An Experimental Study of Decentralized Matching

Federico Echenique[*] Alejandro Robinson-Cortés[†] Leeat Yariv[‡§]

October 14, 2024

Abstract

We present an experimental study of decentralized two-sided matching markets with no transfers. Experimental participants are informed of everyone's preferences and can make arbitrary non-binding match offers that get finalized when a period of market inactivity has elapsed. Several insights emerge. First, stable outcomes are prevalent. Second, while centralized clearinghouses commonly aim at implementing extremal stable matchings, our decentralized markets most frequently culminate in the *median* stable matching. Third, preferences' cardinal representations impact the stable partners participants match with. Last, the dynamics underlying our results exhibit strategic sophistication, with agents successfully avoiding cycles of blocking pairs.

Keywords: Decentralized Matching, Experiments, Market Design.

[*]Department of Economics, UC Berkeley, fede@econ.berkeley.edu

[†]Department of Economics, University of Exeter, a.robinson-cortes@exeter.ac.uk

[‡]Department of Economics, Princeton University, lyariv@princeton.edu

[§]We thank Drew Fudenberg, Jacob Goeree, Matthew Jackson, Gabriel Katz, Erik Madsen, Sergio Montero, Muriel Niederle, Charles Plott, Andrew Schotter, and Alistair Wilson for useful conversations and suggestions. We also thank the Editor and three anonymous reviewers for very helpful comments. Financial support from the National Science Foundation (SES 0963583), and the Gordon and Betty Moore Foundation (grant 1158) is gratefully acknowledged.

1 Introduction

This paper presents an experimental investigation of decentralized two-sided matching markets with no transfers. We consider two-sided markets, such as labor markets consisting of workers and firms, marriage markets comprising women and men, and so on. We study the outcomes emerging from free interactions between market participants. Would unconstrained decentralized interaction produce stable outcomes?[1] When there are multiple stable matchings, would some have stronger drawing power? What dynamics would decentralized interactions follow? Our experiments are designed to answer these questions.

Matching theory offers predictions for the set of plausible market outcomes, namely the set of stable matchings, or the core, under two basic premises: all agents are completely informed of others' preferences, as well as their own; and agents can match with one another freely. Our experiments are designed to test the fundamental predictions of this model. While the assumption that market forces would guide matchings to stability has governed much of the literature, it has never been tested. In this regard, our approach is reminiscent of some of the original general equilibrium experiments (e.g., Smith, 1962; Plott and Smith, 1978), which examined whether markets reach an equilibrium without imposing constraints on the sequential actions that lead them there.

In the main treatments of our experiment each side of the market is composed of 8 participants. Market participants are fully informed of their own and other participants' match payoffs, the payoffs derived from all possible pairings. Preferences are designed so that each market participant has either one, two, or three stable match partners. As part of our design, we utilize several cardinal utility representations for the same ordinal preference profiles, varying the utilitarian welfare of each side of the market as well as the marginal returns from

[1] A stable outcome corresponds to a pairing of agents from both market sides such that no agent prefers to sever their partnership and no pair of agents prefer to pair with one another over remaining with their assigned partners.

matching with one partner as opposed to a more preferred one. At any point during a market's operation, agents on each side are free to make match offers to any agent on the other side of the market, up to one offer at a time, as well as accept any offer that arrives, up to one at a time. Importantly, agents can keep making and receiving offers while matched, so matches are tentative until the market ends, which occurs after 30 seconds of inactivity.

Our experiments are designed to allow the exploitation of blocking opportunities, the main theoretical driving force behind stability. Nevertheless, there are reasons to suspect stability will be difficult to achieve through decentralized interactions, or at the very least require many offers, many rejections, and a long time. Indeed, the leading dynamics that the theoretical literature has offered assume sequential formation of blocking pairs (Knuth, 1976). Roth and Vate (1990) proved that there is always *at least one* path of such sequential blocking pairs leading to a stable matching. However, not every sequence of blocking pairs culminates in a stable matching. Knuth (1976) showed that such paths can also generate indefinite cycles.[2] In fact, the sequential formation of blocking pairs can have an almost chaotic behavior (Rudov, 2022). From a computational perspective, finding the set of stable matchings is a hard problem in general, see Gusfield and Irving (1989), and even markets with 8 individuals on each side entail $2 \times 8 \times 8 = 128$ match payoffs that participants would need to internalize in order to identify the market-wide stable matching.

In our data, *stable outcomes appear in a predominance of cases*, with 88% of markets reaching a stable matching. Despite the existence of cyclic "traps," participants steer clear of cycles, and converge quickly to a stable matching, both in terms of offer volume, which

[2]For example, suppose each market side has two agents—f_1 and f_2 on one, c_1 and c_2 on the other—and that all agents view all others as acceptable. Assume f_i's preferred partner is c_i, while c_i's preferred partner is f_{3-i}, $i = 1, 2$. Suppose, say, that f_1 and c_1 are the only matched participants initially. Then, (f_2, c_1) is a blocking pair. If this match is formed, (f_2, c_2) becomes a blocking pair, which, if formed, makes (f_1, c_2) a blocking pair. Finally, if f_1 and c_2 match, then (f_1, c_1) is a blocking pair. Upon matching, the market returns to its starting point.

averages 46 offers per market, and in terms of market duration, which takes less than two minutes on average.

In general, stable outcomes exhibit a lattice structure, with two extremal stable matchings: one preferred by all the agents on one side of the market and one preferred by all the agents on the other side. These extremal stable matchings are computationally simple to identify via the Gale and Shapley (1962) Deferred Acceptance (DA) algorithm. Since the DA algorithm underlies many centralized clearinghouses in the field, centralized outcomes frequently correspond to extremal stable matchings for the reported preferences. As noted, the full set of stable matchings is computationally challenging to compute. In principle, however, in our markets with multiple stable matchings, decentralized interactions could yield any stable matching, including the *median* stable matching, the matching ranked by all agents between the two extremal stable matchings (Teo, Sethuraman, and Tan, 2001). Matching theory is mute regarding the selection of stable matchings, as well as on the effects of cardinal preference representations: stability accounts for only ordinal rankings of partners.

In our data, *the median stable matching emerges as the modal outcome*. In fact, 77% of pairings are between median stable match partners, and 80% of stable matchings correspond to the market-wide median stable matching. We also find that the *cardinal representation of preferences significantly influences the selection of stable matchings*. In particular, the side of the market that has "more to lose" by forgoing their most preferred matching tends to establish its most preferred matching more frequently. These findings are potentially useful for policy. When considering centralized interventions, our results highlight the potential importance of cardinal preference information, on top of ordinal rankings, for counterfactual decentralized outcomes. They also suggest that features of median stable matchings—utilitarian welfare, distributional attributes, and the like—may be relevant for assessing when centralized clearinghouses might be particularly beneficial.

The underlying dynamics of interactions are at the heart of the outcomes we observe. With field data, it is naturally challenging to track the complete protocol of market participants' interactions.[3] Our data allow a lens into decentralized dynamics and a comparison with some dynamics proposed by the theoretical matching literature, including the sequential formation of blocking pairs (e.g., Roth and Vate, 1990; Ackermann, Goldberg, Mirrokni, Röglin, and Vöcking, 2011), and the dynamics underlying the DA algorithm of Gale and Shapley (1962) (e.g., Dworczak, 2021).

Our last set of results pertains to the dynamic paths markets follow to reach stability. The speed at which stable pairings form decreases over markets' operation: a substantial fraction of agents find stable partners quickly, followed by a slow-down of market activity while the remaining agents take longer to find their stable matches. While, as expected, agents do not appear to see through the market-wide matching and make offers to stable partners immediately, *participants exhibit strategic sophistication and avoid blocking pairs that could trap them in a cycle*. Participants make more frequent offers to blocking partners, particularly to those who benefit more from the pairing. Furthermore, not all offers from blocking partners are accepted. These patterns help explain the convergence to the median stable matchings, as well as the avoidance of cycles and consequent quick convergence to stability that we observe.

We also simulate some of the dynamics the theoretical literature has proposed to gain a comparison benchmark for our data. These simulations indicate that existing models do not organize our data well. Indeed, existing models are, by and large, non-strategic, and do not incorporate the history of past interactions. Our results open the door to new models, allowing strategic behavior as we observe.

In addition to our main treatments, we also report results from two auxiliary treatments that investigate our main results' robustness, in terms of market size and bargaining power.

[3] One exception is Hitsch, Hortaçsu, and Ariely (2010), who use detailed data from an online dating site.

The first set of treatments involves larger markets, with 15 participants, instead of 8, on each side. The second set of treatments allows agents only on one side of the market to make offers. Qualitatively, the findings from our main treatments continue to hold: stable matchings, and in particular median stable matchings, are very frequent. Similarly, the dynamic patterns we identify in our main treatments are still present in larger markets, as well as when only side can make offers.

2 Literature Review

Our experimental design corresponds to the standard model of two-sided matching markets, see Roth and Sotomayor (1990); which belongs to the realm of cooperative game theory. The model provides clear predictions: outcomes coincide with the core of the market, the set of stable matchings. In that respect, our experimental results provide a strong experimental validation of the theory underlying the stability notion.

Several authors propose particular dynamic decentralized processes by which one-to-one matchings are created; see, e.g., Haeringer and Wooders (2011), Ferdowsian, Niederle, and Yariv (2023), and Pais (2008). These papers usually impose some structure on the process by which offers are made and accepted. The main focus of the decentralized matching literature is on the identification of conditions under which stability is likely to arise in equilibrium. Complete information of the prevailing preferences, as in our experiments, allows for stability to emerge in equilibrium, while more stringent demands on preferences and equilibrium selection are required for stability to be the unique prediction. To the extent of our knowledge, the literature is silent on the selection of stable matchings when multiple ones exist in a market.

While the experimental literature on matching markets has grown rapidly in recent years (Hakimov and Kübler, 2021), there are only a few studies of decentralized markets. Nalbantian and Schotter (1995) analyze several procedures for matching with transferable utility, decen-

tralized matching among them, where agents have private payoff information. Nalbantian and Schotter (1995) include private negotiations between potential match partners. Offers in our treatment are private as well; only accepted offers become public, but they are non-binding. Kagel and Roth (2000) analyze the transition from decentralized matching to centralized clearinghouses when market features lead to inefficient matching through unraveling. Haruvy and Ünver (2007) study the effects of repetition on a two-stage decentralized matching protocol. Agranov, Dianat, Samuelson, and Yariv (2023) allow for transfers in small decentralized markets that follow protocols similar to ours, where information about preferences is either complete or incomplete. Transfers and incomplete information make stability elusive, particularly when preferences are submodular. Pais, Pintér, and Veszteg (2020) study the effects of information and costly-offer frictions on outcomes and find that enough frictions may make stability difficult to achieve. Finally, Niederle and Roth (2009) also look at an incomplete information setting in which one side of the market (the firms) makes offers to the other side (the workers) over three experimental periods. They study the effects of offer structure on the information that gets used in the final matching and consequent market efficiency.[4]

There is also a methodological link between the current paper and some of the experimental work studying financial markets and general equilibrium predictions in the lab, see for instance Smith (1962), Plott and Smith (1978), or the survey in Chapter 6 of Kagel and Roth (1995). As in our paper, the underlying predictions of general equilibrium theory pertain to outcomes, and by and large shine through in experiments; this despite the precise dynamics leading to these outcomes not having been imposed by the experimenters.

[4]There is a growing experimental literature studying centralized matching systems, e.g., Bergstrom, Bergstrom, and Garratt (2013), Harrison and McCabe (1996), Chen and Sönmez (2006), Pais and Pintér (2008), Echenique, Wilson, and Yariv (2016), Featherstone, Mayefsky, and Sullivan (2022), and Featherstone and Niederle (2016).

3 Theoretical Preliminaries

We start by reviewing the one-to-one two-sided matching model, and the theoretical results that are pertinent to our paper.

Let F and C be disjoint, finite sets. We call the elements of F "foods" and the elements of C "colors." We use the language of foods and colors in our experimental design, but these sets can stand for firms and workers in labor markets, men and women in heterosexual marriage markets, etc. A *matching* is a function $\mu : F \cup C \to F \cup C$ such that for all $f \in F$ and $c \in C$,

(i) $\mu(c) \in F \cup \{c\}$,

(ii) $\mu(f) \in C \cup \{f\}$,

(iii) $f = \mu(c)$ if and only if $c = \mu(f)$.

Whenever $a \in F \cup C$ is unmatched under μ, we write $\mu(a) = a$; when f and c are matched under μ, then $c = \mu(f)$ (and $f = \mu(c)$).

A *preference relation* is a linear order (a complete, transitive, and antisymmetric binary relation). In particular, we assume preferences are *strict*. A preference relation for a food $f \in F$, denoted $P(f)$, is understood to be over the set $C \cup \{f\}$, with f representing the possibility of being unmatched. Similarly, for $c \in C$, $P(c)$ denotes a preference relation over $F \cup \{c\}$. For simplicity and consistency with our experimental design, we ignore individual rationality by assuming that each food (color) prefers any color (food) over remaining unmatched. A *preference profile* is a list P of preference relations for foods and colors, i.e.,

$$P = \Big((P(f))_{f \in F}, (P(c))_{c \in C} \Big).$$

Denote by $R(f)$ the weak version of $P(f)$. That is, $c' \, R(f) \, c$ if either $c' = c$ or $c' P(f) c$. The definition of $R(c)$, the weak version of $P(c)$, is analogous.

