DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

CLEARINGHOUSES FOR TWO-SIDED MATCHING: AN EXPERIMENTAL STUDY

Federico Echenique
California Institute of Technology

Alistair J. Wilson
New York University

Leeat Yariv
California Institute of Technology

SOCIAL SCIENCE WORKING PAPER 1315

December 2009

Clearinghouses for Two-Sided Matching: An Experimental Study*

Federico Echenique† Alistair J. Wilson‡ Leeat Yariv§.

December 18, 2009

*We are thankful to the following for their help and advice: Clayton Featherstone, Guillaume Fréchette, Andy Schotter, and Emanuel Vespa. We are very grateful to Walter Yuan, who programmed our experimental interface and helped us in various ways with the implementation. We gratefully acknowledge financial support from the Gordon and Betty Moore Foundation, and from the Lee Center at Caltech.

†Division of Humanities and Social Sciences, California Institute of Technology, fede@hss.caltech.edu

‡Department of Economics, New York University, alistair@nyu.edu.

§Division of Humanities and Social Sciences, California Institute of Technology, lyariv@hss.caltech.edu

Abstract

We study the performance of two-sided matching clearinghouses in the laboratory. Our experimental design mimics the Gale-Shapley (1962) mechanism, utilized to match hospitals and interns, schools and pupils, etc., with an array of preference profiles. Several insights come out of our analysis. First, only 48% of the observed match outcomes are fully stable. Furthermore, among those markets ending at a stable outcome, a large majority culminates in the best stable matching for the receiving-side. Second, contrary to the theory, participants on the receiving-side of the algorithm rarely truncate their true preferences. In fact, it is the proposers who do not make offers in order of their preference, frequently skipping potential partners. Third, market characteristics affect behavior and outcomes: both the cardinal representation and the span of the core influence whether outcomes are stable or close to stable, as well as the number of turns it takes markets to converge to the final outcome.

1 Introduction

1.1 Overview

Many two-sided matching markets function through centralized clearinghouses: medical residents to hospitals, rabbis to congregations, high-school students to schools, commissioned officers to military posts, college students to dorms, etc. All use highly structured procedures to generate matches. Clearinghouses have the advantage that they can be designed to implement desirable outcomes at the market level. In particular, many of the extant clearinghouses aim at implementing *stable* outcomes.[1] In this paper we inspect such clearinghouses using laboratory experiments. Our goal is to gain insights into when, in fact, stable outcomes emerge, as well as on how market participants respond to incentives within such markets.

Consider the example of the National Resident Matching Program (NRMP) in the United States, which has been operating since 1952. In 2009, a total of 36,000 physicians and 4,300 hospitals (herein referred to as workers and firms, respectively) participated in the process, each submitting their preference lists to the central clearinghouse. In the NRMP, each physician submits preferences over hospitals' job openings, and each hospital submits their preferences for the physicians applying for each opening. Matches are then computed by following a version of the *DA* algorithm(first described in Gale and Shapley 1962). Namely, the clearinghouse emulates a process through which workers make employment offers in order of their *submitted* preferences, and firms, at each stage, hold on to the best offers they have received for each vacancy, as determined by their *submitted* preferences. At some point, all workers have an offer held or find themselves in a situation where all the firms they would consider working for have rejected them. At

[1] Stable matchings are characterized by two conditions: i) no agent prefers to remain by themselves over their allocated match; and ii) no two agents prefer to match to one another over their allotted partners.

this terminating point, currently held offers are converted to matches, and the market's outcome is determined.

This type of clearinghouse has appealing theoretical properties. Whenever all market participants submit their preferences truthfully, the generated matching is stable, with every resulting match being the most-preferred stable partner for the proposing side. In terms of incentives, workers (the proposing side) have no motive to misrepresent their preferences—it is a weakly dominant strategy for each to reveal their true preferences. However, firms (the side that receives proposals) have an incentive to submit "truncated" preferences; that is, list the workers in the true order of preference, but shorten their lists, declaring some of the workers unacceptable. In fact, firms are able to obtain their most preferred stable matching by each truncating workers ranked below their most-preferred stable partner.[2]

In order to determine outcomes and responses to incentives under such clearinghouses, field data can be very useful, but have inherent drawbacks: true preferences are not observed, interactions between participants outside of the clearinghouse are difficult to gauge, and the information subjects have regarding others' preferences is unclear. This is why experiments, which allow for a fully controlled environment, are particularly valuable in gaining complementary insights into the functioning of matching clearinghouses.

In this paper we report results from an array of experiments in which subjects have to go through the steps of the *deferred-acceptance (DA)* algorithm. The two sides of the market alternate—with workers proposing a match to firms, and firms accepting at most one held offer among those proposals received and any held worker from the previous turn—with the process repeating until there are no new proposals. A worker is regarded as truthful if they propose to firms in the order of their preference. A firm

[2]Truthful revelation by all participants when there are multiple stable matchings is generally not an equilibrium of the full-information game, and the identified truncation strategy is a strong equilibrium for firms, cf. Roth and Sotomayor (1990, Theorems 4.6 and 4.17).

truncates if it rejects proposals that are actually preferred to the status quo, not matching with any worker at all.

Our markets are comprised of eight subjects on each side, and participants have complete information on everyone's preferences.[3] Our subjects participate in a variety of markets, varying over multiple characteristics: in *market complexity*, as captured by the number of stable matchings (either one, two, or four), and the number of rounds required for the DA algorithm to converge under truth-telling; in *the incentives to manipulate or report untruthfully*, captured through the size of the core (i.e., the number of stable matchings) and the degree of manipulation required by the receiving side to produce their preferred matching; and, finally, in the markets' *cardinal representation of preferences*, controlled by the payoff differences between matches for any particular subject.

There are several findings that come out of our analysis: First, stable matchings are not the norm, *as only about one-half of our markets generate a stable matching.* Of these, when markets have multiple stable outcomes, *approximately* 29% *generate the proposer-best stable matching.*

Second, *market characteristics are important in determining outcomes.* For instance, both the cardinal representation and the span of the core have a significant effect on whether outcomes are stable, the overall distance of the observed outcomes from the core, and the number of turns it takes markets to converge to an outcome.

Third, individual behavior diverges in a consistent manner from the theoretical predictions. In particular, *workers are not truthful and firms do not optimally truncate.* Specifically, we find that workers "skip down" their preference lists. For example, a worker might propose to their third-best firm, skipping the favorite and second-favorite firm; then, if rejected by their third favorite, the worker might skip down to the fifth; and so on. This behav-

[3]Having complete information serves as a natural first step in understanding participants' responses to incentives, void of issues pertaining to belief updating and learning that would arise in environments with incomplete information.

ior is clearly at odds with the theory. For firms, we do not detect optimal truncations: Firms typically choose the best alternative out of any set of proposals. They do not strategically reject proposals, as the theory predicts, but instead reject fairly constantly offers from workers in the lower part of their preference order, with little reaction to market structure.

Our analysis suggests that *proposers are sophisticated in their "skipping" behavior*. Workers consider the position they themselves are held in their target firm's preferences when making a proposal decision. For example, if a worker's first-best firm ranks them as largely undesirable, that worker is less likely to propose to them. Furthermore, our results suggest that a behavioral notion of risk might play an important role in workers' choices.

From a methodological perspective, our experimental design differs from the few existing experiments on matching mechanisms (which we review in detail). We require subjects to "produce" the matching by going through the steps of the algorithm. We could, instead, have asked subjects to report their preferences, and then produce the matching statically via the DA algorithm. This alternative design would have the advantage that it mirrors the actual markets, where workers and firms submit preference lists. We argue, however, that the game induced by our design is essentially equivalent to the static procedure, and that our own design has some methodological advantage.

We show that, under a rather natural restriction on the set of strategies, the mechanism we implement and the standard DA mechanism are isomorphic. If workers' strategies depend only on the set of possible proposals, and firms' only on the proposals received, there are standard, simple conditions under which the strategies are equivalent to submitting a preference. In addition, the submitted preferences give rise to the same outcome as the preference-revelation mechanism would.

Our design has two important advantages. The first advantage is that our design enhances subjects' understanding of their strategic incentives. The DA algorithm is difficult to understand, and our concern is that experimental subjects will find it very difficult to map reported preferences into matching outcomes. As a result, agents may behave truthfully simply because it is an easy criterion for how to act in the experiment. Our design is easier to understand since subjects make the proposals, and decide which to accept in a piecemeal fashion. The strategic issues are still complicated, but they are complicated solely due to the strategic uncertainty faced by the players. The mapping from actions to outcomes is clearer. The second advantage is related to experimenter demand. If we first give the subjects a preference list, and then proceed to ask them what their preferences are, we might contaminate our experiment. The subjects could infer that the experimenters are interested in whether they report truthfully or not. This may trigger affective reactions regarding lying, as well as attempts to comply with what they perceive as the experimenters' expectation.

1.2 Related Literature

Laboratory experiments focusing on two-sided matching have been relatively scarce. In terms of design, Haruvy and Ünver (2007) is the closest to ours. Haruvy and Ünver's main motivation is to study repeated interactions between firms and workers, and the predictive power of the DA algorithm with regard to these situations; it is not, like our own experiment, designed to examine the strategic behavior *within* the DA algorithm. They run a version of the sequential "offers by workers, responses by firms" game in 4×4 markets. There are several important differences with respect to our design: i) workers are allowed to repeat offers, thereby creating a larger wedge between the game played and the DA algorithm; ii) workers and firms are paid for the results in every turn of the sequence (not only the ultimate matching); and, iii) their design incorporates automated respondents in some sessions, robots that automatically accept the best offer, as well as incomplete information in preferences. They find a substantial number of repeat offers (that most centralized clearinghouses do not allow) and significantly less "skipping" by the proposers than we find.

Harrison and McCabe (1992) implement the preference-revelation DA mechanism in one 3×3 (3 workers and 3 firms) market and one 4×4 market. Similar to our own, their design entails common knowledge of all market participants' preferences. However, unlike our design, Harrison and McCabe have subjects repeat play of the market multiple times, and replace many market roles with computers programmed to play truthfully. In their environment, outcomes are more in line with the theoretical predictions than ours. However, they do observe a small degree of "skipping," as well as firms failing to successfully manipulate the mechanism.

