An Experimental Study of Decentralized Matching

Federico Echenique* Leeat Yariv†‡

May 30, 2011

Abstract. We present an experimental study of complex decentralized one-to-one matching markets, such as labor or marriage markets. In our experiments, subjects are fully informed of everyone's preferences and can make arbitrary non-binding match offers that are realized only when a certain period of market inactivity has elapsed. We find three main results. First, stable matches are the prevalent outcome. Second, in markets with multiple stable matches, the median stable match is selected most frequently. Third, the cardinal representation of ordinal preferences substantially impacts which stable match gets selected. Furthermore, the endogenous dynamic paths that lead to stability exhibit several persistent features.

Keywords: Decentralized Matching, Experiments, Market Design.

---

*Division of the Humanities and Social Sciences, Caltech, http://www.hss.caltech.edu/~fede

†Division of the Humanities and Social Sciences, Caltech, http://www.hss.caltech.edu/~lyariv

‡We thank Drew Fudenberg, Jacob Goeree, Matthew Jackson, Muriel Niederle, Charles Plott, Andrew Schotter, and Alistair Wilson for useful conversations and suggestions. Gabriel Katz and Erik Madsen provided us with superb research assistance. Financial support from the National Science Foundation (SES 0963583) and the Gordon and Betty Moore Foundation is gratefully acknowledged.

1. Introduction

1.1. Overview. Matching theory is widely used as a normative guide in the field of market design, and as a positive tool offering predictions that can be tested in sociological and economic investigations. The current paper offers an experimental inquiry of the predictive power of matching theory for decentralized matching markets. Concretely, we study whether a decentralized market will reach a stable matching outcome, and if so, which stable matching will be selected.[1] We are also interested in the endogenous market dynamics that takes place in decentralized settings.

Much of the extant literature on matching markets, both theoretical and empirical, has investigated centralized markets: markets where a central clearinghouse dictates who matches with whom. There are many examples in which matching markets are centralized (the medical residency match, school allocations, the U.S. market for reform rabbis, etc.). Nonetheless, many centralized markets come after decentralized market interactions have taken place. Furthermore, many markets are not fully centralized (college admissions in the U.S., the market for law clerks, junior economists, and so on). The analysis of decentralized markets is therefore critical to the design of institutions.

Stability is a central notion in the discussion of two-sided matching markets. In practice, stable outcomes are frequently the goal for implementation. One of the leading reasons is that systems that do not yield stable outcomes have been observed to lead to market collapse through pre- or post-decentralized market interactions.[2] In fact, folk arguments assert that if market participants can interact freely at any stage, they will ultimately install (possibly inefficiently) a stable outcome. This argument has been used not only for defining the objective outcomes of matching institutions, but also for broader econometric exercises. Indeed, the assumption that observed outcomes are stable can serve as the basis for estimating the underlying preferences in the market.[3]

Recent theoretical work has studied decentralized markets and the conditions under which

[1] A matching is stable if no agent prefers to leave the market over their allocated match and no pair of agents prefers to be matched to one another over their given match partners.

[2] The leading example is the unraveling and subsequent collapse of the market for medical interns, see Roth and Sotomayor (1992). A recently documented example is that of gastroenterology fellows, see Niederle and Roth (2003).

[3] See, e.g., Chiappori, Fortin, and Lacroix (2002), Choo and Siow (2006), Heckman, Matzkin, and Nesheim (2006), and Hitsch, Hortacsu, and Ariely (2006).

stable outcomes may make sense as market predictions.[4] The empirical study of decentralized markets has, however, been rather limited. Indeed, decentralized markets are challenging to analyze using field data – preferences and information regarding preferences are unobserved by the econometrician, and the protocol of market interactions is rarely available.[5] Experiments are useful in that they allow us to control for all of these elements.[6]

In addition, stable matchings are typically not unique, and there is no theory of "equilibrium selection" for stable matchings. It is important to understand, at least empirically, if markets systematically favor a stable matching with particular characteristics. All centralized mechanisms make a choice of which stable matching to implement. For example, in the medical residency match, the stable matching that is best for the residents and worst for the hospitals is implemented (if participants report their preferences truthfully). A decentralized market might settle for a different stable outcome.

Our paper is an experimental investigation of two themes. The first theme is stability: will decentralized matching markets yield stable outcomes? The second theme is selection: when there are multiple stable matchings, which one will be selected? What are the factors affecting the selection?