Fix a preference profile P. We say that a pair (c, f) *blocks* μ if $c \neq \mu(f)$, $c\, P(f)\, \mu(f)$, and $f\, P(c)\, \mu(c)$. In words, (c, f) is a blocking pair if c and f prefer to be matched to one another over their assigned matches under μ. In the absence of individual rationality, a matching is *stable* if there is no pair that blocks it. Denote by $S(P)$ the set of all stable matchings.

Theorem (Gale and Shapley, 1962)**.** *$S(P)$ is non empty, and there are two matchings μ_F and μ_C in $S(P)$ such that, for all $f \in F$, $c \in C$, and $\mu \in S(P)$,*

$$\mu_F(f)\, R(f)\, \mu(f)\, R(f)\, \mu_C(f),$$
$$\mu_C(c)\, R(c)\, \mu(c)\, R(c)\, \mu_F(c).$$

The matchings μ_F and μ_C coincide when the market has a unique stable matching. The matching μ_F is called *food optimal*, while μ_C is called *color optimal*, and both are known as the *extremal stable matchings*. The matching μ_F is preferred by all foods to any other stable matching, and all colors prefer any stable matching to μ_F. Analogously for μ_C. The proof of the Gale-Shapley Theorem is constructive, and uses the Deferred Acceptance (DA) algorithm to identify one of the extreme matchings, μ_F or μ_C. Beyond its theoretical role in establishing existence, DA is the algorithm often used in centralized markets. For instance, the National Resident Matching Program uses a variation of it (Roth and Peranson, 1999).

The set of *stable partners* of an agent $a \in F \cup C$ is the set of agents that are matched to a under some stable matching, i.e., $\{\mu'(a) \mid \mu' \in S(P)\}$.[5] Likewise, a pair $(f, c) \in F \times C$ is said to be a *stable pair*, or a *stable match*, if f and c are matched under some stable matching.[6] Therefore, in a stable matching, all matches are stable, and all agents are matched to a stable partner. Nonetheless, in unstable matchings, it might be reasonable for some agents who are

[5]Stable partners are also referred to as "achievable partners" in the literature (e.g., Roth, 1984).

[6]Note the distinction between a *match* and a *matching*. A match refers to a pair (f, c) who are matched. A matching refers to the function μ describing all the matches in a market.

not stable partners to be matched: from an individual perspective, it makes sense to form matches in which no agent has a blocking partner, regardless of whether the match is stable. Similarly, while in markets with a unique stable matching every agent has a unique stable partner, the number of stable partners may differ among agents, and, in particular, may be less than the number of stable matchings when not unique.

Consider a preference profile P for which $S(P)$ has an odd number K of matchings, and denote the partners of agent a in each of these matchings by $a_1, \ldots, a_K$ (which may not all be distinct). A *median stable matching* is a matching $\mu \in S(P)$ such that, for all agents $a \in F \cup C$, $\mu(a)$ is a's median partner among a's stable partners under $P(a)$. That is, $\mu(a)$ occupies the $\frac{K+1}{2}$-th place in a's preference among $a_1, \ldots, a_K$. We refer to $\mu(a)$ as the *median stable partner* of a.

In general, median stable matchings are guaranteed to exist, see Teo et al. (2001)—in fact, they also exist when K is even. Median stable matchings present a compromise between the two sides of the market. Interestingly, there are no known simple algorithms that generate median stable matchings. Certainly, one can search for all stable matchings of a market and then identify a median one. From a computational perspective, however, this can potentially be quite demanding as the problem of finding all stable matchings is computationally hard (see Gusfield and Irving, 1989, for general references; Irving and Leather, 1986, show that determining the number of all stable matchings is generally #P-complete; while Cheng, 2008, shows that finding the median stable matching is hard). These results contrast with the problem of finding a color- or food-optimal stable matching, which can be done in polynomial time by using DA.

The notion of stability, as well as the ranking of the different stable matchings, are ordinal in nature. In particular, the theory does not allow for refined predictions on the basis of *how much* agents prefer certain partners to others.

4 Experimental Design

Our experimental design corresponds to a decentralized one-to-one, two-sided market with no transfers. The two sides are termed colors and foods, and contain the same number of participants each. In each round, each participant is randomly assigned a role: "red," "blue," etc. if a color; "apple," "banana," etc. if a food. A participant can match with one and only one participant from the other side of the market, each match resulting in a potentially different monetary payoff. All participants observe all potential payoffs from a numerical matrix on the experimental interface. If a participant is unmatched, they earn a payoff of 0.[7]

We implemented three *main treatments* with 8 participants on each side of the market, and two additional *auxiliary treatments* intended to check for robustness with respect to bargaining power and market size. In our main treatments, over the course of the experiment, participants are free to propose a match to anyone on the other side of the market. At any point in time, participants observe all current matches through a panel of the experimental interface. Importantly, participants can make an offer while (tentatively) matched, and offers can be made to any member of the opposing side of the market, including participants who are already matched. If a matched agent accepts a new offer, their existing match is undone. When receiving an offer, a participant has 10 seconds to respond. Each market ends after 30 seconds have transpired without any new offer.

We ran our main treatments across five experimental sessions, each consisting of 2 practice rounds and 10 real rounds.[8] Participants took part in a different market each round. Markets

[7]The instructions and the set of payoff matrices we used are available at the following link: https://sites.google.com/view/decentralized-matching.

[8]Four of the sessions included 32 participants, so we ran two markets simultaneously, for a total of 20 non-practice experimental rounds per session. In the remaining session, we only had 16 participants for a total of 10 non-practice experimental rounds. Hence, in total, we ran 90 experimental rounds across five sessions. Due to software malfunctions, 5 markets were not presented as intended to participants. We drop these from the data. In total, we have 85 rounds of experimental market data for our main treatments. Including data from the dropped rounds does not alter our results qualitatively.

were designed with two objectives in mind. First, in order to see whether cardinal representations of preferences matter, we implemented multiple cardinal representations of the same ordinal markets. Specifically, for any ordinal preference—say, red prefers apple to banana—there are many ways by which these preferences can be presented cardinally. For example, red receiving \$50 and \$10, or \$5 and \$4, from matching with apple and banana, respectively, would both correspond to the same ordinal ranking. Our design entailed 6 underlying ordinal descriptions of markets with 17 different cardinal representations. Second, in order to study the endogenous selection of stable matchings, we designed our experimental markets so that all participants had either one, two, or three possible stable partners. When all agents have only one stable partner, the market as a whole has a unique stable matching. For each fixed number of stable partners (ranging from 1 to 3), we use several markets differing in market participants' ordinal and cardinal preferences.[9]

The top panel of Table 1 summarizes our three main experimental treatments. The following is a general description of the match payoffs we used.

Unique stable matching. We used 4 different ordinal markets that have a unique stable matching: assortative preference markets, where participants on each side of the market agree on the ranking of participants on the other market side; markets with assortative preferences on one side, where only the members of one side of the market are in agreement; a market including a fully egalitarian matching, providing all agents the same payoff, which is unstable; and a "generic" market with a unique stable matching without agents agreeing on the ranking of others on either side. Each of the first three markets was implemented via one cardinal representation. The last, generic market was implemented via four cardinal representations. In two of these we varied how aligned the interests were across the market: that is, if μ is

[9]The mapping from rounds to payoff matrices varied; in some sessions we used the same markets across rounds, and in others we shuffled them. We see no strong evidence of learning in either.

Table 1: Description of treatments

Agents per side	*Proposing sides*	*#Exp. Mkts.*	*#Ordinal Mkts.*	*#Cardinal Mkts.*	*#Sessions*	*#Particp's*	*#Stable matchings*	*Avg. #stable partners*
Main treatments								
*Unique stable matching**								
8	2	30	4	7	5	144	1	1.00
Two embedded 4-by-4 markets†								
8	2	35	1	6	5	144	2×2	1.75
5 stable matchings & 3 stable partners								
8	2	20	1	4	3	80	5	3.00
Auxiliary treatments								
Unilateral offers								
8	1	9	1	2	2	80	1	1.00
8	1	19	1	4	2	80	2×2	1.75
8	1	15	1	3	3	112	5	3.00
Large markets‡								
15	2	8	3	3	3	90	1	1.00
15	2	4	2	2	2	60	3	2.93
Total		140	11	22	12	346		

Notes. The table reports all experimental markets (rounds within a session) across our three main treatments (*unique stable matching*, *four-by-four*, and *five stable matchings*), and two auxiliary treatments (*unilateral offers* and *large markets*). For each treatment, the table reports: the number of agents on each side of the market, the number of sides that can make match offers, the number of experimental markets ran in this treatment, and the ordinal and cardinal representations used across them, the number of sessions in which at least one market in the corresponding treatment was played, the number of participants who played in at least one market in this treatment, and the number of stable matchings and the average number of stable partners in the markets in the treatment. The last row reports the totals across all treatments.

*One of the ordinal markets (out of 4) had a salient egalitarian unstable matching, which was used in 4 experimental markets (out of 30). Eight experimental markets (out of 30) were run with two alternative ordinal payoff matrices: four of them had one-sided aligned preferences, and the other four had two-sided aligned preferences.

†In each of the two embedded markets, there were eight agents, four on each side. Three agents on each side had two stable partners, and one agent had one, for an average of 1.75 stable partners per agent.

‡In the large markets with a unique stable matching, five (out of 8) had one-sided aligned preferences. In the large markets with multiple stable matchings, in one of the markets every agent had three stable partners, while in the other one some agents had two stable partners and most had three (avg = 2.87).

the stable matching, we computed the correlation of the vectors $(u_f(\mu))_{f\in F}$ and $(u_{\mu(f)}(\mu))_{f\in F}$, where $u_a(\mu)$ denotes the payoff of agent $a \in F \cup C$ in the matching μ. We created one market in which the correlation was -0.9 and one where it was 0.9. We used two additional cardinal representations: one in which, for each agent, the difference in utilities between matching with the agent's k'th and $k+1$'th choices was 20¢ and one in which these marginal differences were 70¢. Altogether, we used 7 different cardinal markets with a unique stable matching.

Two embedded four-by-four markets. We used one ordinal market in which each agent had two possible stable match partners. These were constructed so that there were two 4×4 embedded markets, where any agent within a submarket preferred to match with anyone from that submarket over anyone from the other. We varied the overall utilitarian efficiency of each matching, the utilitarian efficiency of foods relative to colors from each matching, the distribution within each matching,[10] as well as the marginal loss for either side of the market from switching from their more preferred stable matchings to their less preferred ones (higher for foods or for colors). Overall, we used 6 cardinal markets of this sort.

Five stable matchings and three stable partners. We used one ordinal market represented cardinally in 4 ways: one in which the marginal differences between utilities derived from matching with one's k'th and $k+1$'th most preferred partners was 20¢, one in which it was 70¢, one in which for foods it was 20¢ and for colors 70¢, and one in which these differences were 20¢ for both market sides, but colors' payoffs were all shifted up by \$1. In these markets, while each individual has precisely three distinct stable partners, there are five different market-wide stable matchings.

[10] Since egalitarian motives appear frequently in experiments, we were concerned that some form of altruism would be driving our results. We therefore designed payoffs so that in some treatments, fully egalitarian matchings were unstable (see the description of our markets with a unique stable matching above). Furthermore, we included cardinal representations in which certain *stable* matchings were more egalitarian than others.

In addition to our three main treatments, with 8 participants on each side who could all make match offers, we implemented two auxiliary treatments that differed in either the relative bargaining power each market side had or the market size.[11] The bottom panel of Table 1 summarizes our two auxiliary experimental treatments.

Unilateral offers. We ran treatments with 8×8 markets and payoffs as in our main treatments where only foods could make offers. Otherwise, markets operated as in our main treatments.

Large markets. We also ran several treatments with larger markets, containing 15 participants on each side. We concentrated on markets with either a unique stable matching or 3 stable matchings. We used 3 distinct markets with a unique stable matching: assortative preferences on one side, assortative preferences with a fully egalitarian unstable matching, and preferences that were not assortative on either side. We used two markets with 3 stable matchings (in one market, all agents had three stable partners; in the other one, most agents had three stable partners, while some had two).

All sessions took place at the California Social Science Experimental Laboratory (CASSEL), using a modification of the multi-stage software. All participants were UCLA undergraduates and each participant participated in only one session. The average payment per

[11] We ran four sessions with unilateral markets. In three sessions there were 32 participants, so we ran two markets at the same time, each consisting of 16 participants, 8 per side. In the remaining session we only had 16 participants. However, we exclude the first session from the analysis, which had 32 participants, due to a software glitch that affected how the termination rule was implemented. While none of our results change qualitatively if we include this session in the analysis, some markets terminated before 30 seconds of inactivity and some after in this session. This glitch was fixed between the first and the second session. Additionally, for a similar reason as with some markets in our main treatments (see footnote 8), some markets were not presented as intended to participants, so we drop them from the analysis. In total, this adds up to 43 experimental markets in our unilateral treatments. We ran three sessions with large markets, each with 30 participants, 15 on each side. Since large markets took longer to converge than our main and unilateral treatments, we ran fewer rounds. In two sessions, we ran five non-practice rounds, and in one, we ran two, for a total of 12 experimental markets in our large treatment.

participant was \$40 in our main treatments, \$33 in the treatments in which only foods were able to make offers, and \$24 in the large 15×15 market treatments.[12] All of these were combined with a \$5 show-up fee.

5 Market Outcomes

Three main findings emerge from our experiments regarding final outcomes. First, across our treatments, *most outcomes are stable*. Furthermore, market outcomes that are unstable are close to stable; they are close in the sense that both the number of blocking pairs, as well as the unrealized payoff gains from not forming them, are small. Second, in our treatments with three stable partners, *most agents are matched to their median stable partner*. Even when we "handicap" one side of the market, so that it cannot make any offers, we continue to see the median as the modal outcome. Last, *cardinal representations of preferences affect the particular stable matchings that get selected*. Specifically, higher cardinal incentives to colors make the color optimal matching more likely to be selected; similarly for foods.