A number of experimental papers seek to compare the different centralized mechanisms that are used in practice. Chen and Sönmez (2006) compare DA with the Boston and the Top Trading Cycle mechanisms. Their focus is on the school-choice problem, hence they have strategic agents on only one side of the market. Chen and Sönmez implement a preference-revelation design, in which agents know their own preferences, but not the preferences of the other participants. In terms of manipulation, they find that proposers (workers) do not misrepresent their preferences. Featherstone and Niederle (2008) also compare DA with the Boston mechanism. They find results that are similar to one of our own, that proposers do not necessarily follow their dominant strategy to truthfully reveal, and skip highly-ranked potential matches that are very unlikely to accept them. However, they attribute the effect to weak market-specific incentives for the skipping player; our own experiments indicate that this effect is more systematic. Featherstone and Mayefsky (2010) test the DA and Boston mechanisms under incomplete information on the preferences of others, implementing the proposing side as mechanical truth-tellers. They find some evidence for limited manipulation of the submitted preferences by receivers. Their results point to DA being harder to manipulate than Boston, as the degree of truncations is smaller than the theoretically identified optimal amount. Trading Cycles mechanisms in the laboratory under incomplete information. Automating the proposing side of the market to reveal truthfully, they find greater manipulation by subjects in the Boston mechanism, but that the Top Trading Cycle mechanism dominates the other two procedures when assessed over both truth-telling and the efficiency of matches.

Finally, a few papers experimentally examine decentralized markets. Echenique, Katz, and Yariv (2010) examine behavior in decentralized markets and find that outcomes are in most cases stable. Their study focuses on selection, and they find that the median stable matching tends to emerge. Kagel and Roth (2000) analyze the transition from decentralized matching to centralized clearinghouses, when the market features lead to inefficient matching through unraveling. Nalbantian and Schotter (1995) analyze several procedures for matching with transferable utility, decentralized matching among them, where agents are informed of their own payoffs, but not anyone else's.

2 Dynamic Design of Centralized Matching

Our paper is an experimental study of strategic behavior in centralized matching markets. To motivate our approach, consider the following game, described in Roth and Sotomayor (1990, page 79):

1. Actions in the market are organized in stages. Each stage is divided into two periods. Within each period, each worker and firm must make decisions without knowing the decisions of other workers and firms in that period.

2. In the first period of the first stage, each worker may make at most one proposal to any firm he chooses (and is also free to make no proposal). Proposals can only be made by workers.

3. In the second period of the first stage, each firm that has received proposals may freely reject any or all of them immediately. A firm may

also keep at most one worker "engaged" by not rejecting their proposal.

4. In the first period of any stage, any worker who was rejected in the preceding stage may make at most one proposal to any firm he has not previously proposed to (and been rejected by). In the second period, each firm may reject any or all of these proposals, including that of any worker who proposed in an earlier stage and was kept engaged. A firm may keep at most one worker engaged by not rejecting his proposal.

5. If, at the beginning of any stage, no worker makes a proposal, then the market ends, and each worker is matched to the firm they are currently engaged with. Workers who are not engaged with any firms, and firms who are not engaged with any workers, remain unmatched.[4]

The game imitates the steps in the DA algorithm (see Section 3 for a description). In actual centralized matching markets, workers and firms submit preferences to a central matching authority (as is the case in the National Residents Matching Program). The authority then uses the submitted preferences as inputs to the DA algorithm, instituting the resulting matching. In contrast, in the game above, workers and firms decide on proposals at each step; a matching emerges sequentially through their actions.

Roth and Sotomayor present the game as an introduction to strategic issues in matching. There is a notion of "*straightforward behavior*" in the game. A worker behaves straightfowardly if their proposals go from the most-preferred firm to the second-most-preferred firm, then to the third-most preferred, and so on. A firm behaves straightforwardly if at each step it accepts the most-preferred proposal. Straightfoward behavior corresponds naturally to truthful behavior in the centralized mechanism. The strategic issue is whether agents will behave straightforwardly (or truthfully).

[4]The description is literally Roth and Sotomayor's, with the only difference that we recast men/women as workers/firms.

We directly adopt the above game within our experimental design (detailed in Section 4). Roth and Sotomayor's use of this game is pedagogical, our reasons are similar. We want subjects to grasp the relation between their actions and the resulting outcomes. Subjects best understand the incentives they face when directly experiencing the steps involved in the matching process. In contrast, with the preference-revelation game, subjects need to map each declared profile into an outcome of the algorithm: This map is complicated, and it is difficult to ensure that laboratory subjects have a clear understanding of the DA algorithm in the lab.

A second reason for adopting the above game is related to experimenter demand: If we *provide* subjects with a preference ranking, and then proceed to ask them to *submit* a preference ranking, we worry that subjects will infer the experimenters' motives. They may, as a result, act with a different motivation from that we sought to induce. By asking them to present a preference we present a cue that the experiment is assessing whether they will behave truthfully or not. This cue may trigger behavior related to the consequences of lying, and/or complying with the experimenters' expectations. The resulting experimenter-demand effects could act in either direction, and are inseparable from the behavior we desire to assess.

Theoretically, under some plausible restrictions on behavior, the above game and the direct revelation game induced by the DA algorithm are effectively equivalent. In what follows, we describe some of the theoretical background for our investigation as well as the formal requirements for this equivalence.

3 Theoretical Background

3.1 The Underlying Model

Let W and F be disjoint, finite sets. We call the elements of W workers and the elements of F firms. The sets W and F can represent medical residents and hospitals, men and women, parents and schools, etc., that are to be matched to one another in the market. A *matching* is a function $\mu : W \cup F \rightarrow W \cup F$ such that for all $w \in W$ and $f \in F$,

1. $\mu(f) \in W \cup \{f\}$,

2. $\mu(w) \in F \cup \{w\}$,

3. $w = \mu(f)$ if and only if $f = \mu(w)$,

where the notation $\mu(a) = a$ means that participant a is unmatched under μ and $f = \mu(w)$ denotes that w and f are matched under μ. We denote the set of all possible matchings, given the sets W and F, as $\mathcal{M}$.

A *preference relation* is a linear, transitive, and antisymmetric binary relation (all preferences are strict, no worker or firm is indifferent over two distinct partners). A preference relation for a worker $w \in W$, denoted P_w, is understood to be over the set $F \cup \{w\}$ Similarly, for a firm $f \in F$, P_f denotes a preference relation over $W \cup \{f\}$. If any participant a prefers remaining unmatched to being matched with another participant $a\prime$ ($aP_a a\prime$), we will say that the match $\mu(a) = a\prime$ is *not individually rational* ,or *unacceptable*, for a. We will assume that each worker (firm) prefers every firm (worker) to remaining unmatched.[5]

A *preference profile* is a list P of preference relations for workers and firms:

$$P = \left((P_w)_{w \in W}, (P_f)_{f \in F} \right).$$

As is standard, for $i \in W \cup F$, we denote by P_{-i} the profile of preferences for all agents but i. Let $\mathcal{P}$ be the set of all possible preference profiles, and for an agent $i \in W \cup F$, let $\mathcal{P}_i$ denote the set of possible preferences for i.

We assume that preferences are *strict*. Denote by R_w the weak version of P_w. So $f' R_w f$ if $f' = f$ or $f' P_w f$. The definition of R_f is analogous.

[5]This fits our experimental design where remaining unmatched is the worst outcome.

Fix a preference profile P. We say that a pair (w, f) *blocks* the matching μ if $fP_w\mu(w)$, and $wP_f\mu(f)$. A matching is *stable* if it is individually rational and there is no pair that blocks it.[6] Finally, denote by $S(P)$ the set of all stable matchings.

3.2 Centralized Mechanisms

A *mechanism* is a function $\phi : \mathcal{P} \rightarrow \mathcal{M}$ that assigns a matching to each preference profile. A mechanism is *stable* if $\phi(P) \in S(P)$ for all $P \in \mathcal{P}$.

Gale and Shapley (1962) proved that every preference profile admits a stable matching, and provided the following algorithm to identify one:

Deferred-Acceptance Algorithm.

STEP 0 *The set A_0 of* active *workers consists of all the workers. All firms have no tentative partners.*

For $k = 1, 2, \ldots$, repeat the following until A_k is empty:

STEP k *:*

- *Each worker w in A_{k-1} proposes to the highest-ranked firm according to P_w, across all of the firms w has not proposed to in previous steps of the algorithm.*
- *Each firm f chooses the best partner (according to P_f), out of the set of workers that proposed to f in step k, and f's tentative match from step $k-1$. This choice is f's new tentative match; reject all other proposals.*
- *The set A_k formed from the set of all active workers rejected in this step: either their proposal to a firm was rejected, or they were tentatively matched in step $k-1$, and rejected in favor of a new proposal.*

[6] Since we assume that partners are always acceptable, any matching is individually rational under the true preferences.

Stop

The tentative matching at the end of the last step is the output matching.

Theorem (Gale-Shapley). *$S(P)$ is non empty, and there are two matchings μ_W and μ_F in $S(P)$ such that, for all w, f, and $\mu \in S(P)$,*

$$\mu_W(w)\ R_w\ \mu(w)\ R_w\ \mu_F(w)$$
$$\mu_F(f)\ R_f\ \mu(f)\ R_f\ \mu_W(f).$$

The matching μ_W is called *worker-best* while μ_F is called *firm-best*. Beyond its theoretical role in establishing existence, the DA algorithm is often used in centralized markets. For instance, the National Resident Matching Program uses a close modification of the DA algorithm (where physicians serve as proposers, or workers, and hospitals as the receivers, or firms).

A mechanism ϕ defines a direct revelation game: the normal-form game where the agents in $W \cup F$ simultaneously report their preferences, so the strategy space of agent i is $\mathcal{P}_i$, and the outcome of a profile P is given by $\phi(P)$. Denote by ϕ_{DA} the mechanism defined by the DA procedure.

For an agent $i \in W \cup F$, truth-telling is a *weakly dominant strategy* if, for any preference profile P_i' different from the true preferences P_i, and any profile $\tilde{P}_{-i}$ of all agents but i, it is true that

$$\phi(P_i, \tilde{P}_{-i})(i)\ R_i\ \phi(P_i', \tilde{P}_{-i})(i)$$

A mechanism is *strategy proof* if truth-telling is weakly dominant for all agents. As it turns out, we have the following (see Roth and Sotomayor 1990):

Theorem (Strategy Proofness in Stable Mechanisms).

1. *In ϕ_{DA}, truth-telling is weakly dominant for workers.*

2. *No stable mechanism is strategy proof.*

Fix a preference profile P, and suppose that all workers w truthfully choose P_w as their strategy in the direct-revelation game. We consider the *induced* game among firms, where firms simultaneously choose a preference profile $\tilde{P}_f \in \mathcal{P}_f$. A *Nash equilibrium* of the induced game is a profile of preferences $(P'_f)_{f \in F}$ such that,

$$\phi((P_w)_{w \in W}, (P'_{f'})_{f' \in F})(f) R_f \phi((P_w)_{w \in W}, \tilde{P}_f, (P'_{f'})_{f' \in F \setminus f})(f)$$

for all $f \in F$ and $\tilde{P}_f \in \mathcal{P}_f$.

