}[v(k_l(\alpha, p, u_w), \tilde{P}_{-w})] > \mathbb{E}_{p^{\alpha'}}[v(k, \tilde{P}_{-w})] \quad \forall \ k < k_l(\alpha, p, u_w),
\]

34
so that \( k_l(\alpha', p, u_w) \geq k_l(\alpha, p, u_w) \).

Similarly, \( k_h(\alpha, p, u_w) \) satisfies
\[
E_{p^{\alpha'}} [v(k_h(\alpha, p, u_w), \hat{P}_w)] \geq E_{p^{\alpha}} [v(k, \hat{P}_w)] \quad \forall \; k < k_h(\alpha, p, u_w).
\]

From (8.0.4), we must then have
\[
E_{p^{\alpha'}} [v(k_h(\alpha, p, u_w), \hat{P}_w)] \geq E_{p^{\alpha}} [v(k, \hat{P}_w)] \quad \forall \; k < k_h(\alpha, p, u_w),
\]
so that \( k_h(\alpha', p, u_w) \geq k_h(\alpha, p, u_w) \). \( \square \)

References