In all our treatments, learning across rounds did not appear to have significant effects on either outcomes or behavior. Therefore, we present aggregate results across rounds. We return to the within-round dynamics underlying our final outcomes in Section 8.

5.1 Stability in Experimental Markets

Virtually all agents match through our markets' operations, and a large fraction of markets culminate in a stable matching, as shown in Table 2. The table summarizes the overall outcomes in our main treatments. Over 99% of agents are matched when markets terminate. Furthermore, 88% of markets are fully stable: no agent in the whole market has a blocking

[12] Standard deviations were \$3, \$13, and \$9, respectively.

partner. Markets with three stable partners exhibit slightly fewer market-wide stable outcomes, 75%. As a measure of stability at the match or pair level, we check the proportion of matched pairs in which no member was part of a blocking pair. In over 95% of matches, no member has a blocking partner, with little variation across markets. The treatment with the lowest proportion of blocking partners is the one in which every agent has three stable partners. Even in this case, in 90% of matches, no agent can form a blocking pair.[13]

Not all our markets reached full stability, but markets that were not fully stable were close to stable. Even in unstable markets, Table 2 indicates that, in the majority of matched pairs (61%), no agent has a blocking partner. Indeed, only 20% of agents have a blocking partner when markets finalize (compared to 2% of agents across all markets). The left panel of Figure 1 presents the empirical cumulative distribution functions (CDFs) of the maximal number of disjoint blocking pairs across all treatments, main and auxiliary.[14] The figure contains the distribution of blocking pairs pertaining to *markets in which outcomes were not fully stable.*[15] In our main treatments, most outcomes are close to stable: most markets culminating in unstable outcomes have only one or two disjoint blocking pairs.

We can also use payoffs to measure distance to stability. For each agent, we compare the payoff they received from their final match, with the payoff they would have received had they matched with their most preferred blocking partner at the end of the market. That is, we compute the maximum immediate losses agents incur from not exploiting blocking opportunities. Agents in markets that did not culminate in a stable matching lost, on average, 12.8% of their final market payoff, equivalent to around 28¢. By contrast, the same average

[13]The use of color and food labels in our markets did not seem to have any effect. For example, if one considers banana and mango to be associated with yellow, apple and cherry with red, and kiwi and pear with green, there is no significant increase in the corresponding matches relative to any other classification.

[14]In our main treatments, when no agents are matched, the maximal number of disjoint blocking pairs is 8, the number of potential disjoint pairs. When a stable matching is in place, there are 0 blocking pairs.

[15]Including all markets in the figure would not allow to visualize the fine-grained distribution of blocking pairs, as the prevalence of stable matchings implies a large spike at zero blocking pairs.

Table 2: Outcomes in main treatments

	Unique stable matching	*Two embedded 4-by-4 markets*	*5 stable matchings & 3 stable partners*	*All main treatments*
#Mkts. with stable matching / #Mkts	27 / 30	33 / 35	15 / 20	75 / 85
% Mkts. with stable matching	90.00	94.29	75.00	88.24
Avg. % pairs w/o blocking partners (BPs)	95.32	98.65	90.00	95.44
Avg. % pairs w/o BPs | unstable matching	53.24	76.39	60.00	61.25
Avg. % agents with ≥ 1 *BP*	2.50	0.71	5.00	2.35
Avg. % agents with ≥ 1 *BP | unstable matching*	25.00	12.50	20.00	20.00
Avg. # of BPs per agent | ≥ 1 *BP*	1.35	1.00	1.23	1.22
Avg. % unmatched agents	0.42	0.36	0.00	0.29
Avg. % unmatched agents | ≥ 1 *BP*	4.76	25.00	0.00	6.43

Notes. The table reports the following final outcomes for each of our *main treatments*: (i) number and (ii) percent of final matchings that are stable; avg. number of final pairs (matches) in which no agent has a blocking partner across (iii) all markets and (iv) markets with an unstable final matching; average percent of agents with at least one blocking partner across (v) all markets and (vi) markets with an unstable final matching; (vii) average number of blocking partners per agent across agents with at least one blocking partner; average number of unmatched agents (viii) among all agents, and (ix) among agents who have at least one blocking partner.

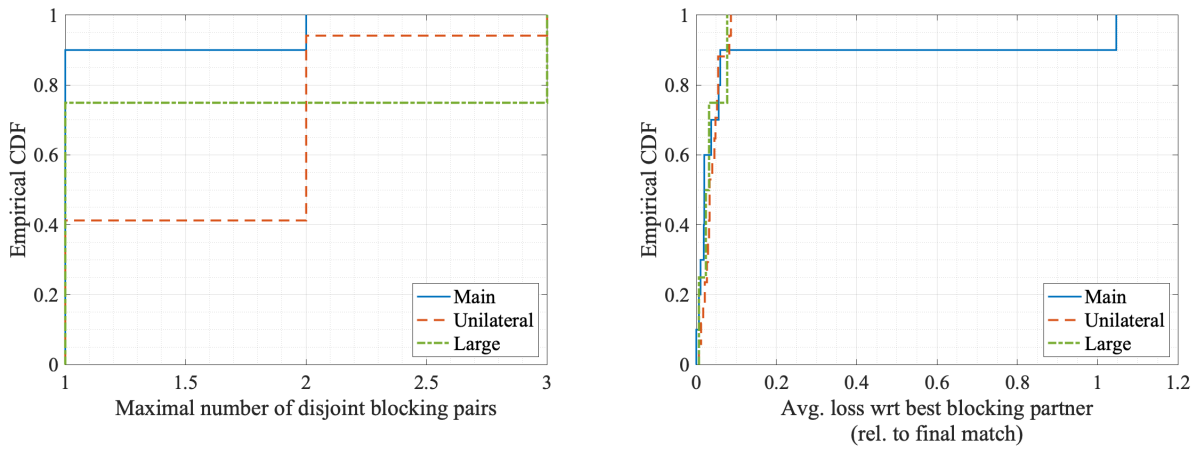


Figure 1: Distance to stability in unstable markets

across all markets is 1.5%, equivalent to around 3¢. The right panel of Figure 1 reports the CDFs of this measure across all unstable markets. Even in unstable markets, most agents had very small payoff losses due to instability. Indeed, the median loss is equivalent to 2.1% of the final payoff, which comes down to 10¢.[16]

In terms of the duration of market interactions, convergence to the stable matching was rapid. On average, in our main treatments the final matching was reached after 1.62 minutes (sd = 1.00). While markets terminated 30 seconds after the last offer was made, the last offer usually took place after the final matching was reached (i.e., the last offers are generally not accepted and do not change the matching). On average, the last offer took place after 2.28 minutes (sd = 1.35), and markets terminated 30 seconds afterwards, with 89% of markets terminating before the five minute mark. We return to the dynamics underlying these observations in Section 8.

[16]In the Online Appendix, we plot CDFs for alternative measures of distance to stability across the distinct treatments: total number of blocking pairs, and the average immediate loss with respect to the best blocking partner (i) in absolute terms, (ii) relative to average market payoffs, and (iii) relative to the average payoff across random matchings (which with high likelihood are unstable). Roughly speaking, all measures point to the same conclusion: the majority of agents in markets that did not reach full stability had few blocking partners and did not incur substantial immediate payoff losses (Echenique, Robinson-Cortés, and Yariv, 2024).

5.2 The Emergence of Median Stable Matches

The median stable matching has strong drawing power, and cardinal representations of preferences affect outcomes, as shown in Table 3. The top panel of the table reports the distribution of final matchings and matched pairs across markets in our main treatments in which each agent has three stable match partners. A key difference between the markets is the level, and marginal differences, of payoffs within each side of the market. We use the notation of "x-y marginals" to denote a market in which the marginal difference in utilities between one partner and the next-best partner was x cents for foods and y cents for colors. For example, in markets with 20-70 marginals the payoff differences to the next-best partner was 20¢ for foods and 70¢ for colors. In the 20-20 marginals market with 100 color shift (labeled as 20-20$_{+100}$ in the table), a 100¢ (or \$1) was added to the color payoffs of the 20-20 market.

The median stable matching is the modal outcome: 80% of stable matchings correspond to the median stable matching, entailing *all* participants being matched to their median stable partner. Moreover, *none* of these markets converged to an extremal stable matching, food-optimal or color-optimal. Recall that the markets in which each agent has three stable partners have five market-wide stable matchings. Hence, there are two stable matchings that are neither the median nor extremal. We emphasize that, in these markets, a median matching had to be "discovered" by participants: it is quite hard to infer the set of stable matchings from looking at the payoff tables (an 8×8 table with 128 numerical entries); almost impossible without using a computer program. We return to the dynamics underlying the convergence to the median in Section 8.

Shifting attention to individual match-level outcomes, we observe similar patterns. 77% of the final matched pairs that were stable correspond to median stable partners. From the perspective of individual behavior, agents care about stable partners, not market-wide matchings. It is remarkable that the vast majority of stable matches—which comprise 97%

Table 3: Selection of stable matching and cardinal effects

Marginal Utility Differences (foods–colors)	*20–20*	*20–20$_{+100}$*	*20–70*	*70–70*	*All*
Main: markets with three stable partners					
# markets	5	5	5	5	20
Market-level outcomes (matchings)					
% *stable*	60	40	100	100	75
% *median* | *stable*	67	100	60	100	80
% *non-extremal* | *stable*	100	100	100	100	100
% *food-optimal* | *stable*	0	0	0	0	0
% *color-optimal* | *stable*	0	0	0	0	0
Individual-level outcomes (matches)					
% *stable*	95	93	100	100	97
% *median* | *stable*	64	63	80	100	77
% *food-optimal* | *stable*	11	0	0	0	3
% *color-optimal* | *stable*	24	37	20	0	20
Unilateral offers (foods propose): markets with three stable partners					
# markets	0	5	5	5	15
Market-level outcomes (matchings)					
% *stable*	–	20	20	60	33
% *median* | *stable*	–	0	100	100	80
% *non-extremal* | *stable*	–	100	100	100	100
% *food-optimal* | *stable*	–	0	0	0	0
% *color-optimal* | *stable*	–	0	0	0	0
Individual-level outcomes (matches)					
% *stable*	–	93	88	95	92
% *median* | *stable*	–	49	77	83	70
% *food-optimal* | *stable*	–	14	17	17	16
% *color-optimal* | *stable*	–	37	6	0	14

Notes. The table reports (i) % of markets which final matching corresponds to the median, non-extremal, food-optimal, or color-optimal stable matching, and (ii) the percentage of final matches that are food-optimal, color-optimal, or median stable matches (i.e., part of a stable matching). The table reports results for all experimental markets with five stable matchings, in the *main treatments* (top panel) and *unilateral offers* (bottom panel). The results are disaggregated into cardinal treatments, according to the marginal utility of each side (payoff difference from less preferred to more preferred partner). The second column corresponds to markets in which utility differences are 20¢ on each side, with the payoffs of colors shifted upwards by $1.

of all final matched pairs—correspond to the median. That is, in the vast majority of cases, agents ended up in stable matches in which neither of the two parties were matched to their most preferred stable partner: no side of the market systematically got their way. Of the remaining stable matches, the majority were color-optimal stable matches (20% out of 23%), and the rest were food-optimal matches.[17]

Stability depends on agents' ordinal preferences alone, not on their cardinal representation. Of course, our experimental participants receive different monetary payments depending on who they match with, and one might expect the different monetary magnitudes to play a role. It is interesting, then, that in our data, different cardinal representations seem to affect more *which* stable matching gets selected than whether a stable matching is achieved, as shown in Table 3. Participants on the side of the market (foods or colors) having "more to lose" tend to partner with their most preferred match at higher frequencies (significant at any conventional levels). We return to the effects of cardinal preference representations on outcomes in Section 7.

6 Robustness to Bargaining Power and Market Size

We investigate the robustness of our results along two dimensions: the bargaining power afforded to each side of the market, and the market's size. As for bargaining power, in our *unilateral design*, only one side of the market is allowed to make offers. Specifically, only foods are able to make offers to colors. The effect of market size is studied by means of our *large-market design*, which resembles our main treatments but with 15 agents on each market side. Table 4 summarizes our findings.

[17]The preferences in these markets, both ordinal and cardinal, were not symmetric across the two sides, which explains the imbalance.

Table 4: Outcomes in unilateral offers and large markets

	Unilateral offers	*Large markets*	*Main*
#Mkts. with stable matching / #Mkts	26 / 43	8 / 12	75 / 85
% Mkts. with stable matching	60.47	66.67	88.24
Avg. % pairs w/o blocking partners (BPs)	83.30	93.33	95.44
Avg. % pairs w/o BPs | unstable matching	57.76	80.00	61.25
Avg. % agents with ≥ 1 *BP*	9.60	3.61	2.35
Avg. % agents with ≥ 1 *BP | unstable matching*	24.27	10.83	20.00
Avg. # of BPs per agent | ≥ 1 *BP*	1.19	1.04	1.22
Avg. % unmatched agents	0.87	0.00	0.29
Avg. % unmatched agents | ≥ 1 *BP*	3.00	0.00	6.43

Notes. The table reports the following final outcomes for markets with *unilateral offers* and in *large markets*: (i) number and (ii) percent of final matchings that are stable; avg. number of final pairs (matches) in which no agent has a blocking partner across (iii) all markets and (iv) markets with an unstable final matching; average percent of agents with at least one blocking partner across (v) all markets and (vi) markets with an unstable final matching; (vii) average number of blocking partners per agent across agents with at least one blocking partner; average number of unmatched agents (viii) among all agents, and (ix) among agents who have at least one blocking partner. For ease of comparison, we replicate the rightmost column of Table 2, corresponding to markets in our *main treatments*, in the last column.