The following result is well known (again, see Roth and Sotomayor 1990):

Theorem (Equilibrium Outcomes in Stable Mechanisms). *Consider any stable mechanism implementing the worker-optimal stable matching for any reported preferences. In the game induced from truth-telling by the workers, the set of Nash equilibrium outcomes coincides with the set of stable matchings.*

3.3 Outcome Equivalence

We present the main intuition behind the equivalence between our game and the DA direct-revelation game. The main difference arises because agents in our game can condition their actions on the sequencing of events. We impose several simple restrictions on strategies (which we later scrutinize using our experimental data) that make the differences between the two games irrelevant. In Appendix A.2 we present a formal analysis of this comparison.

Heuristically, a strategy for a particular proposer maps a sequence of past proposals (with their corresponding outcomes) into a current proposal. We first restrict strategies to only depend on available proposals. For example, if w_1 can only propose to f_1 or f_2, his choice should be independent of the precise sequence of (rejected) proposals that ended with f_1 and f_2 as the remaining choices. While this restriction seems realistic, it is easy to write down examples that violate it. For example, w_1 may choose f_1 over f_2 when he proposed to f_3 once and was rejected. But he may choose f_2 instead if his proposal to f_3 was initially accepted, and rejected several rounds later.

The second restriction is standard in choice theory. A strategy for a worker is a mapping from sets of available firms into a proposal; for each set F' of firms, either some $f \in F'$ is proposed to or no proposal is made. The strategy is then a choice function that can take empty-set values. Under standard conditions from choice theory (such as the congruence axiom of Richter (1966)), we can represent such a strategy with a preference relation.[7]

We make analogous assumptions on firms' behavior. A firm's strategy is a decision on which proposal to accept, given any set of proposals made by the active workers, and any worker whose proposal the firm holds. Again, the restrictions we impose are of two types. First, strategies cannot depend on histories *per se*. Second, strategies obey certain minimal consistency requirements across time, so that they can be represented as preference relations.

We show that a profile of strategies, once represented as a profile of preference relations, generates the same outcome as the one that would have been generated in the preference-revelation game ϕ_{DA}. Hence, the incentives faced by workers and firms in both games are the same.

4 Experimental Design

Our experimental sessions implement a sequence of markets involving two sides, which we neutrally term *colors* and *foods*. In each round, subjects are randomly assigned a role (*red*, *blue*, etc., if a color; *apple*, *banana*, etc., if a food). There are 8 roles in each group, totaling 16 subjects in a market. Subjects can match with at most one subject from the opposite group, and derive different monetary payoffs from each match. Subjects are fully informed on

[7]Namely, we can find a preference ranking P_w such that for any set of available firms F', if there is some acceptable firm under P_w, the one that is the most preferred according to P_w in F' is proposed to. The restrictions are reminiscent of the weak axiom of revealed choice, assuring consistency of observed behavior.

all the potential payoffs for every possible match in the market by a table on their computer screens, as depicted in Figure 1, where the first number in each cell is the corresponding color's payoff in cents, the second number the corresponding food's payoff.[8] Remaining unmatched results in a payoff of 0. For the remainder of the paper, we replace the neutral food/color labels with the worker/firm frame used previously.

	Apple	**Banana**	**Kiwi**	**Cherry**	**Mango**	**Pear**	**Grape**	**Peach**
Red	(360,125)	(210,175)	(60,375)	(110,425)	(160,475)	(10,425)	(310,475)	(260,325)
Blue	(160,475)	(360,125)	(260,275)	(210,475)	(60,225)	(110,175)	(10,225)	(310,475)
Green	(260,375)	(110,325)	(360,125)	(310,325)	(210,425)	(60,475)	(10,375)	(160,375)
Magenta	(310,325)	(160,425)	(110,225)	(360,125)	(260,275)	(10,275)	(60,425)	(210,175)
Yellow	(260,275)	(310,275)	(160,425)	(60,175)	(360,125)	(10,375)	(210,275)	(110,225)
Pink	(10,425)	(210,375)	(60,325)	(160,375)	(310,375)	(360,125)	(110,175)	(260,425)
Cyan	(110,225)	(260,225)	(160,175)	(60,275)	(210,325)	(310,325)	(360,125)	(10,275)
Orange	(260,175)	(210,475)	(310,475)	(10,225)	(160,175)	(110,225)	(60,325)	(360,125)

Figure 1: Example Matching Payoffs

In each round, subjects interact within a protocol that mimics the DA algorithm with workers proposing—the Roth-Sotomayor game discussed in Section 2. Subjects on differing sides of the market take turns, each composed of two stages. In the first stage, each worker can make (at most) one proposal to a firm. In the second stage, each firm that receives proposals, can hold on to at most one offer rejecting all others. Then, in the third stage, workers who do not have a held proposal can again make offers. In the fourth stage, firms that receive new offers chose at most one offer to hold, rejecting all others. And so on.[9] In each proposing stage workers have 30 seconds to

[8]Full instructions are available at:
http://sites.google.com/site/galeshapley/

[9]The first of these two stages consist the first turn, the next two the second turn, and so on. Note that if a firm who held an offer in stage k decides to hold a new offer in stage $k+2$, the worker held in stage k is automatically rejected (and free to make an offer in

decide whether and to whom they will propose. Firms have 25 seconds to respond to offers (with a failure to respond to any proposal within the time limit interpreted as a rejection of all proposals).

In order to induce the Roth-Sotomayor game, we impose a restriction that workers may not make repeat proposals. So, after proposing to and receiving a rejection from a particular firm, a worker cannot make any additional proposals to that same firm. An experimental round ends whenever there are no new proposals within a worker-proposing stage.[10] As the round progresses, subjects only observe their own interactions, they do not observe any proposal/rejection in which they are not directly involved. For instance, a subject playing the role of a worker knows the order and stages in which they proposed, and similarly the stages they were rejected in. They do not, however, observe who else proposed to a particular firm at any time, who the firms rejected, *etc.* Similarly, firms only observe proposals made to them, and their own hold/reject behavior. When the round ends, each held proposal becomes a match, and the corresponding firm and worker receive their resulting payoffs (according to the common match-payoff table).

Each experimental session is composed of 2 unpaid practice rounds followed by 15 paid rounds. Each round uses match payoffs corresponding to one of 6 preference profiles for the market participants.[11] A detailed summary of the markets used in the sessions appears in Table 1. The number of times each market was run appears under the N-column.[12]We designed the

stage $k+3$).

[10]This end condition can have three potential causes: i) All the workers have held proposals and therefore none is available to make an offer; ii) All workers without held proposals have no firms to which they have not made a proposal, so no unheld firm can make an offer; iii) Some workers without a held proposal choose not to make a proposal in this stage, and the remaining workers have no new proposals to make.

[11]Subjects have no knowledge on the sequence of markets, and only observe a particular round's payoff table at the start of that round. Additionally, rows and columns are randomly permuted so as to disguise obvious patterns like that in the main diagonal of Figure 1.

[12]The full set of markets we used is available at:
http://sites.google.com/site/galeshapley/

Table 1: Markets Used

Market	Arrangement	Stable	Truncation		Core Span		Avg. Payoff		DA	N
		Matchings	F-Best	Unstable	W	F	W	F	Turns	
(I)	W-F	1	-	1	-	-	\$2.50	\$2.50	8	4
(II)	W-F	1	-	7	-	-	\$2.50	\$3.48	8	8
	F-W	1	-	1	-	-	\$3.48	\$2.50	2	4
(III)	W-F	2	5	8	1.00	1.75	\$2.85	\$2.73	4	4
	W-F Dev 1	1	-	5	-	-	\$2.85	\$2.79	4	8
	W-F Dev 2	1	-	8	-	-	\$2.60	\$3.60	8	8
	F-W	2	4	5	1.75	1.00	\$3.60	\$2.35	1	4
(IV)	W-F	2	1	4	1.00	5.13	\$3.60	\$1.25	1	8
	F-W	2	7	8	5.13	1.00	\$3.81	\$3.10	11	8
(V)[a]	W-F	2	1	3	1.75	2.00	\$3.10	\$2.00	5	28
	W-F Dev 1	1	-	3	-	-	\$2.53	\$2.85	15	8
	F-W	2	4	5	2.00	1.75	\$3.00	\$2.22	6	16
	F-W Dev 1	1	-	5	-	-	\$2.85	\$2.53	6	8
(VI)	W-F	4	7	7	1.00	0.75	\$3.35	\$3.10	3	4
All		1.67	1.83	4.77	1.21	1.23	\$3.04	\$2.64	6.1	**120**

[a]This market was run with marginal payoffs of 20¢ and 50¢ for both the W-F and F-W arrangements.

markets to vary over several dimensions, detailed as follows:

Market "Complexity." All but one of our markets have either a unique stable outcome or two disjoint, stable matchings.[13] We designed the markets to vary in the number of turns[14] required for the DA algorithm to converge under truth-telling, as well as the sensitivity of outcomes to truncation by firms (the receiving side of the market). The latter is captured in two ways: First, we calculate the number of workers that firms must truncate in order to achieve their most-preferred stable matching, assuming that workers behave truthfully and firms truncate jointly.[15] Second, we calculate the minimal number of workers firms must jointly (and uniformly) truncate to generate an unmatched partner. This measure captures the sensitivity of stable matchings to truncation.

Cardinal Representation. Match payoffs in cents are constructed from each market's ordinal preference profile. The marginal decrease between an agent's n-th and $(n+1)$-th favorite partners is fixed at 50¢ in the majority of markets. In order to gauge the effects of cardinal representations within our markets, we use marginal decreases of just 20¢ in one of the baseline markets, Market (V).[16] The average payment across agents (and across stable

[13] One market, (VI), was designed to provide some insight on the effects of market size on outcomes and was composed of two 4×4 embedded markets (so that any agent within a submarket preferred to match with anyone from that submarket over anyone from the other). Each of the sub-markets entailed two stable matchings, leading to 4 market-wide stable matchings. However, each participant in this market had at most two potential stable match partners.

[14] Defining a turn as the two consecutive stages, a proposal from workers, and the stage directly following it in which firms respond.

[15] Formally, this number is calculated as follows. Suppose for each firm i the firm-optimal stable matching assigns a worker ranked as $r_i \in \{1, .., 8\}$ ($r_i = 1$ corresponding to firm i's most preferred worker). We compute the minimal number $t \in \{1, ..., 8\}$ such that if each agent i truncates the bottom $\min\{t, 8 - r_i\}$ workers, the worker-optimal stable matching would be implemented under the worker-proposing DA mechanism if workers behave truthfully. Small truncation values t correspond to smaller required (joint) deviations from truthful revelation to implement the firm-optimal stable matching.