In our baseline experiments, each side of the market was composed of 8 subjects. Subjects were fully informed of all participants' cardinal preferences. Payoffs were designed so that each market participant had either one, two, or three stable match partners. Furthermore, for each ordinal preference profile in the market, we designed several cardinal utility representations, differing in the utilitarian welfare of each side of the market as well as the marginal returns from matching with one partner as opposed to a more preferred one. In each period of our experiment, *agents on each side were 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 arrived (up to one at a time)*. Markets ended after 30 seconds of inactivity. Payoffs were then determined according to the matches created, where unmatched agents received a payoff of 0.

[4] See, e.g., Haeringer and Wooders (2009), Niederle and Yariv (2011), and references therein.

[5] There are a number of studies that document certain aspects of decentralized markets such as their timing, see for example Roth and Xing (1997).

[6] In fact, there are various experimental studies that assess the performance of particular (mostly centralized) matching markets. For example, Chen and Sonmez (2006), Frechette, Roth, and Unver (2007), Nalbantian and Schotter (1995), and Niederle and Roth (2009).

We note that the (cooperative game) theory of one-to-one matching offers predictions on the *set of plausible outcomes,* namely the stable matchings (or the core), under two basic premises on the underlying markets: that all agents are completely informed of other participants' preferences (as well as their own) and that agents can freely match with one another. In that respect, our design aims at mimicking as much as possible these premises. In particular, agents have complete information on everyone's preferences and are free to make offers to one another in a rather unconstrained manner.[7] The design allows us to inspect the organic selection of stable outcomes, when they are reached, and the endogenous path that generates them.

There are three main insights that come out of our experiments. First, *stable outcomes appear in a predominance of cases.* 85% of our markets ended up in stable outcomes, whereas in markets with one or two stable matchings, over 90% of the markets generated stable outcomes. Furthermore, when markets did not yield a stable outcome, the resulting matchings were very close to stable ones in terms of payoffs and number of blocking pairs.[8] Our markets are complex enough that eyeballing a stable matching is arguably impossible (in fact, from a computational perspective, finding the set of stable matchings in generally a difficult problem). It is therefore interesting that experimental markets find a stable matching, and do so relatively quickly (in time, rarely exceeding 5 minutes, as well as in offer volume, which averages 61 distinct offers per market).

Second, in markets in which each agent had three stable match partners, *the median stable matching*[9] *emerged as the modal outcome* – 73% of matched pairs corresponded to median stable match partners. This is particularly interesting when contrasted with the leading clearinghouses used in the field (e.g., the variations of the Gale-Shapley, 1962, algorithm used for the medical resident match in U.S. and the U.K.) that implement a stable matching that is the most preferred by one side of the market, assuming truthful reports of participants'

[7] This is reminiscent in spirit of some of the original general equilibrium experiments, which examined whether markets reach an equilibrium without imposing constraints on the sequential actions that lead them there, see our discussion below.

[8] A blocking pair is one comprised of agents who both benefit from matching with one another over staying with their current match.

[9] Stable matchings are ordered so that there is a stable matching that is the most preferred to each side of the market, and is precisely the least preferred to the other, the third, *median*, stable matching is ranked by all the agents in between the other two.

preferences. Surprisingly, when we ran a treatment in which only one side of the market could make offers, the median matching is still selected most of the time.

The last insight pertains to the selection and the cardinal representation of preferences. Stability is an ordinal concept, and so the set of stable matchings in a market does not depend on the precise match utilities each participant associates with their potential match partners. In the lab, these *cardinal representations have a clear effect on the selection of stable matchings.* In particular, the side of the market that has "more to lose" by forgoing their most preferred matching (say, because their marginal loss from shifting from their most preferred stable matching to a less preferred one is greater than the other side's), tends to get their most preferred matching more frequently.

In order to check for the robustness of these results, we ran two sets of additional treatments. In the first set we tested larger markets, ones containing 30 participants (with 15, instead of 8, subjects on each side). The three insights still hold – stable matches, and in particular median stable matches, are frequent. In the second set, we allowed for only one side of the market to make offers. Indeed, in many decentralized markets (such as the market for junior academic economists), only one side of the market effectively makes offers. We observe similar results in these markets. Stable matchings still occur habitually at a rate of 87%. Furthermore, despite the absence of offers from one side of the market, median stable matchings remain the modal empirical outcome, occurring in 59% of the markets in which agents had three stable match partners.