Unilateral Offers. A natural possible explanation for the prevalence of the median stable matches is that, in our experiments, participants on both sides of the market could make offers. In a sense, they had equal bargaining power. In contrast, in Gale and Shapley's DA, only one side makes offers while the other side decides which offers to accept. DA produces the optimal matching for one side, so it is possible that allowing both sides of the market to make offers is at the root of the frequent median matches we observe.[18] This conjecture is interesting, particularly in view of the fact that several real-world matching markets that operate in a decentralized manner allow one side greater, if not sole responsibility for making offers: e.g., the job market for academics, the marriage market in certain cultures, etc. To explore the effects of bargaining or proposal power, in our *unilateral offers* treatments, only foods were allowed to propose matches. While stability rates are not as high as in our main treatments, we find that outcomes are mostly stable and also correspond to median outcomes.

With unilateral offers, outcomes are qualitatively similar to those in our main treatments, as seen in the bottom panel of Table 3 and the first and third columns of Table 4. In these treatments, 99% of participants are matched. In 83% of the final matches, no agent is in a blocking pair, compared with 95% in the markets in which both sides can make offers. At the market level, 60% of the markets culminate in a stable matching, compared with 88% in the main treatments (see Table 2).[19] As Figure 1 shows, the observed unstable matchings are close to stable, although not as close as in our main treatments. Across the markets that fail to reach stability, on average, around a quarter of agents has at least one blocking partner, compared to a fifth in the main treatments. Similar to the main treatments, agents who have blocking partners, have 1.2 such partners on average. The convergence time in unilateral

[18] In Section 8, we contrast the dynamics we observe in the experiment with a two-sided version of DA (Dworczak, 2021), and still observe that median stable matches are more prone to occur in our experiment than when both sides propose in DA.

[19] The corresponding percent of final matches in which no agent is in a blocking pair (final matchings that are stable) for markets with a unique stable matching, two embedded four-by-four markets, and with five stable matchings, in unilateral markets, are 94% (78%), 91% (74%), and 67% (33%).

markets is similar to our main treatments: on average, it takes 1.64 minutes to reach the final matching (sd = 1.04), the last offer takes place after 1.69 minutes (sd = 1.10), and markets terminate 30 seconds afterwards, with 95% of markets terminating before 5 minutes.

Regarding selection, two important observations emerge. First, in markets with four stable matchings (two per each embedded submarket), the stable matchings favored by foods, the proposers, are much more frequent. In our main treatments, when both sides can make offers, 27% of markets converge to the food-optimal stable matching, 21% to the color-optimal, and the rest to a non-extremal stable matching.[20] By contrast, in markets in which only foods can make offers, 86% of the markets converge to the food-optimal stable matching, and none to the color-optimal stable matching (the rest to one of the two non-extremal ones). The same finding emerges when looking at individual matches across all markets, not only those reaching stability. When both sides can make offers, among all the agents who have two stable partners (75%), 54% of final matches are food-optimal, and the rest are color-optimal. With unilateral offers, a higher proportion of the final matches are food-optimal (90%).

For the markets with two embedded four-by-four markets, results are consistent with extremal outcomes, but this is no longer true in markets with three stable partners. Indeed, for markets in which each agent has three stable partners, median outcomes are again the most common. As the bottom panel of Table 3 shows, 70% of final matches that are stable correspond to median stable matches when only foods are able to make offers, compared to 77% when both sides can do so. Similarly, among markets that reach full stability, 80% reach the median stable matching, same as when both sides can make offers. At the individual level, there is an increase in the number of matches that are optimal for foods, the proposing side, from 3% to 16%, but this comes mainly at the expense of matches that are optimal for colors, the receiving side. Nonetheless, cardinal incentives matter for the distribution of outcomes as

[20] In these markets, non-extremal stable matchings are the ones in which one of the 4×4 submarkets is at the food-optimal stable matching, and the other submarket is at the color-optimal one.

they do in our main treatments: when foods have more to lose from forgoing higher-ranked partners, color-optimal matches are less common.

Market Size. The outcomes in our large-market treatments are similar to those in our main treatments, as seen in the second and third columns of Table 4. Our large markets involve 30 participants, which is quite large as experimental markets go. They do not, of course, come near the size of some real-world matching markets—such as, for example, the medical residents market—but it is a comfort that duplicating the size of our the markets in our main treatments does not upset our main results.

In our large-market treatments, in 93% of the final matches, no agent was part of a blocking pair. In terms of market-wide outcomes, while 67% of matchings are fully stable, the number of blocking pairs within unstable markets is, again, very small, see Figure 1. Our large experimental markets exhibit at most 3 disjoint blocking pairs, and in 75% of cases just one. Indeed, as Table 4 reports, even in markets that fail to reach full stability, the vast majority of agents have no blocking partners (89%, as compared to 96% across all markets, unstable and stable). Agents who have blocking partners have, on average, just one; and there are no agents left unmatched. Interestingly, the seemingly more complex markets with three stable matchings always culminate in a stable matching in our large 15×15 markets.

As one might expect, convergence is slower in large markets. On average, it takes 4.47 minutes for large markets to reach a final matching (sd = 3.16). The last offer takes place after 4.86 minutes (sd = 3.06), and markets terminate 30 seconds afterwards, with half terminating before the five-minute mark.

As for the selected stable match partners, the results are similar to those in our main treatments, although arguably more extreme. In the large market treatments, *every* agent who has three stable partners matches with their median stable partner. Hence, the market-wide stable matchings are exclusively median stable matchings.

7 Cardinal Incentives and Social Preferences

A matching delivers cardinal (monetary) payoffs to market participants, and these payoffs may be more or less fair: they may be more or less equal across participants. The degree of fairness of a matching can reasonably be expected to influence whether it is chosen: a taste for egalitarian outcomes is well-documented in experimental economics (the literature has suggested different types of social preferences, for surveys see, e.g., Chapter 4 in Kagel and Roth, 2020, and Fehr and Gächter, 2000). Fairness considerations may, in particular, be important given our finding that median matches are very common.

Our design entailed several markets with identical ordinal preference profiles, but different cardinal utility representations. Focusing on such variations, we have already seen that the side that faces steeper cardinal incentives is more likely to achieve its optimal stable matching. We now inspect the impact of payoff *distributions*. We show that equality of payoffs across match partners makes a matching more likely, *as long as it is stable.*

We first analyze markets with a salient egalitarian, but unstable, matching. In these markets, one unstable matching entails identical payoffs to all market participants, and comparable utilitarian welfare to that generated by the unique stable matching.[21] Participants consistently fail to select the unstable egalitarian matching. In our main treatments, none of these markets end up in the unstable, albeit more equal, matching outcome.

Our larger markets, in which more payoff variation could be introduced, generate similar findings. For example, we ran one 15×15 market in which there was an unstable matching under which all agents received exactly \$4; there was a unique stable matching in which the average payoff was also \$4, which was much more unequal.[22] While the final outcome in these

[21] By specifying identical payoffs to all market participants, we avoid the challenge of identifying the individuals that are relevant for participants' social preferences.

[22] Its Gini index was 26. For a country's income distribution, this is in the Scandinavian range, but it appears starkly unequal compared to perfect egalitarianism with no utilitarian efficiency loss.

markets is not always stable—in some instances a few blocking pairs remain—participants clearly avoid the egalitarian unstable matching.

Second, we use the markets with two stable matchings to assess the effects of cardinal utilities on the selection of food- and color-optimal stable matchings.[23] Markets are more likely to converge to the extremal stable matching, food- or color-optimal, which has the lowest payoff variation. That is, when the dispersion of payoffs is relatively high in, say, the color-optimal matching, markets tend to achieve the food-optimal matching, and vice-versa. Specifically, we compute the *coefficient of variation* of agents' payoffs in a given matching: the standard deviation of payoffs divided by their mean, which is a "scale free" measure of the dispersion in payoffs across agents. We then compute the ratio of the coefficient of variation at the food-optimal stable matching over that at the color-optimal stable matching. Markets with a high ratio are those in which the payoff variation in the food-optimal stable matching is high, relative to that in the color-optimal stable matching. When the ratio is above the median ratio in our data, the food-optimal matching obtains 11.7% of the time, and the color-optimal 35.3%. When it is below the median, they obtain 43.8% and 6.3%, respectively. The difference between these values is significant at any conventional level of confidence. The implication is that when the variance in payoffs at the color-optimal matching is relatively high, we tend to get more food-optimal outcomes, and vice-versa.

To summarize, our results suggest that egalitarian, or fairness, considerations play a role in selecting outcomes, but they are not so strong as to trump stability.[24]

[23] Recall that each of these markets was constructed by embedding two smaller 4×4 markets, which allows us to gain more payoff variations. In these markets, 99% of final matches are stable. In addition, the embedding of the "submarkets" was effective—indeed, there are only 6% of cross-offers and 1% of cross-matches (the acceptance probability of cross-offers is 0.05, compared with 0.42 of non-cross offers).

[24] Our results cannot be directly compared with the experimental bargaining literature that has documented a preference for fairness. The nature of bargaining in our experiments is substantially different from other experiments, such as the dictator game, where the compromise between the two sides is obvious. In fact, as noted, identifying the set of stable matchings in our markets is arguably not trivial.

8 Market Dynamics

We now turn to the dynamics underlying the final outcomes discussed till now. Several insights emerge from our analysis of the observed dynamics. First, markets progress quickly early on, with many agents finding stable matches in a short amount of time, and slow down as they reach the final matching. Second, agents' behavior exhibits *strategic sophistication*. Specifically, when making proposals, participants target highly preferred partners, but also take receivers' payoffs into consideration: they are more likely to make proposals to blocking partners that gain substantial payoff benefits from the match. Furthermore, participants are successful at avoiding cycles of blocking pairs. While it is challenging to determine exactly how participants manage to avoid cycles, we observe that agents incorporate past interactions into their decisions and reject some offers from blocking partners. Finally, our experiments challenge the dynamic models that have been proposed in the literature. We use simulation exercises to show that *observed dynamics in our data differ from those generated by some of the standard dynamic models in the literature* that, for the most part, are non-strategic and do not incorporate past interactions.

8.1 Evolution towards Stable and Median Outcomes

Markets reach stability gradually, with the vast majority of blocking opportunities, especially the most profitable ones, vanishing during the initial stages of markets, as shown in Figure 2. On the x-axis, we plot the normalized time within a market, where 0 indicates the start of the market and 100 the time at which the final matching is reached (on average, 1.62 minutes across our main treatments). On the y-axis, we plot the total number of blocking pairs (top-left panel), the number of agents in at least one blocking pair (top-right panel), the maximum number of disjoint blocking pairs (bottom-left panel), and the average immediate payoff loss with respect to the best blocking partner of each agent (bottom-right panel). The solid lines

correspond to averages across all markets in our main treatments, and the shaded regions to a standard deviation above and below the average at every point. On average, after 20% of time has elapsed, half of all blocking opportunities vanish, and the average gains per agent from forming blocking pairs have are cut in half. Nonetheless, over half the agents still have at least one potential blocking pair. As markets approach the final matching, the four measures dip towards zero. This dip reflects the fact that the vast majority of markets in our main treatments converge to stability: all four measures equal zero when outcomes are stable.[25]

The pattern of stable match formation echoes the patterns depicted in Figure 2: a significant fraction of agents find their stable match partners early on. Figure 3 depicts the evolution of the shares corresponding to agents matched to a stable partner (in blue), matched to an unstable partner (in orange), and remaining unmatched (in green).[26] Markets in our three main treatments are shown in the top row, panels (a)–(c). The plots further differentiate between *final* and *transient* partners. Two agents are *final partners* (or, equivalently, form a *final match*) if matched to one another at the market's culmination; a pair of matched agents is *transient* if their match is not final.[27] The shares of agents matched to a final (transient) stable (unstable) partner are shown in dark (light) blue (orange). The figure shows the path by which agents find stable and final partners from the early stages of market activity. For instance, when 30% of time has elapsed, over 40% of agents are matched with a stable partner in our main treatments, with the vast majority constituting final partners. This fraction is even higher in markets in which agents have more than one stable partner: over 50% in four-by-four markets, and over 65% in markets with five stable matchings. However, we also observe a

[25] A similar picture emerges when differentiating markets according to their number of stable matchings; as well as for our auxiliary treatments entailing larger markets and unilateral offers. See the Online Appendix for more details (Echenique et al., 2024).

[26] A similar picture emerges when replacing the x-axis with the percentage of offers made so far. See the Online Appendix for more details (Echenique et al., 2024).

[27] Under this definition, a final match *may* break and then form again; all transient matches eventually break.