[16] In theory, payoff representations of preferences do not affect incentives in the complete information DA mechanism, nor do they matter for the set of stable matchings.

matchings when there were two) is between \$2.50 and \$3.20.[17] The average payoff, given truthful revelation on both sides of the market, is separately given in the *Average Payoff* column for workers and firms. Given truthful behavior, workers should earn an extra 40¢ per round—though across market variations this difference varies from \$1.00 less than firms through to \$2.35 more.

Incentives to Manipulate and Core Size. Three markets with multiple stable matchings, Markets (III), (IV), and (V), are duplicated with the roles of workers and firms reversed.[18] This reversal provides information on the effects from the differing incentives of proposers and receivers. The reversed markets are indicated in the *Arrangement* column of Table 1, where W-F is the original setting and F-W is the market in which roles are reversed. In addition, we alter two of these markets, (III) and (V), by making minor modifications to preferences—changing the ranking of just one participant so as to produce a blocking pair—and, thereby induce a similar market with a unique stable matching. For Market (III), two different modifications are introduced to make the worker-optimal and firm-optimal stable matching from the original market the unique stable outcome (with resulting markets denoted by W-F Dev 1 and W-F Dev 2, respectively). In a similar fashion, for Market (V) W-F Dev 1, we introduce a small change to the W-F market to focus on the original's worker-optimal stable matching. The same deviated market has the roles reversed in F-W Dev 1, to achieve the firm-optimal stable outcome as the unique matching when compared against the F-W orientation.

Markets also differ in the size of the core. For each worker we calculate the distance in rank position between their best and worst stable partner,

[17]For each profile of preferences, we chose payoffs to minimize this average under two constraints: i) this average is above \$2.50; and ii) each subject's payoffs from any match exceeds 5¢.

[18]Keeping any particular subject i's preference profile fixed, we switch their market side, firms to workers, and vice versa. This is equivalent to retaining labels and using the firm-proposing DA algorithm.

and average these values across workers. We call the resulting number the workers' "core span." The analogous calculation is also given for the firms. Core spans vary between 0 (when the stable matching is unique) and 5.13.[19] A large core span for one side corresponds to greater incentives for achieving that side's optimal stable partner.

Our sessions were all run at the California Social Science Experimental Laboratory (CASSEL), and implemented using a variation of the multi-stage software. In total, 128 subjects were recruited; all were UCLA undergraduates, and each subject participated in just one session. The average payment per subject was \$41 (with a standard deviation of \$5), combined with a \$5 show-up payment.

5 Aggregate Outcomes

The main indicators for the aggregate results across our experimental markets appear in Table 2. There are several layers we go through in our analysis below. First, we show that a significant fraction of our markets did not end up in a stable matching, and that the matchings they did end up in are suggestive of workers rather than firms behaving in an untruthful fashion. Second, when markets entail multiple stable matchings, we inspect which ones get selected. In line with subjects not behaving truthfully, we see that a large majority of markets end up at or close to the firm-best stable matching. Last, we study the tangible outcomes subjects experienced in our markets, namely time spent and payoffs earned.

[19]When there are two stable outcomes, the two matchings were designed to be disjoint—that is, every worker and firm's best and worst stable partner are different—so the core span is at least 1 in these markets.

Table 2: Aggregate Outcomes

Market	Arrangement	Stable	W-Best (Closer)	Distance	Unmatched	Δ Payoff		Turns	N
						W	F		
(I)	W-F	25.0%	-	0.71	6.3%	-3.3¢	3.3¢	9.3	4
(II)	W-F	50.0%	-	0.92	1.6%	0.0¢	-17.5¢	8.9	8
	F-W	25.0%	-	1.41	9.4%	-22.4¢	8.6¢	9.0	4
(III)	W-F	50.0%	50.0% (50%)	0.78	3.1%	-22.6¢	-53.2¢	7.3	4
	W-F Dev 1	37.5%	-	1.03	1.6%	-8.73¢	15.1¢	6.0	8
	W-F Dev 2	87.5%	-	0.69	0.0%	0.0¢	-5.5¢	8.4	8
	F-W	50.0%	50.0% (25%)	0.84	6.3%	-58.3¢	-11.7¢	8.0	4
(IV)	W-F	62.5%	0.0% (0%)	1.79	6.3%	-62.5¢	-4.1¢	4.0	8
	F-W	62.5%	100.0% (100%)	1.20	0.0%	-22.66¢	-58.6¢	8.0	8
(V)	W-F	53.6%	0.0% (7.1%)	1.01	3.1%	-64.3¢	-5.7¢	10.7	28
	W-F Dev 1	62.5%	-	1.13	4.7%	-2.5¢	0.8¢	8.3	8
	F-W	18.8%	33.3% (37.5%)	0.86	3.1%	-39.5¢	-25.2¢	11.4	16
	F-W Dev 1	25.0%	-	1.52	6.3%	-44.1¢	34.2¢	10.1	8
(VI)	W-F	75.0%	66.7% (75%)	0.20	0.0%	-15.6¢	-29.7¢	3.5	4
All		48.3%	28.6% (18.3%)	1.05	3.3%	-26.2¢	-10.6¢	8.8	**120**

5.1 Proximity to Stable Matchings

Our experimental markets do not consistently produce a stable outcome. In fact, *just half of the markets result in a stable matching*—48% for the markets with a unique stable matching and 49% of those with multiple stable matchings. The *Stable* column contains the percentages for markets in that arrangement that end in a stable matching.

Markets resulting in an unstable matching have an average of 5.5 blocking pairs. For the 62 unstable markets, 32 have unmatched subjects, while the remaining 30 markets have all the participants matched (with an average of 3.5 blocking pairs per unstable market). Examining the available blocking pairs, we can classify markets into two broad categories. First, there are markets with *available* blocking pairs: blocking pairs that could still form at the final stage of the market, but do not. Available blocking pairs necessarily involve unmatched subjects.[20] Alternatively, there are the *unavailable* blocking pairs: blocking pairs that cannot form because the worker in the pair was either previously rejected by the firm in the pair, or is held by another firm, and subsequently has no agency to make a proposal to form the blocking pair.

Of the 32 markets in which some subjects end the process unmatched, 8 markets had an available blocking pair. Of the remaining 24 markets with unmatched subjects, the unmatched workers were rejected by every blocking firm. In particular, this implies that all unmatched firms must have rejected the unmatched workers at some point in the round. Column *Unmatched* in Table 2 provides the fraction of unmatched workers by market arrangement.

The high rates of unstable outcomes must be due to deviations from truthful reporting. We can associate unstable outcomes to the agents who were responsible by misrepresenting their preferences. To that end, it is useful

[20] This must be a pair comprised of an unmatched worker and a firm such that: i) the firm has not rejected the worker; and ii) the firm is either unmatched or prefers the worker to her current match.

to consider unstable markets that terminate with all participants matched. Consider a worker-firm blocking pair (w, f) for one such matching μ. The blocking pair can be of two types, either: i) firm f previously rejected w (equivalent to f submitting a preference report ranking their ultimate match $\mu(f)$ preferable to w, in contradiction to the definition of (w, f) blocking μ); or ii) worker w never proposes to firm f, but is matched to another firm $\mu(w)$ (equivalent to w stating the current match $\mu(w)$ is preferred to f). From the 30 unstable, fully-matched markets, a third have blocking pairs in both categories, 20% only have blocking pairs corresponding to category (i), where the firms effectively misstate their preferences, and the remaining 47% have only blocking pairs corresponding to category (ii), where workers misstate their preferences. This is suggestive of the substantive misreporting that takes place in our markets and, in particular, the volume of outcomes that are a direct result of proposers misstating their preferences.[21] We further examine the behavior that produces these results in Section 7.

Given the prevalence of markets culminating in unstable matchings, it is interesting to see *how far* the resulting matchings are from the set of stable matchings. We use subjects' preference rankings to create a distance measure for all markets at an unstable outcome. Specifically, we measure the average distance in ranking for each individual between their final match (defining the unmatched outcome as rank 9), and the closest rank of a stable match partner. The results are in the column titled *Distance* in Table 2. On average, subjects were approximately one position away from a stable-match partner across all unstable matches, corresponding to an approximate loss of 50¢ per person (the exception being those markets with lower marginal differences between partner payoffs, where this loss was 20¢).[22]

[21]Markets that are not fully matched produce a similar result. As outlined above, *all* unstable matchings with unmatched subjects and no available blocking pairs *must* have reached this situation through at least one firm misstating their preferences. For the 24 markets with unmatched subjects and no available blocking pairs, 13 contain blocking pairs with workers who failed to make a proposal to the blocking firm.

[22]The overall distance measure for each market arrangement may be calculated by mul-

5.2 Selection of Stable Matchings

We examine those markets with multiple stable matchings and ask which matching the observed outcome is closest to. The *W-Best* column gives the fraction of stable outcomes that were W-optimal, while the figure in parentheses is the fraction of rounds in which the market's outcome is closer to the worker-best outcome than the firm-best, measured in the same manner as *Distance*.

For the markets with multiple stable matchings that yield a stable outcome, 28.6% of the outcomes are at the proposer-best, the outcome resulting from truth-telling in the DA mechanism. However, there is large variation across market arrangements. All stable outcomes in Market (IV) W-F are at the firm-best match, all those under the F-W orientation at the worker-best. Referring back to the truncation column in Table 2, we see that this market is particularly sensitive in the W-F arrangement, reaching the firm best under only a very small preference truncation. Conversely, attaining the firm-best outcome in the F-W order requires extreme truncation by all agents. Inspection of the other markets suggests this as a general trend: When the truncation requirement is low, the stable matching implemented is always the firm-best. With middling-levels of truncation required, the stable matching varies, and with extreme truncation required, the stable matching is always the worker-optimal.

5.3 Tangible Outcomes: Time and Payoffs

5.3.1 Unraveling Time

On average, each market takes approximately 9 turns to finish (see column *Turns*), with the average turn taking 21.5 seconds.Using truth-telling behavior as a benchmark, we compare the actual number of turns taken against

tiplying our distance number by the percentage of unstable matches in a market, as all stable matches are by definition distance 0 from a stable matching.

the predicted number. We find that markets take an extra 2.5 turns to end, and that only 24 out of 120 markets end within the truth-telling number of turns. In particular, these observations suggest that on the market level, any behavior intended to shorten time spent in the experiment was unsuccessful.

5.3.2 Average Payoffs

Consider the average worker in our average market. Conditional on the worker-best outcome being chosen, his expected payoff is \$3.02 per round; and if the firm-best is chosen, it is \$2.57. The observed figures are closer to the latter, lower, prediction, \$2.66. Conducting the same exercise for the firm side of the market, the average firm's expected payoff varies between \$2.66 per round if the worker-best outcome is selected, and \$3.09 under the firm-best. The observed value is \$2.91, in between these two figures.[23] These figures are suggestive of outcomes being closer to the firm-best matching but not strongly so.