Our design also allows us to track the dynamic path by which our subjects reach stability. As mentioned, subjects converge to a stable matching rather rapidly. We use several discrete choice models to explain the making of and responses to offers. We find that subjects are strategically sophisticated: when making an offer they seem to put themselves in the place of the recipient of the offer, and gauge the likelihood that the offer is accepted. The interplay between match utilities and the likelihood of offer acceptance seem to explain a substantial portion of market activity.[10]

[10] We also use simulations to examine whether dynamic models that have been offered in the theoretical literature fit our data, namely the Roth and Vande Vate (1990) dynamics and the Gale-Shapley (1962) algorithm with proposers that are randomly chosen. Neither seem to match properly the features of the empirical dynamics.

In what follows, we present a review of the related literature. We provide some basic theoretical background in Section 2. We then describe our experimental design in Section 3. A description of the aggregate outcomes appears in Section 4, with an analysis of the impacts of the cardinal representations of preferences appearing in Section 5. The experimental dynamics and the corresponding individual behavior are studied in Section 6. Section 7 concludes.

1.2. Literature Review. The paper ties to two separate strands of literature: theoretical and experimental. Theoretically, our experimental design corresponds to the cooperative game theory model of matching markets (see Roth and Sotomayor, 1990). This setting has crisp predictions: outcomes coincide with the core of the market, the set of stable matchings. In that respect, our experimental results provide a strong confirmation of the theory underlying the notion of stability.

Several recent papers propose particular dynamic decentralized processes by which one-to-one matchings are created (see, e.g., Haeringer and Wooders, 2009, Niederle and Yariv, 2011, and Pais, 2008). These papers usually impose some structure on the process by which offers are made and accepted (namely, offers are made by only one side of the market). The main focus of this literature is the identification of conditions under which stable outcomes are likely to arise as equilibrium outcomes. Roughly speaking, complete information of the prevailing preferences (as is the case in our experiments) allows for stability to emerge in equilibrium, while a certain amount of correlation between agents' preferences is required for stability to be the unique prediction. To the extent of our knowledge, the extant literature is silent on the selection of stable matchings when multiple ones exist in the market.

Experimentally, there are fairly few laboratory studies of decentralized matching. Kagel and Roth (2000) analyze the transition from decentralized matching to centralized clearing-houses, when market features lead to inefficient matching through unraveling. Nalbantian and Schotter (1995) analyze several procedures for matching with transferable utility, decentralized matching among them, where agents have private information about payoffs. Nalbantian and Schotter 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. 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.[11]

There is also a conceptual 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.

2. Theoretical preliminaries

We start by reviewing the underlying cooperative 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. The sets F and C can be metaphors for men and women, firms and workers, etc. that are to be matched to one another in the market. A *matching* is a function $\mu : F \cup C \to F \cup C$ such that for all $f \in F$ and $c \in C$,

1. $\mu(c) \in F \cup \{c\}$,
2. $\mu(f) \in C \cup \{f\}$,
3. and $f = \mu(c)$ if and only if $c = \mu(f)$.

Denote the set of all matchings by $\mathcal{M}$. The notation $\mu(a) = a$ means that a is unmatched under μ; $c = \mu(f)$ denotes that f and c are matched under μ.

A *preference relation* is a linear, transitive, and antisymmetric binary relation. A preference relation for a food $f \in F$, denoted $P(f)$, is understood to be over the set $C \cup \{f\}$. Similarly, for $c \in C$, $P(c)$ denotes a preference relation over $F \cup \{c\}$. For simplicity and consistency with our experimental design, we assume that each food (color) prefers any color

[11] There is a growing experimental literature studying centralized matching systems, e.g., Bergstrom and Garratt (2010), Harrison and McCabe (1996), Chen and Sonmez (2006), Haruvy and Ünver (2007), Pais and Pintér (2008), Echenique, Wilson, and Yariv (2010), Featherstone and Mayefsky (2010), and Featherstone and Niederle (2011).

(food) over remaining unmatched. A *preference profile* is a list P of preference relations for foods and colors, i.e.,

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

We assume that preferences are *strict*. That is, no food or color is indifferent over two different partners.

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. A matching is *stable* if there is no pair that blocks it. (We ignore individual rationality since we restrict attention to preferences under which all agents prefer to be matched to another agent rather than remaining unmatched.) Denote by $S(P)$ the set of all stable matchings.