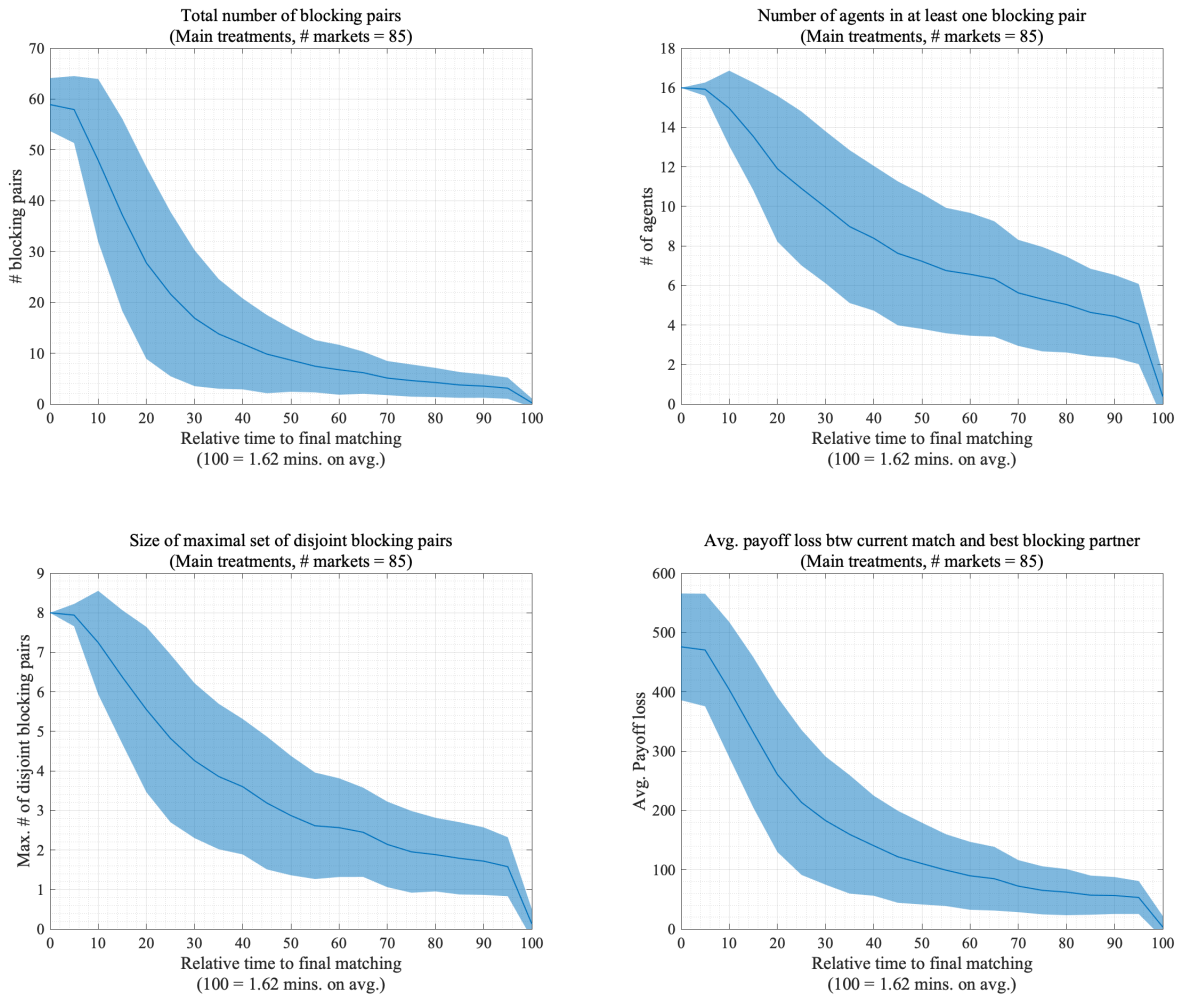
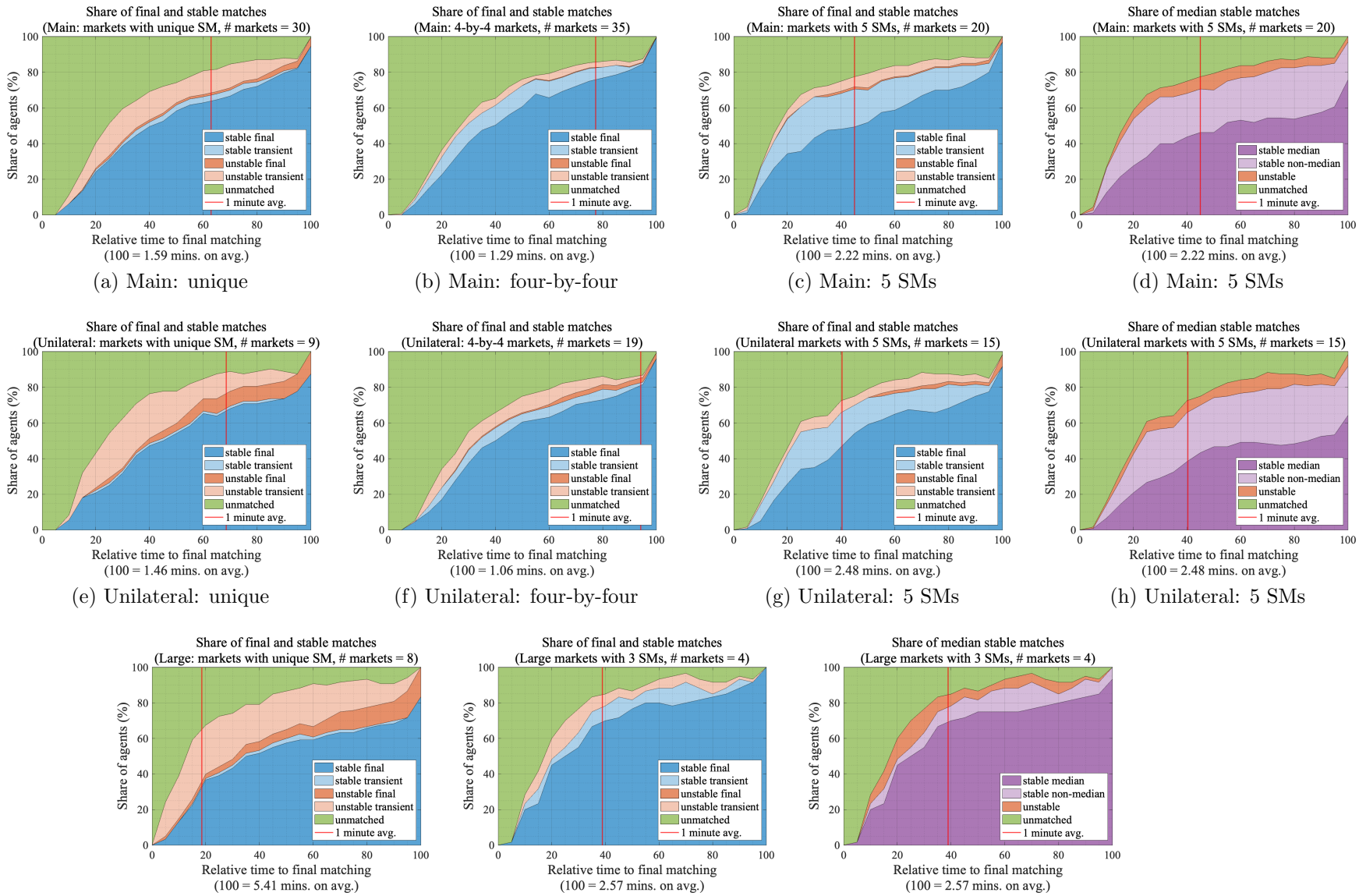


Figure 2: Distance to stable matching over time

larger fraction of transient stable pairs in these markets, indicating that agents match with more than one stable partner throughout the course of the market.[28] The second row of Figure 3 depicts analogous dynamics in our unilateral markets, see panels (e)–(g). Patterns appear remarkably similar to those observed in our main treatments, with agents in unilateral markets also finding stable and final partners at early market stages, and markets converging

[28] Figure 2 and Figure 3 are reminiscent of the dynamics observed by Roth and Xing (1997) in the entry-level market for clinical psychologists. Roth and Xing (1997) document more intense activity at early stages of the market, with a majority of agents finding their final match early on, and a few agents remaining unmatched toward the final stages of the market.



(a) Main: unique (b) Main: four-by-four (c) Main: 5 SMs (d) Main: 5 SMs

(e) Unilateral: unique (f) Unilateral: four-by-four (g) Unilateral: 5 SMs (h) Unilateral: 5 SMs

(i) Large: unique (j) Large: 3 SPs (k) Large: 3 SPs

Figure 3: Progression of final, stable, and median stable matches over time

at similar speeds. The figure also illustrates the similarity in convergence speeds across our main treatments and our unilateral treatment. In particular, whether or not both sides can make offers, convergence speeds are ranked similarly by the number of stable matchings. Last, the bottom row of Figure 3 shows the same shares for our large markets, see panels (i) and (j). While large markets take longer to converge, especially those with a unique stable matching, many agents still find stable and final partners early on.[29]

The formation of median stable matchings is also rapid initially, and then slows down, as seen in the rightmost panels of Figure 3, which plot the shares of agents who are matched to their median stable partner (in dark purple), to a non-median stable partner (in light purple), and to an unstable partner (in orange). For example, when 30% of the time has elapsed in our main treatment with five stable matchings, 40% of agents are matched with their median stable partner, and this figure rises to 77% by the time the final matching has been reached.

8.2 Market Activity across Treatments

The time and volume of activity it takes for markets to reach a final matching is small, as shown in Table 5. In our main treatments, with 8 agents on each market side, the number of distinct (directional) offers is $8 \times 8 = 64$. Regardless of the number of stable matchings, the average number of offers observed till an experimental market terminates is lower than 64. The average number of matches is also limited, and the number of partners each participant has ranges between 2 and 3, depending on market features. Market "complexity" affects the volume of interactions, where markets with more stable matchings exhibit more offers and more repeat matchings. Notably, the activity required under DA offers a poor predictor of behavior in our data. In terms of offer volume alone, under DA, there are no repeat matches

[29] We also analyzed the first proposal made by agents in our experimental markets. The starkest differences are observed in large markets, where agents are less likely to target their first proposal at their highest-ranked partner or at a stable partner. See the Online Appendix for more details (Echenique et al., 2024).

Table 5: Volume of market activity across main and auxiliary treatments

	Unique stable matching			*Four-by-four*		*3 Stable Partners*		
	Main	*Unilateral*	*Large*	*Main*	*Unilateral*	*Main*	*Unilateral*	*Large*
# Mkts.	30	9	8	35	19	20	15	4
# offers	44.8	18.1	198.9	39.7	16.8	59.2	29.9	92.8
# offers per agent (avg.)	2.8	1.1	6.6	2.5	1.0	3.7	1.9	3.1
offers per minute (avg.)	22.3	12.6	33.3	22.5	17.3	21.3	12.0	32.7
time to last offer (mins.)	2.2	1.5	5.8	2.0	1.0	2.9	2.6	3.0
% accepted offers	41.3	71.4	38.6	46.5	76.0	40.4	56.2	39.5
# matches	15.9	12.8	69.2	15.9	12.5	24.6	15.9	37.0
# matches per agent (avg.)	2.0	1.6	4.6	2.0	1.6	3.1	2.0	2.5
matches per minute (avg.)	11.6	9.1	14.3	14.9	12.5	11.5	7.2	15.5
time to final matching (mins.)	1.6	1.5	5.4	1.3	1.1	2.2	2.5	2.6
% matches formed are repeated	14.0	5.6	27.6	19.3	7.7	27.2	11.5	23.3
% matches formed break	21.4	6.3	34.7	10.4	12.0	17.4	11.2	9.3
# offers in food-proposing DA	21.2	16.0	87.8	12.0	12.0	13.0	13.0	36.0
# offers in color-proposing DA	18.3	15.0	86.1	10.0	10.0	14.0	14.0	36.0
# matches in food-proposing DA	11.0	13.0	18.4	10.0	10.0	11.0	11.0	24.0
# matches in color-proposing DA	10.3	11.0	31.0	8.0	8.0	10.0	10.0	23.0

Notes. The table reports averages across all our experimental markets in each of our main and auxiliary treatments of: number of offers made by both sides, average number of offers made per agent, average number of offers made per minute, time (minutes) of the last offer made, percentage of offers that are accepted, number of matches (same as accepted offers), average number of matches (not necessarily distinct) per agent, average number of matches formed per minute, time (minutes) to reach the final matching (last accepted offer), % of matches that are repeated (proposer and receiver had already been matched previously), % of matches that break (a match breaks if any of the two agents is matched to someone else at a later point in time), and number of offers made and matches formed in the food/color-proposing versions of the DA algorithm.

by definition, and the number of offers the algorithm entails is lower than the number observed in our data, although of the same order of magnitude.

We also see some differences across treatments. Within all our treatments, convergence time and offer volume are arguably low. Nonetheless, we observe a lower offer volume—roughly a half, on average—in our unilateral treatments relative to our main treatments. This finding reflects, in part, the higher acceptance rates observed in our unilateral markets: on average, more than half of the offers in our unilateral markets are accepted. In addition, the activity volume we observe in large markets does not align with the number of stable matchings.

On average, large markets with a unique stable matching exhibit close to 200 offers and 70 matches, compared to 93 offers and 37 matches in markets with three stable partners.

8.3 Offers and Acceptances through Time

Proposal and acceptances exhibit different features as markets progress, illustrated in Table 6. The table reports average summary statistics regarding offers made and accepted at distinct stages in our main treatments. We split every market into three segments of equal length relative to the time at which the last offer is made: the early stage (first third), middle stage (second third), and later stage (last third).[30]

The top panel of Table 6 echoes our observations up to now. Offers, as well as their acceptance rates, occur at greater intensities early on in our markets and consistently decline as markets converge. The second panel shows that, while the majority of offers are made to blocking partners, this is more common during early stages, with 80% of offers directed at blocking partners in the first tercile, and only 36% directed at blocking partners in the last tercile. In part, this pattern is mechanical in nature: as markets converge toward stability, fewer blocking pairs are available.[31] Nevertheless, the rate at which blocking partners accept offers increases over time, from 61% in the first tercile to 71% in the last tercile. Importantly, offers from blocking partners are rejected at substantial rates throughout, a point we soon return to. In markets with 8 agents on each side, there are 64 distinct pairs. Given the volume of offers made and accepted, repeat matches are inherent. Indeed, the table illustrates that repeat offers and matches become more frequent as markets progress.