Column Δ-*Payoff* provides the average difference in the actual payment from that of the best outcome by market side (that is, the sub-column corresponding to workers contains the difference between the average worker's payoff per round and under the worker-best stable matching. Similarly for the sub-column corresponding to firms[24]). This column contains similar information to the *Distance* and *W-Closer* columns, but provides some interesting differences. In some markets the average matched firms achieve better outcomes than their most preferred stable match partner. In these markets, there is a unique stable outcome, and the average matched worker is faring worse. As will be echoed in the individual analysis below, the reason for these results is that workers in these markets propose to a firm that is ranked below their stable match partner, one that values them more highly

[23] Accounting for unmatched subjects raises these observed averages by approximately 8¢.

[24] These averages are conditional on agents being matched, unmatched subjects are not affecting these figures directly.

than their own stable partner. The ensuing results of such "skips" lead to an unmatched firm and worker. In fact, conditional on being matched, firms earn, on average, 6¢ more than the payoffs generated by their stable match partners in those markets with a unique stable matching. In markets with multiple stable matchings, both sides fare poorly, though the firms are closer in dollar and relative terms to their best stable outcomes.[25]

6 Market Characteristics

We learn from the previous discussion that there is one aspect of a market that predicts *which* stable matching the market produces. How complex the market is to manipulate for firms, as measured by the minimal level of truncation required by firms to establish their preferred matching, is a good predictor of whether the outcome is the worker- or the firm-best matching. We now formalize this idea, and inspect other market characteristics that affect outcomes. Table 3 provides results from several descriptive regressions explaining different dimensions of observed outcomes, using the market characteristics outlined in Section 4 as regressors. The first column outlines the broad effect these design metrics have on a market's duration, the observed number of turns. The next three measures relate to stability: the minimum distance to a stable outcome; the total number of blocking pairs; and a dummy variable indicating whether the final matching was stable or not. Finally, the last column looks at the proximity to the worker-best matching, the left-hand-side variable being a dummy taking the value of 1 when the resulting matching is closer to the worker-best outcome, and restricting the data to those markets with multiple stable matchings.

Formally, we use the following regressors: The first, *Round No.*, takes values from 1 to 15 and represents the position in the sequence of markets

[25] At the end of the experiment, we asked subjects to reflect on the experiment and express their preference over having the role of worker or firm. In line with these payoff differences, 79.6% expressed a preference for the firm role.

Table 3: Descriptive Outcome Regressions

	Turns	Distance	Blocking Pairs	Stable Outcome	Closer to W-Best
Round No.	-0.058	-0.176	0.161	0.005	-0.004*
	(0.082)	(0.137)	(0.126)	(0.005)	(0.002)
Low marginals for workers	0.001	0.101***	0.116**	-0.390***	-0.001
	(0.025)	(0.030)	(0.053)	(0.024)	(0.202)
Low marginals for firms	0.065***	-0.028*	-0.026**	0.107***	0.274***
	(0.015)	(0.013)	(0.010)	(0.023)	(0.053)
Worker core span	0.145*	-0.046	-0.011	0.037	0.102***
	(0.080)	(0.138)	(0.118)	(0.049)	(0.044)
Firm core span	-0.096	0.002	0.060	0.024	-0.048
	(0.056)	(0.091)	(0.078)	(0.020)	(0.055)
Truncation level for F-best	-0.109	-0.039	-0.094	-0.017	0.135***
	(0.066)	(0.125)	(0.100)	(0.037)	(0.032)
N	120	120	120	120	72

Stable Outcome and Closer to W-Best give the marginal effects from a probit regression; all other columns are elasticities obtained from an OLS regression. Standard errors given in parentheses below the estimates, and are clustered by market. Significance levels indicated as follows:***=99%, **=95%, and *=90%.

within an experimental session–the first paid round takes value 1, the last round value 15. The next two regressors capture the effect of the low 20¢ marginals (as opposed to the standard 50¢) on outcomes. The final three regressors are metrics from Table 1, corresponding to the average distance (core span) between the extremal stable matchings, for workers and firms, respectively, and the truncation level firms are required to use to produce the firm-optimal stable matching (*Truncation-F* from Table 1).

We first note that *Round No.* does not carry much explanatory power in our regressions, indicating limited learning or convergence throughout an experimental session.

In terms of market attributes, the different columns highlight several points. First, there does not seem to be any clear pattern in the number of turns taken to conclude a market, as the regression is jointly and individually insignificant at standard levels. Second, the regressions on measures of market stability indicate that low-powered incentives seem to have the strongest effect: Low marginals for proposers significantly increase instability across all three measures. Low marginals for the receivers, the firms, have the opposite effect, increasing outcome stability. Finally, consistent with the casual observation in Section 5, we find that the greater the required truncation levels, and the weaker the firm's incentives, the more likely it is that the observed outcome is closer to the *worker-best* matching. Greater worker incentives (namely, a larger distance between the two stable matchings for the workers) have the same effect.

7 Individual Behavior

The previous section depicts aggregate market outcomes, frequently corresponding to instability. But these aggregate measures are the product of 16 individuals' choice sequences. We now begin an analysis of the individual responses within the experiment.

Figure 2: Distributions of Straightforward/Truthful Play, by Group.

An important finding of the paper is that workers (proposers) do not behave "straightforwardly," in the sense defined in Section 2. That is, their proposals do not track their preference rankings. Firms' (receivers') behavior, on the other hand, is largely straightforward: firms (tentatively) accept proposals from the most-preferred workers in the vast majority of cases. Figure 2 presents the empirical distribution for straightforward, or truthful, play by workers and firms, where each data point represents the fraction of interactions in which a specific subject in each group reveals straightforwardly.

The results are striking. The theory predicts that workers will truthfully reveal their preferences, and firms will strategically misrepresent to achieve better outcomes, most notably (and simply) by truncating preference orderings. In the experiment, over half the subjects acting as firms behave truthfully in *all* their experimental interactions within this role, with two-thirds reporting truthfully more than 90% of the time. The distribution of truth-telling for workers is more uniform—and stochastically dominated by that for firms—with approximately one-third of the workers behaving straightforwardly less than half of the time. In what follows we analyze individuals' behavior in detail.

7.1 Truncation and Skipping

As mentioned above, in the DA procedure workers have a dominant strategy to truthfully reveal their preferences; conversely, if markets have more than one stable matching, firms have a strategic incentive to misrepresent. One easy way of implementing a better stable matching is by truncating the true preferences—disingenuously stating that an acceptable match is less preferred than remaining unmatched.

If every firm were to truncate their preferences below their firm-best stable partner, then the firm-best stable matching would be the resulting outcome. In fact, in many markets a lower degree of truncation can be used by firms to produce their preferred stable matching (see Table 1). Given our data, we can check for the extent of truncation firms are using by direct inference: when an unmatched firm rejects all those proposing in a turn, this is equivalent to stating that the proposals all come from (purportedly) unacceptable workers.[26]

Table 4(a) presents the probability of rejecting all those proposing conditional on the true ranking of the *best* proposer. That is, for any rank k, we track all the events at which a firm (with no tentative acceptances) receives proposals, the best of which is from their k'th ranked partner. The number of these events across all rounds is in column N, and the number in the first round is in the column titled N_1. We calculate the fraction of times that all of these proposals were rejected both across rounds, and in the first round.

[26]We do not observe truncations in any other case. For instance, consider the situation in which two workers, w and $w^{'}$ propose to a particular firm in the same round, and $w^{'}$ is accepted. In this situation we cannot infer whether w was acceptable or unacceptable, all we observe is that $w^{'}$ is preferred to w, and that $w^{'}$ is preferred to no match at all.

When the proposer is the first-best (rank 1), this figure is close to zero. In fact, *truncations within the upper-half of the preference ordering are rare.* As the ranking of the best offer falls (toward 8) the truncation probabilities increase, reaching a probability of rejection of 58.2% for the lowest-ranked proposer. This truncation behavior does not qualitatively differ between the first and subsequent periods, and both exhibit large probabilities in the final two positions of the preference ordering.[27]

The results could be influenced by the large number of observations in particular markets (the two arrangements W-F and F-W of Market (V), for instance). Analyzing each of the markets separately does not change the results drastically.[28]

However, the truncation strategies are not the complete story. The theory makes clear that proposers have a dominant strategy to truthfully reveal their preferences. We now analyze whether workers follow this dominant strategy, and move in sequence through their preference list. Table 4(b) details the probability with which workers propose to the highest-ranked firm from those available: the set of firms that the worker has not yet proposed to. The overall probability is 66.2%, consistent with our initial observations that a substantial number of workers do not make offers in order of their true preferences. The table also reports how these probabilities vary with what may serve as a natural proxy for the likelihood that the proposal will be accepted, the ranking of the worker in the eyes of the targeted firm, the reflected ranking. Specifically, we report the probabilities workers propose to their most-preferred available firm, conditioning on how they themselves stand in the ranking of that firm.[29] In order to provide some control over

[27]Explicitly controlling for the turn number in which the decisions is made generates no effect.

[28]For Market (V), the probability of truncating within the top half of the preference ordering is 3.2% for the W-F treatment, 3.3% in the F-W treatment, and 2.4% for all other markets. For the bottom half, the respective numbers are 31.2%, 46.2%, and 30.6%.

[29]For instance, the first row of the table, corresponding to a worker rank of 1, details the probability with which a worker proposes to the best firm that has not been ruled

Table 4: Individual Behavior

(a) Truncation Levels

Best offer rank	Prob. of rejecting (%) All turns	Prob. of rejecting (%) First turn	Subsample N	Subsample N_1
1	0.2 (0.2)	0.0 (–)	551	80
2	1.1 (0.6)	0.0 (–)	472	118
3	3.2 (1.4)	4.5 (2.1)	402	132
4	8.8 (2.2)	12.1 (3.6)	317	115
5	21.0 (4.8)	19.4 (6.4)	119	36
6	21.8 (5.1)	20.0 (9.7)	87	15
7	45.6 (7.6)	63.6 (14.6)	57	11
8	58.2 (6.9)	50.0 (10.7)	55	22
All	6.6	9.1	2060	529

(b) Skipping

Worker rank	Prob. of proposing (%) All turns	Prob. of proposing (%) First turn	Subsample N	Subsample N_1
1	93.5 (2.6)	96.8 (3.1)	92	32
2	78.4 (2.9)	86.1 (3.9)	208	79
3	66.9 (2.6)	81.6 (3.3)	317	141
4	76.6 (1.9	75.3 (3.2)	487	186
5	61.6) (2.5)	59.1 (4.0)	383	154
6	72.4 (2.4)	60.6 (5.0)	351	94
7	62.1 (2.7)	38.1 (7.5)	314	42
8	46.6 (2.3)	39.2 (3.7)	483	176
All	66.2	9.1	2635	904

Standard errors, clustered by subject, are given in parentheses below sample probabilities. Results were obtained via Probit models with appropriate dummies. Maximum-likelihood estimates for the marginal effects are the subsample mean, as dummy variables induce a partition over the data, but the errors are modeled jointly across each column.

any time effects within a market, we again report separately the probabilities for the first turn within a round (with N and N_1 the number of observations over all turns and over the first turn, respectively).