Gale-Shapley 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, c, and $\mu \in S(P)$,*

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

Note that 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*. 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 what is often referred to as the Gale-Shapley deferred acceptance algorithm to identify one of the extreme matchings μ_F or μ_C. Beyond its theoretical role in establishing existence, the Gale-Shapley algorithm is often the algorithm used in centralized markets. For instance, the National Resident Matching Program uses a variation of the Gale-Shapley algorithm.

Suppose P is a preference profile for which $S(P)$ has an odd number K of matchings. A *median stable matching* is a matching $\mu \in S(P)$ such that for all agents $a \in F \cup C$, $\mu(a)$ occupies the $\frac{K+1}{2}$'th place in a's preference among the agents in $\{\mu'(a) \mid \mu' \in S(P)\}$: $\mu(a)$ is

a's median partner among a's stable-matching partners.

In general, median stable matchings are guaranteed to exist (see Teo and Sethuraman, 1998; this is also the case when K is even). In a way, median stable matchings present a compromise between the two sides of the market. Interestingly, there are no known simple algorithms that generate the median stable matching. Certainly, one can search for all stable matchings of a market and then identify the 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 illustrate that determining the number of all stable matchings is generally #P-complete; 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 the Gale-Shapley algorithm.

Note that 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.

3. Experimental Design

Our baseline design was a decentralized market involving two groups of 8 subjects each (comprising the two sides of the market), which we neutrally termed colors and foods. In each round, each subject was randomized a role ("red," "blue," etc. if a color; "apple," "banana," etc. if a food). A subject could match with one and only one subject from the opposite group. They derived different monetary payoffs from matching with different subjects. All subjects observed all potential payoffs from a numerical matrix on the experimental interface.[12] Remaining unmatched generated a payoff of 0.

Over the course of the experiment, agents were free to propose a match to any agent on the other side of the market. At any point in time, subjects observed all current matches through a panel of the interface. Importantly, subjects could make an offer while being matched, and the target of offers could be currently matched. If a matched agent accepted a new offer, the previous match was undone. When receiving an offer, a subject had 10 seconds to respond.

[12] Full instructions are available at: https://sites.google.com/site/decentralizedmatching/

Each market ended after 30 seconds of inactivity.[13]

Each experimental session entailed 2 practice rounds and 10 real rounds. We designed the different markets with two factors in mind. First, in order to see whether cardinal representations of preferences matter we chose 7 preference profiles, so-called ordinal markets, and 20 cardinal representations of them (which we spell out below).[14] Second, in order to check for the endogenous selection of stable matches, we designed our experimental markets so that in each of them all subjects had either one, two, or three stable match partners (in particular, the first case corresponds to markets with a unique stable matching).

As mentioned, for each fixed number of stable match partners (ranging from 1 to 3), we used several markets differing in ordinal and cardinal preferences of market participants.[15]. The following is a general description of the payoffs used.

Unique stable match partner**.** We used 5 different ordinal markets: aligned preference markets (each side's participants having aligned preference rankings), markets with aligned preferences on one side,[16] a market in which a fully egalitarian matching (providing all agents the same payoff) which was not stable existed, and two markets with the same unique stable matching, in which the correlation of payoffs across the market was different (-0.9 and 0.9). For the market in which this correlation was lower, 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¢. These generated 7 different cardinal markets.

Two stable match partners. 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 (so that 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 match, the utilitarian benefit of foods relative to colors from each match, the distribution within each

[13]In less than 5% of the markets, activity stalled and markets were terminated after 5 minutes. We discuss the duration of markets when we describe our results.

[14]Recall that, in theory, the set of stable matchings depends only on the preference rankings (or ordinal preferences) of market participants.

[15]The full set of matrices used are also available at: https://sites.google.com/site/decentralizedmatching/

[16]Alignment implies that agents on each side of the market agree on the ranking of agents on the other market side. When preferences are aligned on one side, only the members of one side of the market are in agreement. Alignment on either side of the market is sufficient to guarantee a unique stable matching.

match[17], 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.

Three stable match partners. We used one ordinal market represented cardinally in 5 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¢, one in which the reverse occurred (for foods marginal differences of 70¢ and for colors 20¢), and one in which these differences were 20¢ for both sides of the market but colors' payoffs were all shifted up by \$1.[18] In these markets, while each individual has precisely three different stable match partners, there are five different market-wide stable matchings.