[30] Our auxiliary treatments exhibit similar patterns, see the Online Appendix for details (Echenique et al., 2024).

[31] We also considered how the ratio between the number of matches formed among blocking pairs and the size of the maximal disjoint set of blocking pairs evolves as the market progresses. This analysis parallels the message of Figure 2. Initially, there are many outstanding blocking pairs, generating a large volume of offers. Towards the market's termination, activity tapers down at the same time as the number of blocking pairs diminishes. See the Online Appendix for more details (Echenique et al., 2024).

Table 6: Proposing and accepting behavior across time

	1st/3	*2nd/3*	*3rd/3*	*All*
% *offers*	54.4	28.0	17.6	100.0
% *accepted*	(49.9)	(39.3)	(25.4)	(43.2)
# *offers per minute (avg.)*	37.7	10.1	4.0	22.1
# *matches per minute (avg.)*	17.3	4.6	2.0	13.0
% *matches formed are repeated*	8.7	31.5	54.6	19.3
% *matches formed break*	20.0	15.3	3.2	15.9
% *break* | *final*	10.2	6.6	0.8	6.7
% *break* | *stable*	13.3	9.1	2.0	10.0
Characteristics of offers (% accepted)				
% *to blocking partners*	80.0	50.9	35.7	64.8
	(60.7)	(69.7)	(71.3)	(63.4)
% *are only-proposer beneficial*	17.9	42.3	56.5	31.2
	(7.2)	(5.5)	(2.8)	(5.0)
% *are repeated*	12.5	49.7	69.1	32.7
	(35.1)	(33.0)	(31.4)	(28.8)
% *to previous match*	6.9	26.2	37.0	17.8
	(65.2)	(47.9)	(50.6)	(52.7)
In markets with three stable partners:				
% *to best stable partner*	32.3	34.7	39.8	33.5
	(36.6)	(21.3)	(10.2)	(27.2)
% *to median stable partner*	25.8	28.0	20.3	26.5
	(76.9)	(69.5)	(67.7)	(74.6)
% *to worst stable partner*	6.9	7.6	7.2	7.4
	(87.0)	(94.0)	(60.0)	(82.4)

Notes. The table reports averages across markets in main treatments ("thirds" are relative to the time of the last offer) of: avg. number of offers made and accepted per minute, % of repeated matches, % of matches that break, and % of offers made and accepted (% accepted shown in parentheses): to blocking partners, that are only beneficial to the proposer, repeated, and to a previous match. For markets that have five stable matchings and three stable partners, the table also reports the average % of offers made and accepted (% accepted shown in parentheses) to the proposer's best, median, and worst stable partner.

The bottom panel of Table 6 speaks to dynamics and the selection of stable matchings. In markets where agents have three stable partners, the share of offers agents make to their most-preferred stable partner is higher than the share of offers they make to their median stable partner, which in turn is higher than the share they make to their least-preferred stable partner. Acceptance rates follow the opposite pattern: offers made to best stable partners, who are the worst stable partners for receivers, are less likely to be accepted than those made to median stable partners, which in turn are less likely to be accepted than those made to worst stable partners, who are the best stable partners for receivers. From this perspective, matches with median stable partners can be seen as a form of compromise between the two market sides. This form of the "median as a compromise" is also reflected in convergence times. On average, markets converging to the median stable matching converge faster than those converging to a non-median stable matching, which in turn converge faster than markets that fail to reach a stable matching.[32]

8.4 Comparison with Dynamic Models

We now evaluate several theoretical models of market dynamics that serve as benchmarks for the dynamics we observe.[33] First, we consider the dynamics underlying Gale and Shapley's DA algorithm, in which proposers "go down their preference lists" when making offers and receivers accept beneficial offers myopically. We consider two versions of these dynamics: a version of DA in which proposers are chosen uniformly at random from both sides (2RDA),

[32]The corresponding average convergence times are 2.07, 2.21, and 2.60 minutes, respectively. While illustrative, these differences are not statistically significant. We also do not see any clear relationship between the volume of offers foods make in any particular market and the likelihood the market culminates in the food-preferred stable matching. On average, foods make 50.9% of offers in our main treatment with 5 stable matchings and three stable partners. See the Online Appendix for more differences between markets accounting for final matching reached (Echenique et al., 2024).

[33]The theoretical literature on stabilization dynamics in matching markets is limited. There is an active literature on dynamic matching markets, see Baccara and Yariv (2023), and our literature review. However, relatively little attention has been dedicated to how agents reach stable outcomes, analogous to the literature on learning in game theory, or tâtonnement dynamics in competitive markets.

and a version with Compensation Chains (DACC) by Dworczak (2021), where agents dropped by partners who had made them offers initially get priority in proposing themselves. DACC is guaranteed to converge to a stable matching, while 2RDA is not. Second, we consider the dynamics implied by the sequential formation of blocking pairs. We also consider two versions of these dynamics: the Random Paths to Stability (RPS) model of Roth and Vate (1990), whereby blocking pairs are formed sequentially and uniformly at random, and the Random Best Response (RBR) dynamics proposed by Ackermann et al. (2011), which allows (myopic) optimization in a process resembling RPS. Both RPS and RBR are guaranteed to converge to a stable matching. In their standard versions, the four dynamic models we consider only depend on ordinal preferences. We also analyze alternative versions of these algorithms in which we allow the probabilities with which agents are chosen as proposers, or the probabilities different blocking pairs form, to depend on the cardinality of the match payoffs. We provide a detailed description of each algorithm, and discuss these alternative versions, in the Online Appendix (Echenique et al., 2024).

To contrast the proposing behavior in our main treatments with the "going down the list" DA-type dynamics, we look more closely at *when* and to *whom* agents make offers. We call an agent *active* if there exists an agent they prefer to their current match whom they have not made an offer to previously. Across our main treatments, most proposers (84%) are active. In terms of *whom* offers are made to, we classify offers into three overlapping categories: *downward*, *Gale-Shapley*, and *skip someone*. An offer is *downward* if the proposer has not made an offer to an agent whom they prefer less than the receiver. Intuitively, offers are downward if a proposer is going down their preference ranking when making offers.[34] While the vast majority of the offers we observe are downward (74%), it is not uncommon for agents to go up their ranking. An offer is said to be *Gale-Shapley* if the receiver is the proposer's most

[34]The condition for a downward offer relates to offers made beforehand, not to offers that have been received or accepted: an offer made to an agent the proposer prefers to their current match may still be downward.

preferred agent among those the proposer has not made an offer to (all offers are Gale-Shapley in the DA algorithm). A minority of offers are Gale-Shapley (43%), implying that agents either skip potential partners or make repeat offers throughout the market. Accordingly, an offer is said to *skip someone* if there exists an agent the proposer has not made an offer to, whom they prefer to the receiver. Around 32% of offers skip someone.

In sum, while agents tend to go down their preference lists when making offers, they generally do not follow the order of offers prescribed by the DA algorithm: they frequently make repeat offers and skip agents on their preference list. In this regard, our proposing behavior deviates from the dynamics prescribed by DA.

In relation to blocking pair dynamics, while the majority of offers are made to blocking pairs (65%), this number is far from 100%. And, as already noted, a large fraction of offers made to blocking pairs is not accepted (37%). These two observations go against the dynamics prescribed by both RPS and RBR, in which all offers target blocking partners and are accepted. Moreover, as we soon describe, both RPS and RBR yield different patterns than those we observe in our experimental markets with five stable matchings.

8.5 Market Simulations and Cycles of Blocking Pairs

We now compare our experimental markets with simulations from the dynamic models we have discussed, paying particular attention to the role of cycles in the formation of blocking pairs (Knuth, 1976). By comparing simulated dynamics with the observed behavior in our experiments, we evaluate the extent to which agents avoid cycles of blocking pairs in our experiments. We show that participants appear more intentional in forming blocking pairs than some of the theoretical models allow for, and thus avoid falling prey to cycles predicted by these models.

Table 7: Market dynamics and simulations

	Unique stable matching					*Two embedded 4-by-4 markets*					*5 stable matchings and 3 stable partners*				
	Exp't	*2RDA*	*DACC*	*RPS*	*RBR*	*Exp't*	*2RDA*	*DACC*	*RPS*	*RBR*	*Exp't*	*2RDA*	*DACC*	*RPS*	*RBR*
Market activity															
# offers	44.8	36.2	35.3	36.4	19.8	39.7	25.4	23.8	28.1	15.3	59.2	54.2	48.4	682.0	145.0
# matches	15.9	17.2	17.3	36.4	19.8	15.9	15.9	15.1	28.1	15.3	24.6	30.7	27.9	682.0	145.0
% accepted offers	41.3	51.6	53.5	100.0	100.0	46.5	62.3	63.3	100.0	100.0	40.4	56.3	57.3	100.0	100.0
# offers to blocking pairs	27.2	17.2	17.3	36.4	19.8	22.8	15.9	15.1	28.1	15.3	37.3	30.7	27.9	682.0	145.0
% offers to blocking pairs	65.5	51.6	53.5	100.0	100.0	65.0	62.4	63.4	100.0	100.0	63.5	56.3	57.3	100.0	100.0
# accepted | to BP	15.2	17.2	17.3	36.4	19.8	15.1	15.9	15.1	28.1	15.3	22.6	30.7	27.9	682.0	145.0
% accepted | to BP	60.5	100.0	100.0	100.0	100.0	68.0	99.8	99.8	100.0	100.0	59.5	100.0	100.0	100.0	100.0
# matches formed are repeated	2.7	0.9	1.0	9.1	2.8	3.8	1.8	1.4	5.8	1.9	8.2	4.5	3.8	624.4	108.9
% matches formed are repeated	14.0	4.7	5.1	20.2	10.9	19.3	9.3	7.9	17.1	9.6	27.2	12.9	12.1	79.6	58.7
# repeated matchings	1.2	0.0	0.0	0.4	0.1	1.5	0.0	0.0	0.3	0.1	1.8	0.0	0.0	17.1	2.6
% repeated matchings	6.0	0.0	0.0	0.6	0.3	8.0	0.1	0.0	0.6	0.2	6.2	0.1	0.0	1.3	1.0
# match-level cycles	2.5	1.4	1.6	13.5	3.8	4.2	2.6	1.9	8.3	2.6	10.4	6.9	5.7	8.24e21	2.61e7
avg. match-level cycle length	3.3	3.3	3.3	4.4	3.4	3.2	3.3	3.2	3.9	3.2	3.5	3.9	3.5	96.3	12.2
% offers to stable pairs	35.8	31.1	32.3	35.6	53.1	65.2	66.5	67.9	55.2	78.7	67.5	63.0	63.8	62.4	73.2
% accepted | to SP	73.5	89.8	90.5	100.0	100.0	63.1	73.5	73.9	100.0	100.0	51.4	66.7	67.5	100.0	100.0
% proposer is active	88.1	100.0	100.0	98.9	99.4	78.7	100.0	100.0	98.3	99.3	84.7	100.0	100.0	65.9	91.8
% offer is downward	76.7	100.0	99.9	61.0	81.7	77.8	100.0	100.0	64.7	86.1	61.1	100.0	99.8	21.5	44.6
% offer is Gale-Shapley	42.8	100.0	97.7	19.6	31.4	48.6	99.8	98.4	27.4	39.7	31.8	100.0	94.8	9.0	17.9
% offer skips someone	35.6	0.0	1.8	77.5	66.6	23.4	0.2	1.4	67.8	57.5	41.0	0.0	3.9	39.5	59.3
Stability of outcomes															
% final matching is stable	90.0	89.0	100.0	100.0	100.0	94.3	83.3	100.0	100.0	100.0	75.0	46.9	100.0	100.0	100.0
% final pairs are stable	94.7	98.4	100.0	100.0	100.0	99.6	96.5	100.0	100.0	100.0	96.9	93.8	100.0	100.0	100.0

Notes. The table reports average activity and outcome measures across markets in our three main treatments. Within each of the three panels, the first column reports the experimental data (Exp't), and the second through fifth columns report simulations using Two-Sided Random Deferred Acceptance (2RDA), Dworczak (2021)'s Deferred Acceptance with Compensation Chains (DACC), Roth and Vate (1990)'s Random Paths to Stability (RPS), and Ackermann et al. (2011)'s Random Best Response (RBR). For each algorithm and market, we ran 10,000 simulations. In 2RDA and DACC, proposers are chosen uniformly at random (with the corresponding adjustment in DACC). In RPS, blocking pairs are chosen uniformly at random. In RBR, an agent is chosen uniformly at random among those who have at least one blocking partner.

Table 7 reports averages of several summary statistics across 10,000 simulations of each market in our main treatments, with each of the four algorithms: 2RDA, DACC, RPS, and RBR. With one key exception, our experimental markets take more offers to converge than the four algorithms, and exhibit lower acceptance rates, yielding a roughly similar number of ultimate matches. The one exception is in markets with five stable matchings, when compared with the RPS and RBR algorithms. As shown in the last two columns of Table 7, in these markets, both algorithms take a high number of offers to converge to a stable matching—681 and 145, respectively—which is much higher than what we observe in the experiment (avg = 59, max = 109). As we show below, the reason for this difference is due to both RPS and RBR getting "stuck" in cycles of blocking pairs. Participants in our experiments avoid getting stuck.