The results illustrate a clear pattern in proposal behavior; workers are not following their dominant strategy. Instead the *workers are skipping highly ranked firms who are likely to reject them*. This pattern is somewhat more pronounced, though qualitatively similar, for behavior in the first turn of a round.[30]

In many instances this skipping behavior would be inconsequential for outcomes: for instance, if every worker were to skip down to an instituted stable matching, the game would end in a single turn and yield that stable matching. However, in the first turn 19.5% of workers skip down below their own optimal-match partner, and 10.2% skip down to firms ranked below the firm-optimal stable match. Across all turns, conditioning on the availability of the stable partners, 17.2% of workers skip below their own worker-optimal partner, and 8.5% below their firm-optimal partner. We see no qualitative difference between the first and subsequent turns.[31]

Figure 3 illustrates the absolute size of skipping behavior as a cumulative distribution function, classifying the rank of the truthful choice into three categories: the top three, the middle two, and the bottom three. It is easy to see from the graph that the skipping behavior when workers' truthful choices do not rank them highly stochastically dominates the behavior when their truthful choices rank them higher.

out, where that firm ranks them as their top candidate; when the worker's rank is 8, the worker's best current outcome ranks them as the worst match outcome overall.

[30]There is no significant effect from period of play on either the skipping behavior of workers or the truncation decisions of firms. This is true when including the effect as a regressor across all ranks of best proposal, or in subsamples with interacted effects.

[31]This behavior is consistent with the observations of Harrison and McCabe (1992) and Featherstone and Niederle (2008), who observe evidence of proposer skipping in different environments while employing the direct-revelation mechanism.

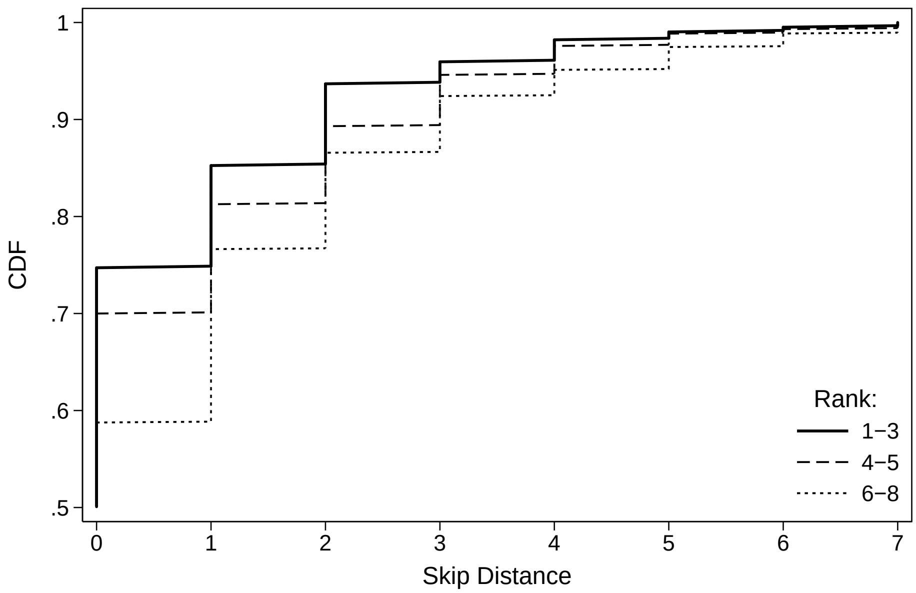

Figure 3: Size of Skip Conditional on Truthful Choice's Ranking

7.2 Risk Aversion

There is complete information in our experimental design, but given how complex the markets are, we explore how strategic risk may have affected subjects' behavior. We imagine that risk-averse workers, who are uncertain about the behavior of others, may want to "skip" in a similar fashion to what we observe.

In order to assess the level of this "riskiness" we make some stark behavioral assumptions. First, we assume that subjects in worker roles internalize the truncation probabilities that the firms are using, but understand their choice as that of a simple static lottery: if the firm they propose to accepts their offer they win the lottery, and earn the payoff from matching with that firm; if the firm rejects, then they lose the lottery, and receive a fixed payment representing the perceived "outside option." Therefore, in the first round, each worker acts as if they are choosing from eight potential lotteries, choosing the best lottery according to their risk preferences.

We parameterize risk preferences with the commonly used constant relative risk aversion specification over monetary payoffs. So the particular choice is made according to:

$$\max_{f \in F} p_f \cdot \pi_f^{1-\sigma} + (1 - p_f) \cdot \underline{\pi}^{1-\sigma},$$

where p_f is the probability choice firm f will reject the proposer, π_f is the monetary payoff from choice f, and $\underline{\pi}$ is a common continuation value.

Calculating p_f via the probabilities given in Table 4(b),we estimate a risk coefficient of σ of 0.60, and an outside-option payment $\underline{\pi}$ of \$2.20.[32] This estimate for the risk aversion is similar to that observed in experiments designed to estimate this parameter through explicit lottery choice.(see Holt and Laury 2002, Table 3).

We stress that the risk hypothesis is only one potential explanation for the observed skipping, and relies heavily on strong behavioral assumptions that reformulate the complex game we use as a choice over simple lotteries.We interpret this estimation as a failure to reject, rather than strong evidence supporting this model of behavior.

8 Conclusion

The paper reports observations from experiments emulating a highly utilized matching clearinghouse, the deferred-acceptance (DA) mechanism. We studied a large set of markets, varying in their complexity, incentives to truthfully reveal preferences, and cardinal representations. Several important insights

[32]We use a non-linear logit specification over the possible lotteries, and restrict attention to the first round of choices to ensure a constant choice size. The 99% confidence interval for σ is $[0.57, 0.63]$, while that for $\underline{\pi}$ is $[172, 258]$ cents, where these intervals are obtained by a bootstrap of size 1,000.

emerge from our experiments. First, less than half of the markets generated a stable matching. Of those markets with multiple stable matchings that resulted in a stable outcome, over 70% concluded in the receiver-best stable matching. Since truthful revelation of preferences generates the proposer-best outcome, these results are suggestive of manipulation. In fact, our second set of insights regards the consistent deviations from truthful behavior. Proposers frequently skipped down their preference ordering, preferring to propose early to those more likely to accept them. Receivers, however, appeared to by and large behave in an effectively truthful manner, accepting the best offer at each point in time. This is in contrast to the underlying theoretical predictions that proposers behave truthfully and receivers not (most simply, by truncating preferences). Last, we show that market attributes have a significant impact on outcomes. For instance, both the cardinal representation and core size influence whether outcomes are stable. They also impact the overall distance of observed outcomes from the core, and the number of turns it takes markets to converge to the final outcome.

The study has potentially important practical implications given the wide use of the DA mechanism. Indeed, consider the medical residents match in the U.S. (the NRMP), involving over $40,000$ participants each year. The behavior we observe in the lab could translate into medical residents from top programs applying to top-tier residencies, while those from less well-regarded schools aiming at middle-ranked hospitals and below. Naturally, outcomes are then very fragile to mistakes (by residents) regarding how low to apply, even if hospitals submit their preferences truthfully. While the centralized system is designed to generate stable matchings, such behavior may cause the system to converge to outcomes that are, in fact, unstable.

The paper opens the door for several directions for future research. In light of the behavior we observe, it would be important to understand formally how fragile outcomes are to particular skipping heuristics by proposers. Furthermore, it would be interesting to discern how certain forms of communication affect outcomes in mechanisms such as those we study (for instance, interviews could be thought of as match quality signals that are revealed prior to the enactment of the centralized clearinghouse). It would also be crucial to determine how incomplete information regarding others' preferences, which is likely in such large markets as the NRMP, shapes outcomes and behavior in matching clearinghouses, particularly in view of the evidence we see suggesting the importance of risk aversion in our experimental markets.

References

Chen, Y., and T. Sönmez (2006): "School choice: an experimental study," *Journal of Economic Theory*, 127(1), 202–231.

Echenique, F., G. Katz, and L. Yariv (2010): "An experimental study of decentralized matching," Discussion paper, Working paper, Caltech.

Featherstone, C., and E. Mayefsky (2010): "Stability and Deferred Acceptance: Strategic Behavior in Two-Sided Matching," Discussion paper, Working paper, Stanford University.

Featherstone, C., and M. Niederle (2008): "Manipulation in school choice mechanisms," Discussion paper, Working paper, Stanford University.

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

Harrison, G., and K. McCabe (1992): "Testing noncooperative bargaining theory in experiments," *Research in experimental economics*, 5, 137–69.

Haruvy, E., and U. Ünver (2007): "Equilibrium Selection and the Role of Information in Repeated Matching Markets," *Economic Letters*.

Holt, C., and S. Laury (2002): "Risk aversion and incentive effects," *The American Economic Review*, 92(5), 1644–1655.

Kagel, J., and A. Roth (2000): "The Dynamics of Reorganization in Matching Markets: A Laboratory Experiment Motivated by A Natural Experiments*," *Quarterly Journal of Economics*, 115(1), 201–235.

Nalbantian, H., and A. Schotter (1995): "Matching and efficiency in the baseball free-agent system: an experimental examination," *Journal of Labor Economics*, 13(1), 1–31.

Niederle, M., and L. Yariv (2010): "Matching Through Decentralized Markets," mimeo, California Institute of Technology and Stanford University.

Richter, M. K. (1966): "Revealed Preference Theory," *Econometrica*, 34(3), 635–645.

Roth, A., and M. Sotomayor (1990): *Two-sided Matching: A Study in Game-Theoretic Modelling and Analysis*, vol. 18 of *Econometric Society Monographs*. Cambridge University Press, Cambridge England.

A Dynamic Clearinghouses

A.1 Static and Dynamic Gale-Shapley Mechanisms

The following example, appearing in Niederle and Yariv (2010), illustrates how weakly dominated strategies on the parts of workers alone do not lead to the same predictions in the static and dynamic versions of the Deferred Acceptance mechanism.