As observed by Knuth (1976), when markets evolve through the sequential formation of blocking pairs, markets may cycle between matchings. Table 7 indicates that, on average, our experimental markets with five stable matchings return to a previous matching 1.8 times. This figure is somewhat higher than in our other treatments (1.2 with a unique stable matching, and 1.5 in four-by-four markets). When simulating our markets with five stable matchings with the RPS and RBR dynamics, markets return to a previous matching, on average, 17.1 and 2.6 times, respectively. These figures are echoed when looking at the matched-pair level. In our markets with five stable matchings, on average, 27.2% of matches are repeated. Under RPS and RBR, 79.5% and 58.7% of matches are repeated, respectively. As in the experiment, both RPS and RBR are more likely to cycle across matchings in these markets than in our markets with fewer stable matchings. However, the differences across treatments are starker than in the experiment. On average, RPS and RBR return less than once to a previous matching in our markets with a unique stable matching and our four-by-four markets.

To glean more insight into our experimental participants' engagement in cycles, we now consider repeated matches in more detail. A *match-level cycle* is a maximal sequence of matches, $(f_k, c_k)_{k=1}^{K}$, in an experimental or simulated market, that starts and ends at the same pair, and in which every two consecutive matches are distinct and share exactly one agent. In particular, $(f_1, c_1) = (f_K, c_K)$, and $f_{k+1} = f_k$ or $c_{k+1} = c_k$, for all $k = 1, \ldots, K-1$. The *length* of a match-level cycle is the number of matches that it contains. For example, if we observe that food f matches with color c, then severs this match to form one with another color, say c', and then severs this second match to match with c again, then this sequence of matches constitutes a match-level cycle of length 3: $(f, c) \to (f, c') \to (f, c)$. Intuitively, match-level cycles describe the paths agents follow when they return to previous partners.[35]

Table 7 reports the average number of match-level cycles and their average lengths in our experimental markets, as well as in their simulated counterparts. The average number of match-level cycles in the experiment is below RPS across all our treatments, and below RBR in our markets with a unique stable matching and with five stable matchings.

Zooming in on our markets with five stable matchings highlights the differences between the dynamics implied by random formation of blocking pairs and the dynamics in our data *at the individual level*. The average number of match-level cycles in our experimental markets with five stable matchings is higher than in our other treatments, but of a similar order of magnitude: 10.4 vs. 2.5 in markets with a unique stable matching, and 4.2 in four-by-four markets. By contrast, when simulating our markets with RPS and RBR, the difference between the number of match-level cycles in markets with five stable matchings and markets with fewer stable matchings explodes. Thus, participants in our experiment do not appear

[35] We include a formal definition of a match-level cycle in the Online Appendix, along with examples and a discussion. In particular, the definition allows for there to be more than a single cycle between two identical pairs. For example, in the sequence of matches (f, c), (f, c'), (f', c), and (f, c), there are two match-level cycles of length 3: $(f, c) \to (f, c') \to (f, c)$, and $(f, c) \to (f', c) \to (f, c)$. See the Online Appendix for further details (Echenique et al., 2024).

to be pursuing blocking pairs randomly, and, consequently, generate matches that lead them back to previous partners less than what the RPS and RBR dynamics imply.[36]

The simulation results in Table 7 also show how the individual dynamics we observe in the experiment are more sophisticated than those implied by the "going down the list" DA-type dynamics. While 2RDA and DACC are very effective at avoiding cycles—never returning back to previous matchings, and generating few repeat matches and match-level cycles—in our experiment, the fraction of offers made to blocking partners is higher, and acceptance rates are lower, than in both of these algorithms. This indicates two features of the data. First, participants in our experiment do not appear to avoid cycles by making offers according to DA-type dynamics ("going down their lists"). Second, the behavior of our experimental participants appears more strategically sophisticated than what DA-type dynamics prescribe: participants take into account the incentives of receivers more often than DA, and are not as myopic as DA-type dynamics prescribe when determining whether to accept offers.

Finally, Table 8 reports features of the dynamics leading to different stable partners across simulations of the four algorithms in our markets with five stable matchings. Both 2RDA and DACC predict similar shares of offers made to proposers' best, median, and worst stable partners. However, the acceptance rates in our simulations differ from the ones we observe in the experiment. In the algorithms, offers made to proposers' best stable partners (i.e., coming from the receivers' worst stable partners) are more likely to be accepted than in the experiment. Hence, while in both 2RDA and DACC agents match with their median stable partners with relatively high frequency— 53% and 63% on average, respectively—the

[36] As mentioned, we simulated different versions of these two algorithms where blocking pairs (in the case of RPS) and proposers (in the case of RBR) are chosen with a probability that is increasing in the profitability of the blocking pair formed. These calibrated algorithms bring the number of offers, repeated matches, and match-level cycles to the same order of magnitude of what we observe in the experiment. While this suggests a channel through which cardinal preferences may play a role in avoiding cycles, using the same calibration in our other treatments with fewer stable partners results in very fast convergence and a mismatch with the data. See the Online Appendix for details (Echenique et al., 2024).

Table 8: Simulated dynamics in markets with five stable matchings

	Experiment	*2RDA*	*DACC*	*RPS*	*RBR*
Target of offers					
% *offers to best stable partner*	33.5	30.3	32.1	19.5	22.7
% *accepted* | *to best SP*	27.2	55.6	57.0	100.0	100.0
% *offers to median stable partner*	26.5	23.4	23.7	23.4	27.8
% *accepted* | *to median SP*	74.6	69.1	71.6	100.0	100.0
% *offers to worst stable partner*	7.4	9.4	8.0	19.5	22.7
% *accepted* | *to worst SP*	82.4	95.6	95.2	100.0	100.0
Market-level outcomes (matchings)					
% *median* | *stable*	80.0	64.1	50.4	39.4	46.2
% *non-extremal* | *stable*	100.0	89.1	75.1	77.5	79.1
% *food-optimal* | *stable*	0.0	5.5	19.4	10.8	8.1
% *color-optimal* | *stable*	0.0	5.3	5.4	11.6	12.9
Individual-level outcomes (matches)					
% *median* | *stable*	76.8	52.9	62.8	58.5	62.6
% *fruit-optimal* | *stable*	2.9	22.0	19.4	10.8	8.1
% *color-optimal* | *stable*	20.4	24.7	17.8	30.7	29.3

Notes. The table reports average activity and outcome measures across markets with five stable matchings. The first column reports the experimental data, and the second through fifth columns report simulations using Two-Sided Random Deferred Acceptance (2RDA), Dworczak (2021)'s Deferred Acceptance with Compensation Chains, Roth and Vate (1990)'s Random Paths to Stability (RPS), and Ackermann et al. (2011)'s Random Best Response (RBR). For each algorithm and market, we ran 10,000 simulations. In 2RDA and DACC, proposers are chosen uniformly at random (with the corresponding adjustment in DACC). In RPS, blocking pairs are chosen uniformly at random. In RBR, an agent is chosen uniformly at random among those who have at least one blocking partner.

frequency at which they do so is lower than what we observe experimentally (77%). In RPS and RBR, the frequency with which proposers make offers to all their stable partners is similar, and acceptance rates are 100% by definition. Hence, while these algorithms converge to median outcomes at times, they do so less frequently than our experimental participants.[37]

8.6 Individual Dynamics

We now analyze the main determinants driving market participants to make and accept offers at the individual level.

Determinants of Offer Targets. At an individual level, offers are driven by yield—the chance that the offer will be accepted—and payoff, as well as by past offers and matches, as Table 9 shows. The table reports the results of multiple conditional logits explaining offer targets (who are offers made to). We use the following regressors. For any offer, the proposer's or receiver's *Payoff Advantage (PA)* is the change in payoff to the proposer, or receiver, if the offer is accepted. We include the proposer PA as a dummy, indicating whether it is positive (i.e., the proposer finds the match profitable given the payoff from their current match). We also differentiate receivers' gains from losses by considering receivers' positive and negative PA separately. In addition, we include receivers' rank in proposers' rank-order list (higher rank means less preferred), whether a receiver is currently matched, whether the pair formed by the proposer and the receiver is a blocking pair, whether the pair has been matched previously, the number of offers the proposer has made to the receiver previously, whether the pair formed by the proposer and the receiver is a stable pair, and whether the offer is downward, Gale-Shapley, or skips someone (as defined above). We estimate several specifications, and test their fit in-sample and out-of-sample using the mean-squared error, the percentage of choices

[37] Strictly speaking, in RPS there are no "proposers," since the algorithm indicates only which blocking pair should be formed. When calculating the average, we compute an expectation, treating each relevant blocking agent as proposer with equal probability.

Table 9: Proposal conditional logit estimations

	(1)	(2)	(3)	(4)	(5)	(6)
Proposer's PA > 0	0.159	0.099	0.165	0.085	0.107	0.194
	(0.033)	(0.031)	(0.031)	(0.057)	(0.061)	(0.062)
Receiver's rank (in prop's list)	-0.077	-0.068	-0.060	-0.086	-0.083	-0.058
	(0.001)	(0.002)	(0.004)	(0.002)	(0.004)	(0.007)
Receiver is matched	0.080	0.032	0.029	0.081	0.014	0.014
	(0.031)	(0.019)	(0.021)	(0.022)	(0.024)	(0.029)
Are blocking pair (BP)	0.148	0.159	0.145	0.192	0.168	0.169
	(0.021)	(0.018)	(0.022)	(0.022)	(0.028)	(0.030)
$\max\{Rec\text{'s } PA, 0\}$	0.028	0.014	0.020	0.039	0.021	0.031
	(0.007)	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)
$\min\{Rec\text{'s } PA, 0\}$	0.044	0.028	0.025	0.047	0.025	0.018
	(0.010)	(0.009)	(0.009)	(0.014)	(0.015)	(0.019)
Matched previously		0.036	0.062		0.013	0.053
		(0.011)	(0.012)		(0.027)	(0.017)
# previous offers (prop. to rec.)		0.019	0.027		-0.037	0.007
		(0.004)	(0.007)		(0.028)	(0.010)
Are stable partners (SP)		0.088	0.093		0.061	0.093
		(0.007)	(0.008)		(0.012)	(0.016)
Offer is downward			0.091			0.007
			(0.026)			(0.025)
Offer is Gale-Shapley			0.108			0.116
			(0.037)			(0.024)
Offer skips someone			-0.009			-0.046
			(0.045)			(0.030)
Treatment	Main	Main	Main	Unilateral	Unilateral	Unilateral
Observations	29,760	29,760	29,760	7,246	7,246	7,246
Adj. R^2	0.390	0.423	0.445	0.458	0.470	0.500
MSE (sample)	7.730	7.439	7.187	6.992	6.823	6.391
MSE (2-fold × valid)	7.751	7.472	7.218	7.028	6.906	6.461
MSE (future | present)	7.315	7.155	6.755	5.958	5.813	5.016
MSE (present | future)	8.241	8.003	7.848	8.571	8.397	7.980
%CorrMaxCP (sample)	56.559	56.279	59.010	61.075	63.118	63.548
%CorrMaxCP (2-fold × valid)	56.228	56.023	59.035	61.075	59.140	62.903
%CorrMaxCP (future | present)	60.114	58.457	62.686	73.816	67.409	76.323
%CorrMaxCP (present | future)	52.629	53.413	54.889	49.737	50.438	53.415
Avg $\mathbb{P}(OK\ Pred)$ *(sample)*	40.570	43.409	45.696	44.482	45.611	49.729
Avg $\mathbb{P}(OK\ Pred)$ *(2-fold × valid)*	40.496	43.306	45.612	44.377	45.343	49.453
Avg $\mathbb{P}(OK\ Pred)$ *(future | present)*	40.646	43.607	46.492	47.509	49.172	55.456
Avg $\mathbb{P}(OK\ Pred)$ *(present | future)*	41.194	42.575	44.461	43.180	43.813	45.624

Notes. The table reports average marginal effects of conditional logits. The response variable indicates the receiver of every offer in the data. Standard errors are clustered at the participant level. The table also reports the mean-squared error (*MSE*) of the predicted choice probability, percentage of choices in which the predicted probability of the alternative chosen in the data is the greatest among all alternatives (*%CorrMaxCP*), and the average probability of correctly predicting the data (*Avg* $\mathbb{P}(OK\ Pred)$). Each is computed in the estimation sample and out of the sample using random two-fold cross-validation, predicting the final five rounds with the first five rounds, and the first five rounds using the final five. See the Online Appendix for more details (Echenique et al., 2024).

in which the alternative chosen in the data has the maximum predicted choice probability, and the average probability of correctly predicting the data (see the Online Appendix for details on how we compute these measures; Echenique et al., 2024). The first three columns report average marginal effects across all our main treatments, while the final three columns focus on our unilateral markets.

The estimates in Table 9 confirm some of the observations already made. First, proposers not only take their own preferences into account, but also those of the receivers. The probability of making an offer to a blocking partner is around 0.15 higher than to someone who is not. Likewise, proposers are more likely to target receivers that find it more profitable to match with them (a sort of "second order reasoning" in game theoretic terms). Second, proposers are more likely to make offers to receivers whom they have already proposed to in the past and, especially, to receivers whom they have already been matched with. Third, proposers are more likely to make downward or Gale-Shapley offers. Last, and perhaps surprisingly, proposers are more likely to make offers to receivers who are stable partners, even after including all the regressors discussed above. As already noted, it is not trivial to identify stable partners in our experimental markets through visual inspection. The draw of stable partners indicates that stability may be intrinsically attractive.[38]

Determinants of Offer Responses. Receivers are more likely to accept offers that are more profitable in monetary terms, disregarding proposers' payoffs, and are more selective the more offers they receive over time, as Table 10 illustrates. The table reports the results of binary logits explaining which offers are accepted with similar regressors to those in Table 9.