Example (Additional Equilibria Outcomes in the Dynamic Version of Deferred Acceptance). Consider a market consisting of workers $\{W1, W2, W3\}$ and firms $\{F1, F2, F3\}$, where all agents prefer to be matched rather than unmatched. Let the induced ordinal preferences $\succsim$ of the three workers and colors be given by:

$$\begin{array}{ll} \mathbf{W1}: & F2 \succ \mathbf{F1} \succ F3 \\ \mathbf{W2}: & F1 \succ \mathbf{F2} \succ F3 \\ \mathbf{W3}: & F1 \succ F2 \succ \mathbf{F3} \end{array}, \qquad \begin{array}{ll} F1: & \mathbf{W1} \succ W3 \succ W2 \\ F2: & \mathbf{W2} \succ W1 \succ W3 \\ F3: & W1 \succ \mathbf{W3} \succ W2 \end{array}.$$

The unique stable matching μ is given below (where we use the convention that each column in the matrix denotes a match between the specified worker and color), $\mu(Wi) = Fi$ for all i. In particular, the Gale Shapley mechanism entails a unique equilibrium in weakly undominated strategies yielding μ. Nonetheless, the matching $\tilde{\mu}$ below (in which $W1$ and $W2$ swap colors relative to μ) can be induced in our dynamic mechanism.

$$\mu = \begin{pmatrix} W1 & W2 & W3 \\ F1 & F2 & F3 \end{pmatrix}, \quad \tilde{\mu} = \begin{pmatrix} W1 & W2 & W3 \\ F2 & F1 & F3 \end{pmatrix}.$$

Indeed, here is a profile in weakly undominated strategies:

Period 1: worker $W3$ makes an offer to $F3$ who accepts.

Period 2: worker $W1$ makes an offer to $F2$ and $W2$ makes an offer to $F1$ who accept.

Upon any deviation, offers from agents other than the stable match or the most preferred match are rejected and all revert to emulating the Deferred Acceptance strategies (in particular, $F1$ rejects an offer from $W3$).

Notice that time plays an important role in the construction of this equilibrium. Indeed, as highlighted Niederle and Yariv (2010), the crucial element driving this construction is the ability of some participants to commit and of others to condition their behavior on observed market outcomes (note that once $W3$ is accepted, he cannot escape $F3$).[33]

A.2 Outcome and Strategic Equivalence

In the dynamic setup, at each period t agents monitor only partial activity in the market. We now describe the information each agent has throughout the game. At the beginning of period t, each worker w observes a history that consists of the (timed) offers the worker made and the responses of firms to those offers, denoted by r for rejection and h for holding (where we use the notational convention that an offer to no firm is denoted as an offer to $\varnothing$ that is immediately rejected):

$$h_{t,w}^{W} \in \left((F \cup \varnothing) \times \{r, h\}\right)^{t-1}.$$

The set of all possible histories at time t for worker w is denoted by $H_{t,w}^{W}$.

In addition, at each period t, suppose firms $f_1, ..., f_{k(t-1)}$ rejected offers from worker w in periods $1, ..., t-1$. Denote by $\tilde{F}_w^t = \left\{f | f \notin \left\{f_1, ..., f_{k(t-1)}\right\}\right\}$ the set of firms that have not rejected worker w yet.

Each firm acts in the second stage of each period t and observes a history that consists of all (timed) offers she received and a (timed) sequence of offers

[33] Interestingly, this equilibrium is not robust in that it is not sequential (for instance, $F1$ would need to believe that other agents will deviate as well when observing an offer from $W3$, but the market does not offer enough monitoring for that).

she held:[34]

$$h^F_{t,f} \in \left(2^W\right)^t \times \left(2^W\right)^t.$$

The set of all possible histories at time t for firm f is denoted by $H^F_{t,f}$.

In addition, at each period t, suppose workers $w_1, ..., w_{k(t-1)}$ made offers to firm f in periods $1, ..., t$. Denote by $\tilde{W}^t_f = \left\{w | w \notin \left\{w_1, ..., w_{k(t-1)}\right\}\right\}$ the set of workers that have not made an offer to firm f.

A strategy for worker w is a collection of mappings $\left\{\sigma^W_{t,w}\right\}$, where $\sigma^W_{t,w} : H^W_{t,w} \to F \cup \varnothing$, and whenever at time t, $\sigma^W_{t,w}\left(h^W_{t,w}\right) \neq \varnothing$ then $\sigma^W_{t,w}\left(h^W_{t,w}\right) \in \tilde{F}^t_w$. A strategy for firm f is a collection of mappings $\left\{\sigma^F_{t,f}\right\}$, where $\sigma^F_{t,f} : H^F_{t,f} \to (W \cup \varnothing)^{2^W \times (W \cup \varnothing)}$. That is, after each history, the firm's strategy specifies which worker (if any) would be held from a menu of worker offers (when possibly already holding an offer).

Note that, for workers, we could, in fact, describe the strategy as: $\sigma^W_{t,w} : H^W_{t,w} \to \{P(w)\}$ (when defining a firm approached at later periods as less preferred).

If agents condition their behavior on time per se, the dynamic setup may, in principle, lead to very different outcomes than the static one. We make the following assumptions:

Assumption (Stationarity) Strategies do not depend on sequencing:

- For any worker w, there exists $\tau^W_w : 2^F \to F \cup \varnothing$, such that whenever at time t worker w is not held and under history $h^W_{t,w}$, $\tilde{F}^t_w$ are the firms he can make an offer to, $\sigma^W_{t,w}\left(h^W_{t,w}\right) = \tau^W_w\left(\tilde{F}^t_w\right)$.
- For any firm f, there exists $\tau^F_f : 2^W \times (W \cup \varnothing) \to W \cup \varnothing$, such that whenever at time t firm f has observed history $h^F_{t,f}$, under which she holds an offer from $f \in F \cup \varnothing$ (where holding an offer from $\varnothing$

[34] An offer of worker w to firm f that is held from period t to t' is recorded as an offer made in periods $t, t+1, ..., t$' that is held by the firm in each of these periods. We use a similar convention for workers.

is interpreted as not holding an offer), and the set of workers who made her an offer in t is $\tilde{W}$, then $\sigma_{t,f}^{F}\left(h_{t,f}^{F}\right) = \tau_{f}^{F}\left(\tilde{W}, w\right)$.

Assume that workers make offers whenever they can.

Stationarity in and of itself does not assure a representation through a preference ranking. Indeed, if $\tau_{w}^{W}(f_1, f_2) = f_1$, but $\tau_{w}^{W}(f_1, f_2, f_3) = f_2$, this would not be consistent with a preference ordering. Namely, a form of independence of irrelevant alternatives is being violated. Furthermore, if $\tau_{w}^{W}(f_1, f_2) = f_1, \tau_{w}^{W}(f_2, f_3) = f_2$, and $\tau_{w}^{W}(f_3, f_1) = f_3$, we would obtain a violation of transitivity when trying to explain behavior through a preference ordering. This is in the spirit of violations of the weak axiom of revealed preferences.

The equivalence between the two types of mechanisms rests on a familiar idea from choice theory.

Let X be a finite set and $\mathbb{B} \subseteq 2^X$. A choice function is a function $C : \mathbb{B} \to X$ such that $C(A) \in A$ for all $A \in \mathbb{B}$. We can associate a binary relation $\succ^C$ with C, where $x \succ^C y$ if and only if there is a set $A \in \mathbb{B}$ with $x, y \in A$ and $x = C(A)$. Note that $\succ^C$ is the revealed-preference relation.

The choice function C satisfies the *congruence axiom* if $\succ^C$ is acyclic; that is, if whenever $x_1, \ldots x_K$ is a sequence in X such that

$$x_1 \succ^C x_2 \succ^C \ldots \succ^C x_K,$$

then it is false that $x_K \succ^C x_1$.

In our setup, each worker w and firm f is characterized by a choice function: τ_{w}^{W} and τ_{f}^{F}, respectively. We say that the congruence axiom holds when all agents' choice functions satisfy the congruence axiom.

Proposition (Equivalence) *Whenever stationarity and the congruence axiom hold, equilibria outcomes in weakly undominated strategies of the static Gale-Shapley mechanism coincide with equilibria outcomes in weakly undominated strategies of the dynamic Gale-Shapley mechanism.*

Furthermore, there is a one-to-one mapping between weakly undominated equilibrium strategy profiles corresponding to the two mechanisms.

B Markets

The ordinal preference profiles for the main market variants we run are given below. Exact payoff tables, deviations for markets, and details on the stable matches for these markets are available on request from the authors.

Fruit preferences	Color preferences
$\mathbf{f_1}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_1}: f_1 \succ f_2 \succ f_3 \succ f_4 \succ f_5 \succ f_6 \succ f_7 \succ f_8$
$\mathbf{f_2}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_2}: f_1 \succ f_2 \succ f_3 \succ f_4 \succ f_5 \succ f_6 \succ f_7 \succ f_8$
$\mathbf{f_3}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_3}: f_1 \succ f_2 \succ f_3 \succ f_4 \succ f_5 \succ f_6 \succ f_7 \succ f_8$
$\mathbf{f_4}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_4}: f_1 \succ f_2 \succ f_3 \succ f_4 \succ f_5 \succ f_6 \succ f_7 \succ f_8$
$\mathbf{f_5}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_5}: f_1 \succ f_2 \succ f_3 \succ f_4 \succ f_5 \succ f_6 \succ f_7 \succ f_8$
$\mathbf{f_6}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_6}: f_1 \succ f_2 \succ f_3 \succ f_4 \succ f_5 \succ f_6 \succ f_7 \succ f_8$
$\mathbf{f_7}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_7}: f_1 \succ f_2 \succ f_3 \succ f_4 \succ f_5 \succ f_6 \succ f_7 \succ f_8$
$\mathbf{f_8}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_8}: f_1 \succ f_2 \succ f_3 \succ f_4 \succ f_5 \succ f_6 \succ f_7 \succ f_8$

Market (I): Assortative

Fruit preferences	Color preferences
$\mathbf{f_1}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_1}: f_3 \succ f_8 \succ f_7 \succ f_4 \succ f_5 \succ f_6 \succ f_1 \succ f_2$
$\mathbf{f_2}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_2}: f_6 \succ f_4 \succ f_5 \succ f_3 \succ f_2 \succ f_7 \succ f_1 \succ f_8$
$\mathbf{f_3}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_3}: f_8 \succ f_5 \succ f_3 \succ f_6 \succ f_1 \succ f_7 \succ f_2 \succ f_4$
$\mathbf{f_4}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_4}: f_5 \succ f_3 \succ f_4 \succ f_1 \succ f_7 \succ f_8 \succ f_6 \succ f_2$
$\mathbf{f_5}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_5}: f_4 \succ f_1 \succ f_2 \succ f_6 \succ f_7 \succ f_3 \succ f_8 \succ f_5$
$\mathbf{f_6}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_6}: f_4 \succ f_7 \succ f_3 \succ f_8 \succ f_5 \succ f_2 \succ f_6 \succ f_1$
$\mathbf{f_7}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_7}: f_5 \succ f_1 \succ f_3 \succ f_7 \succ f_4 \succ f_8 \succ f_6 \succ f_2$
$\mathbf{f_8}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_8}: f_2 \succ f_1 \succ f_3 \succ f_7 \succ f_6 \succ f_5 \succ f_4 \succ f_8$