In our main treatments, larger monetary gains are associated with higher acceptance probabilities, even for proposers who are ranked similarly, echoing the importance of cardinal payoffs

[38] We also conducted similar analysis including cardinal payoff information for both proposers and receivers. Results are qualitatively identical. See the Online Appendix for further details (Echenique et al., 2024).

Table 10: Acceptance binary logit estimations

	(1)	(2)	(3)	(4)	(5)	(6)
Receiver's PA > 0	0.562	0.562	0.495	0.356	0.355	0.417
	(0.044)	(0.043)	(0.034)	(0.113)	(0.113)	(0.121)
Receiver is matched	0.224	0.224	0.200	0.082	0.082	0.048
	(0.037)	(0.037)	(0.032)	(0.062)	(0.062)	(0.066)
$\max\{$*Rec's PA*$, 0\}$	0.034	0.034	0.031	0.029	0.029	0.013
	(0.010)	(0.010)	(0.008)	(0.019)	(0.019)	(0.019)
$\min\{$*Rec's PA*$, 0\}$	0.034	0.034	-0.003	0.354	0.355	0.415
	(0.031)	(0.031)	(0.019)	(0.130)	(0.130)	(0.174)
Proposer's rank (in rec's list)	-0.038	-0.038	-0.015	-0.067	-0.067	-0.052
	(0.005)	(0.005)	(0.004)	(0.008)	(0.008)	(0.009)
Proposer's PA > 0		-0.008	-0.062		-0.035	-0.018
		(0.040)	(0.037)		(0.151)	(0.131)
Matched previously			0.128			0.276
			(0.020)			(0.085)
# previous offers (prop. to rec.)			-0.038			-0.072
			(0.014)			(0.023)
# previous offers (total to rec.)			-0.016			0.004
			(0.005)			(0.008)
Are stable partners (SP)			0.124			0.074
			(0.012)			(0.032)
Treatment	Main	Main	Main	Unilateral	Unilateral	Unilateral
Observations	3,919	3,919	3,919	930	930	930
Adj. R^2	0.376	0.376	0.419	0.365	0.365	0.396
MSE *(sample)*	13.520	13.518	12.515	13.345	13.340	12.609
MSE *(2-fold* $\times$ *valid)*	13.593	13.600	12.675	13.409	13.428	12.876
MSE *(future* $\mid$ *present)*	14.174	14.172	13.286	10.978	10.988	10.652
MSE *(present* $\mid$ *future)*	13.176	13.196	12.234	20.951	20.934	19.948
$\%CorrMaxCP$ *(sample)*	80.684	80.786	82.138	81.398	81.398	82.258
$\%CorrMaxCP$ *(2-fold* $\times$ *valid)*	80.888	80.888	82.164	81.828	82.043	82.688
$\%CorrMaxCP$ *(future* $\mid$ *present)*	79.612	79.669	81.496	86.630	86.630	86.351
$\%CorrMaxCP$ *(present* $\mid$ *future)*	81.411	81.458	82.426	72.504	73.205	73.905
$Avg\,\mathbb{P}(OK\,Pred)$ *(sample)*	72.828	72.830	74.869	73.434	73.439	74.820
$Avg\,\mathbb{P}(OK\,Pred)$ *(2-fold* $\times$ *valid)*	72.814	72.808	74.800	73.362	73.428	74.584
$Avg\,\mathbb{P}(OK\,Pred)$ *(future* $\mid$ *present)*	70.942	70.900	72.855	76.135	76.127	76.265
$Avg\,\mathbb{P}(OK\,Pred)$ *(present* $\mid$ *future)*	74.356	74.313	76.490	68.593	68.451	69.762

Notes. The table reports average marginal effects of binary logits. The response variable is an indicator of whether an offer was accepted. The table also reports the mean-squared error (MSE) of the predicted probability, percentage of choices in which the predicted probability of the alternative chosen in the data is the greatest among all alternatives ($\%CorrMaxCP$), and the average probability of correctly predicting the data ($Avg\,\mathbb{P}(OK\,Pred)$). Each is computed in the estimation sample and out of the sample using random two-fold cross-validation, predicting the final five rounds with the first five rounds, and the first five rounds using the final five.

in our markets. However, when receivers cannot make proposals themselves, as in our unilateral treatments, they place a higher weight on avoiding losses. History matters for receivers: they are more likely to accept offers from proposers with whom they have already matched, but less likely to accept offers from those whom they have already rejected. In our main treatments, receivers are less likely to accept offers as they receive more of them, a pattern that reverse in our unilateral treatments. Similar to proposers, receivers seem to be drawn to proposers who are stable partners, accepting their offers at higher rates.[39]

9 Discussion and Conclusions

This paper presents an experimental investigation of decentralized matching markets. There are three main findings. First, in line with the cooperative solution, decentralized markets often culminate in stable matchings. They do so quickly, in terms of both time and market activity. Second, the median stable matching has very strong drawing power and is frequently selected. Furthermore, cardinal incentives impact the distribution over non-median selected matchings. Roughly speaking, the side of the market that has "more to lose" from forgoing their favorite stable matching, is more likely to implement it. Last, in terms of dynamics, by and large, participants form successive blocking pairs. However, participants are more sophisticated than suggested by some of the naïve dynamics the theoretical literature has focused on, with proposers strategically targeting receivers who value them highly and receivers taking into account past market activity. As a consequence, cycles of blocking pairs are relatively rare in our data.

Our findings are important from a market design perspective. On the one hand, they reinforce the drive to create institutions that implement a stable matching. Indeed, the results

[39] Adding participant fixed effects to these regressions does not alter the results and decreases their predictive power. This suggests that there is no significant unobserved heterogeneity across participants acting as receivers; see the Online Appendix for details (Echenique et al., 2024).

suggest that decentralized processes that precede, or follow, a centralized matching protocol would push outcomes toward stability (and, if not generated by the centralized mechanism in place, may produce inefficiencies having to do with the changes yielding an ultimate stable matching). On the other hand, our findings suggest some important comparisons that underlie the decision to use a centralized clearinghouse to begin with. Many implemented centralized matching markets use a variation of the DA algorithm that, absent incentive compatibility concerns, generate extremal stable matchings, favored by one side of the market. Our decentralized processes did not culminate in an extremal matching often. Therefore, the use of DA should be accompanied with a justification for the choice of stable matching to implement, or at least an appeal to incentive compatibility.

Our results also have implications for the theory of dynamic stabilization in matching markets. Existing models that generate stable matchings through the sequential formation of blocking pairs (such as Roth and Vate, 1990, or Ackermann et al., 2011), or which prescribe an order for the offers made by agents (such as the DA algorithm), do not explain basic features of our data. There is therefore room for further theoretical work that provides foundational guidance on the selection of stable matchings in decentralized markets and takes into account cardinal, not only ordinal, assessments of partners.

Our experiments are designed as a benchmark for decentralized interactions in an idealized setting, allowing a clean examination of the cooperative theoretical predictions. Nonetheless, in applications, a variety of frictions might be at play: incomplete information, offer and time costs, disutility from rejection, and so on. We hope our study opens the door for further investigations of decentralized markets incorporating such frictions.

References

Ackermann, Heiner, Paul W. Goldberg, Vahab S. Mirrokni, Heiko Röglin, and Berthold Vöcking (2011), "Uncoordinated two-sided matching markets." *SIAM Journal on Computing*, 40 (1), 92–106. 4, 37, 39, 43, 49

Agranov, Marina, Ahrash Dianat, Larry Samuelson, and Leeat Yariv (2023), "Paying to match: Decentralized markets with information frictions." *mimeo*. 6

Baccara, Mariagiovanna and Leeat Yariv (2023), "Dynamic matching." *Online and Matching-Based Market Design*. 36

Bergstrom, Carl T., Theodore C. Bergstrom, and Rodney J. Garratt (2013), "Choosing partners: A classroom experiment." *The Journal of Economic Education*, 44 (1), 47–57. 6

Chen, Yan and Tayfun Sönmez (2006), "School choice: an experimental study." *Journal of Economic Theory*, 127 (1), 202–231. 6

Cheng, Christine T. (2008), "The generalized median stable matchings: finding them is not that easy." In *Latin American Symposium on Theoretical Informatics*, 568–579, Springer. 9

Dworczak, Piotr (2021), "Deferred acceptance with compensation chains." *Operations Research*, 69 (2), 456–468. 4, 23, 37, 39, 43

Echenique, Federico, Alejandro Robinson-Cortés, and Leeat Yariv (2024), "Supplement to 'An Experimental Study of Decentralized Matching'." *Quantitative Economics Supplemental Material*. 18, 29, 32, 34, 36, 37, 41, 42, 45, 46, 48

Echenique, Federico, Alistair J. Wilson, and Leeat Yariv (2016), "Clearinghouses for two-sided matching: An experimental study." *Quantitative Economics*, 7 (2), 449–482. 6

Featherstone, Clayton R., Eric Mayefsky, and Colin D. Sullivan (2022), "Why do some clearinghouses yield stable outcomes? experimental evidence on out-of-equilibrium truth-telling." *mimeo*. 6

Featherstone, Clayton R. and Muriel Niederle (2016), "Boston versus deferred acceptance in an interim setting: An experimental investigation." *Games and Economic Behavior*, 100, 353–375. 6

Fehr, Ernst and Simon Gächter (2000), "Fairness and retaliation: The economics of reciprocity." *Journal of Economic Perspectives*, 14 (3), 159–181. 26

Ferdowsian, Andrew, Muriel Niederle, and Leeat Yariv (2023), "Strategic decentralized matching: The effects of information frictions." *mimeo*. 5

Gale, David and Lloyd S. Shapley (1962), "College admissions and the stability of marriage." *The American Mathematical Monthly*, 69 (1), 9–15. 3, 4, 8

Gusfield, Dan and Robert W. Irving (1989), *The stable marriage problem: structure and algorithms*. MIT Press. 2, 9

Haeringer, Guillaume and Myrna Wooders (2011), "Decentralized job matching." *International Journal of Game Theory*, 40 (1), 1–28. 5

Hakimov, Rustamdjan and Dorothea Kübler (2021), "Experiments on centralized school choice and college admissions: a survey." *Experimental Economics*, 24 (2), 434–488. 5

Harrison, Glenn W. and Kevin A. McCabe (1996), "Expectations and fairness in a simple bargaining experiment." *International Journal of Game Theory*, 25 (3), 303–327. 6

Haruvy, Ernan and M. Utku Ünver (2007), "Equilibrium selection and the role of information in repeated matching markets." *Economics Letters*, 94 (2), 284–289. 6

Hitsch, Gunter J., Ali Hortaçsu, and Dan Ariely (2010), "Matching and sorting in online dating." *American Economic Review*, 100 (1), 130–63. 4

Irving, Robert W. and Paul Leather (1986), "The complexity of counting stable marriages." *SIAM Journal on Computing*, 15 (3), 655–667. 9

Kagel, John H. and Alvin E. Roth (1995), *The Handbook of Experimental Economics*, volume 1. Princeton University Press. 6

Kagel, John H. and Alvin E. Roth (2000), "The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment." *The Quarterly Journal of Economics*, 115 (1), 201–235. 6

Kagel, John H. and Alvin E. Roth (2020), *The Handbook of Experimental Economics*, volume 2. Princeton University Press. 26

Knuth, Donald E. (1976), *Marriages Stables*. Les Presses de l'Université de Montreal, Montreal. 2, 38, 40

Nalbantian, Haig R. and Andrew Schotter (1995), "Matching and efficiency in the baseball free-agent system: an experimental examination." *Journal of Labor Economics*, 13 (1), 1–31. 5, 6

Niederle, Muriel and Alvin E. Roth (2009), "Market culture: How rules governing exploding offers affect market performance." *American Economic Journal: Microeconomics*, 1 (2), 199–219. 6

Pais, Joana (2008), "Incentives in decentralized random matching markets." *Games and Economic Behavior*, 64 (2), 632–649. 5

Pais, Joana and Ágnes Pintér (2008), "School choice and information: An experimental study on matching mechanisms." *Games and Economic Behavior*, 64 (1), 303–328. 6

Pais, Joana, Agnes Pintér, and Robert F. Veszteg (2020), "Decentralized matching markets with(out) frictions: a laboratory experiment." *Experimental Economics*, 23, 212–239. 6

Plott, Charles R. and Vernon L. Smith (1978), "An experimental examination of two exchange institutions." *The Review of Economic Studies*, 45 (1), 133–153. 1, 6

Roth, Alvin E. (1984), "Misrepresentation and stability in the marriage problem." *Journal of Economic Theory*, 34 (2), 383–387. 8

Roth, Alvin E. and Elliott Peranson (1999), "The redesign of the matching market for american physicians: Some engineering aspects of economic design." *American Economic Review*, 89 (4), 748–780. 8

Roth, Alvin E. and Marilda Sotomayor (1990), *Two-Sided Matching: A Study in Game-Theoretic Modelling and Analysis*. Econometric Society Monographs. 5

Roth, Alvin E. and John H. Vande Vate (1990), "Random paths to stability in two-sided matching." *Econometrica*, 1475–1480. 2, 4, 37, 39, 43, 49

Roth, Alvin E. and Xiaolin Xing (1997), "Turnaround time and bottlenecks in market clearing: Decentralized matching in the market for clinical psychologists." *Journal of Political Economy*, 105 (2), 284–329. 30

Rudov, Kirill (2022), "Fragile stable matchings." *mimeo*. 2

Smith, Vernon L. (1962), "An experimental study of competitive market behavior." *Journal of Political Economy*, 70 (2), 111–137. 1, 6

Teo, Chung-Piaw, Jay Sethuraman, and Wee-Peng Tan (2001), "Gale-shapley stable marriage problem revisited: Strategic issues and applications." *Management Science*, 47 (9), 1252–1267. 3, 9