Market (II): One Full Aligned

Fruit preferences	Color preferences
$\mathbf{f_1}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_1}: f_5 \succ f_1 \succ f_6 \succ f_2 \succ f_3 \succ f_7 \succ f_4 \succ f_8$
$\mathbf{f_2}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_2}: f_1 \succ f_4 \succ f_2 \succ f_3 \succ f_6 \succ f_8 \succ f_7 \succ f_5$
$\mathbf{f_3}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_3}: f_4 \succ f_1 \succ f_3 \succ f_2 \succ f_5 \succ f_6 \succ f_8 \succ f_7$
$\mathbf{f_4}: c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_7 \succ c_8$	$\mathbf{c_4}: f_3 \succ f_1 \succ f_4 \succ f_2 \succ f_5 \succ f_8 \succ f_7 \succ f_6$
$\mathbf{f_5}: c_5 \succ c_6 \succ c_7 \succ c_8 \succ c_1 \succ c_2 \succ c_3 \succ c_4$	$\mathbf{c_5}: f_2 \succ f_1 \succ f_7 \succ f_5 \succ f_4 \succ f_8 \succ f_3 \succ f_6$
$\mathbf{f_6}: c_5 \succ c_6 \succ c_7 \succ c_8 \succ c_1 \succ c_2 \succ c_3 \succ c_4$	$\mathbf{c_6}: f_7 \succ f_6 \succ f_5 \succ f_8 \succ f_2 \succ f_4 \succ f_1 \succ f_3$
$\mathbf{f_7}: c_5 \succ c_6 \succ c_7 \succ c_8 \succ c_1 \succ c_2 \succ c_3 \succ c_4$	$\mathbf{c_7}: f_6 \succ f_8 \succ f_5 \succ f_7 \succ f_1 \succ f_4 \succ f_2 \succ f_3$
$\mathbf{f_8}: c_5 \succ c_6 \succ c_7 \succ c_8 \succ c_1 \succ c_2 \succ c_3 \succ c_4$	$\mathbf{c_8}: f_8 \succ f_7 \succ f_6 \succ f_5 \succ f_3 \succ f_2 \succ f_4 \succ f_1$

Market (III): Split Two Aligned

Fruit preferences	Color preferences
$\mathbf{f_1}: c_1 \succ c_7 \succ c_8 \succ c_2 \succ c_5 \succ c_4 \succ c_3 \succ c_6$	$\mathbf{c_1}: f_2 \succ f_6 \succ f_3 \succ f_4 \succ f_5 \succ f_7 \succ f_8 \succ f_1$
$\mathbf{f_2}: c_2 \succ c_8 \succ c_3 \succ c_4 \succ c_1 \succ c_6 \succ c_5 \succ c_7$	$\mathbf{c_2}: f_8 \succ f_4 \succ f_6 \succ f_3 \succ f_5 \succ f_7 \succ f_1 \succ f_2$
$\mathbf{f_3}: c_3 \succ c_4 \succ c_1 \succ c_5 \succ c_8 \succ c_2 \succ c_6 \succ c_7$	$\mathbf{c_3}: f_8 \succ f_5 \succ f_1 \succ f_6 \succ f_2 \succ f_4 \succ f_7 \succ f_3$
$\mathbf{f_4}: c_4 \succ c_1 \succ c_5 \succ c_8 \succ c_2 \succ c_3 \succ c_7 \succ c_6$	$\mathbf{c_4}: f_2 \succ f_1 \succ f_6 \succ f_3 \succ f_7 \succ f_8 \succ f_5 \succ f_4$
$\mathbf{f_5}: c_5 \succ c_2 \succ c_1 \succ c_7 \succ c_3 \succ c_8 \succ c_4 \succ c_6$	$\mathbf{c_5}: f_1 \succ f_3 \succ f_6 \succ f_7 \succ f_4 \succ f_2 \succ f_8 \succ f_5$
$\mathbf{f_6}: c_6 \succ c_5 \succ c_8 \succ c_2 \succ c_4 \succ c_7 \succ c_3 \succ c_1$	$\mathbf{c_6}: f_3 \succ f_1 \succ f_5 \succ f_7 \succ f_4 \succ f_8 \succ f_2 \succ f_6$
$\mathbf{f_7}: c_7 \succ c_6 \succ c_2 \succ c_5 \succ c_3 \succ c_1 \succ c_4 \succ c_8$	$\mathbf{c_7}: f_1 \succ f_4 \succ f_3 \succ f_8 \succ f_5 \succ f_2 \succ f_6 \succ f_7$
$\mathbf{f_8}: c_8 \succ c_3 \succ c_1 \succ c_2 \succ c_5 \succ c_6 \succ c_7 \succ c_4$	$\mathbf{c_8}: f_2 \succ f_6 \succ f_3 \succ f_1 \succ f_7 \succ f_5 \succ f_4 \succ f_8$

Market (IV): Two matches, one very unstable

Fruit preferences	Color preferences
$\mathbf{f_1}: c_6 \succ c_1 \succ c_8 \succ c_4 \succ c_3 \succ c_2 \succ c_5 \succ c_7$	$\mathbf{c_1}: f_3 \succ f_6 \succ f_1 \succ f_7 \succ f_5 \succ f_8 \succ f_2 \succ f_4$
$\mathbf{f_2}: c_3 \succ c_1 \succ c_2 \succ c_4 \succ c_5 \succ c_7 \succ c_8 \succ c_6$	$\mathbf{c_2}: f_3 \succ f_8 \succ f_1 \succ f_7 \succ f_2 \succ f_4 \succ f_6 \succ f_5$
$\mathbf{f_3}: c_3 \succ c_6 \succ c_8 \succ c_2 \succ c_1 \succ c_7 \succ c_5 \succ c_4$	$\mathbf{c_3}: f_1 \succ f_8 \succ f_3 \succ f_4 \succ f_2 \succ f_6 \succ f_5 \succ f_7$
$\mathbf{f_4}: c_3 \succ c_4 \succ c_7 \succ c_5 \succ c_1 \succ c_2 \succ c_6 \succ c_8$	$\mathbf{c_4}: f_2 \succ f_1 \succ f_4 \succ f_5 \succ f_7 \succ f_3 \succ f_8 \succ f_6$
$\mathbf{f_5}: c_6 \succ c_1 \succ c_5 \succ c_3 \succ c_2 \succ c_4 \succ c_8 \succ c_7$	$\mathbf{c_5}: f_2 \succ f_8 \succ f_3 \succ f_5 \succ f_1 \succ f_4 \succ f_6 \succ f_7$
$\mathbf{f_6}: c_6 \succ c_2 \succ c_4 \succ c_5 \succ c_1 \succ c_7 \succ c_8 \succ c_3$	$\mathbf{c_6}: f_2 \succ f_8 \succ f_7 \succ f_5 \succ f_4 \succ f_6 \succ f_3 \succ f_1$
$\mathbf{f_7}: c_8 \succ c_7 \succ c_1 \succ c_6 \succ c_2 \succ c_3 \succ c_5 \succ c_4$	$\mathbf{c_7}: f_1 \succ f_2 \succ f_8 \succ f_6 \succ f_5 \succ f_4 \succ f_3 \succ f_7$
$\mathbf{f_8}: c_8 \succ c_1 \succ c_4 \succ c_3 \succ c_2 \succ c_7 \succ c_5 \succ c_6$	$\mathbf{c_8}: f_1 \succ f_5 \succ f_8 \succ f_4 \succ f_3 \succ f_7 \succ f_6 \succ f_2$

Market (V): Two matches, unaligned preferences

Fruit preferences	Color preferences
$\mathbf{f_1}: c_2 \succ c_4 \succ c_1 \succ c_3 \succ c_8 \succ c_7 \succ c_6 \succ c_5$	$\mathbf{c_1}: f_1 \succ f_2 \succ f_3 \succ f_4 \succ f_5 \succ f_6 \succ f_8 \succ f_7$
$\mathbf{f_2}: c_2 \succ c_1 \succ c_4 \succ c_3 \succ c_7 \succ c_6 \succ c_5 \succ c_8$	$\mathbf{c_2}: f_3 \succ f_1 \succ f_2 \succ f_4 \succ f_8 \succ f_5 \succ f_7 \succ f_6$
$\mathbf{f_3}: c_4 \succ c_2 \succ c_3 \succ c_1 \succ c_6 \succ c_5 \succ c_8 \succ c_7$	$\mathbf{c_3}: f_3 \succ f_4 \succ f_1 \succ f_2 \succ f_6 \succ f_7 \succ f_5 \succ f_8$
$\mathbf{f_4}: c_1 \succ c_3 \succ c_4 \succ c_2 \succ c_5 \succ c_8 \succ c_7 \succ c_6$	$\mathbf{c_4}: f_2 \succ f_3 \succ f_4 \succ f_1 \succ f_7 \succ f_8 \succ f_6 \succ f_5$
$\mathbf{f_5}: c_6 \succ c_8 \succ c_5 \succ c_7 \succ c_4 \succ c_3 \succ c_2 \succ c_1$	$\mathbf{c_5}: f_5 \succ f_6 \succ f_7 \succ f_8 \succ f_1 \succ f_2 \succ f_4 \succ f_3$
$\mathbf{f_6}: c_6 \succ c_5 \succ c_8 \succ c_7 \succ c_3 \succ c_2 \succ c_1 \succ c_4$	$\mathbf{c_6}: f_7 \succ f_5 \succ f_6 \succ f_8 \succ f_4 \succ f_1 \succ f_3 \succ f_2$
$\mathbf{f_7}: c_8 \succ c_6 \succ c_7 \succ c_5 \succ c_2 \succ c_1 \succ c_4 \succ c_3$	$\mathbf{c_7}: f_7 \succ f_8 \succ f_5 \succ f_6 \succ f_2 \succ f_3 \succ f_1 \succ f_4$
$\mathbf{f_8}: c_5 \succ c_7 \succ c_8 \succ c_6 \succ c_1 \succ c_4 \succ c_3 \succ c_2$	$\mathbf{c_8}: f_6 \succ f_7 \succ f_8 \succ f_5 \succ f_3 \succ f_4 \succ f_2 \succ f_1$

Market (VI): Four by Four market