In addition, we ran two sets of additional treatments to test for the effects of market size and bargaining power:

Large markets. We ran several treatments with larger markets, containing 15 subjects on each side. We concentrated on markets with either a unique matching or 3 stable matchings. We used 3 cardinal markets with a unique stable matching: aligned preferences on one side, aligned preferences with a fully egalitarian unstable matching, and unaligned preferences. We used two markets with 3 stable matchings (individually and market-wide).

One-sided offers. We also ran several sessions with 8×8 markets (payoffs as in our baseline treatment) in which only foods could make offers (the rest of the market proceeding as before).

All sessions took place at the California Social Science Experimental Laboratory (CASSEL), using a modification of the multi-stage software. Subjects were all UCLA undergraduates and each subject participated in only one session. The average payment per subject was \$39 in our baseline treatments, \$61 in large 15×15 market treatments, and \$46 in the

[17] Since egalitarian motives appear frequently in experiments, we were concerned that some form of altruism would be driving our results. We therefore designed our payoffs so that in some treatments, fully egalitarian treatments were not stable. Furthermore, we included cardinal representations in which certain *stable* matches were more egalitarian than others.

[18] The underlying market used is not symmetric, which is why reversing the marginal differences across sides of the markets does not lead to an equivalent market.

Market Size	Side Making Offers	Number of Stable Match Partners	Type of Markets*	Number of Sessions	Number of Markets Run	Number of Subjects
16	Both	1	O - 5, C - 7	5	26	144 = 4*32 + 16
16	Both	2	O - 1, C - 6	5	39	144 = 4*32 + 16
16	Both	3	O - 1, C - 5	5	25	144 = 4*32 + 16
30	Both	1	O - 2, C - 3	2	8	60 = 2*30
30	Both	3	O - 2, C - 2	2	4	60 = 2*30
16	Foods	1	O - 1, C - 2	4	11	112 = 3*32 + 16
16	Foods	2	O - 1, C - 4	4	42	112 = 3*32 + 16
16	Foods	3	O - 1, C - 4	4	30	112 = 3*32 + 16

* O stands for number of ordinal markets, C stands for number of cardinal presentations of these markets

Table 1: Summary of Treatments

treatments in which only foods were able to make offers.[19] All of these were combined with a $5 show-up fee in our baseline treatments. Table 1 summarizes our experimental treatments.

4. Market Outcomes

We start by describing aggregate market outcomes. There are three main findings that emerge from our analysis. First, *most market outcomes are stable.* Furthermore, market outcomes that are unstable are close to stable; they are close both in terms of the number of blocking pairs and in terms of unrealized payoff gains. Second, in our treatments with three stable partners, *most agents are matched to their median stable partner.* Surprisingly, 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.

4.1. Stability in Experimental Markets. Table 2 summarizes the overall outcomes in our baseline treatments. Almost all agents were matched across markets: overall 98% of the agents (852 of 864) were paired. Furthermore, 95% of matched pairs were stable, with little response to the type of market (in terms of the number of stable matches).[20] From the

[19] Standard deviations were $5, $3, and $4 respectively.

[20] We note that the use of color and food labels in our markets did not seem to have any effects on matchings. 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.

Treatment	Matched	Stable	Market Stable	Instability Payoff Loss PP*
Unique	239/240	95%	90%	16.9 ± 12.2 ¢
Two	351/352	94%	89%	15.0 ± 6.9 ¢
Three	262/272	95%	47%	32.3 ± 10.0 ¢
Total	851/864	95%	76%	27.2 ± 7.2 ¢

* Payoff differences between realized match and utilitarian welfare-maximizing stable match (per person, PP), for markets achieving unstable matches.

Table 2: Aggregate Match Outcomes in the Baseline Treatment

viewpoint of market-wide outcomes, 76% of markets were fully stable, in the sense that all pairs correspond to a stable matching (and there are no blocking pairs). Markets with three stable matchings exhibited less market-wide stable outcomes (namely, 47%).

Even markets that were not fully stable were, in fact, close to stable. This can be seen using different measures of proximity. In terms of payoffs, we can compare the realized payoffs with those achievable through the markets' stable matchings. We find that agents within markets that did not culminate in a stable matching lost between 6% and 7% of the realized total payoffs (translating to 22¢ and 27¢ per person, respectively), the range corresponding to the different stable matchings that could have potentially been selected in markets with multiple stable matchings (the closest one in terms of payoffs, or the most efficient one that is the point of comparison in Table 2). If we average these payoff differences across all markets, stable or unstable, the loss per person is bounded by 2% of realized earnings (between 5¢ and 6¢ per person).[21] Nonetheless, restricting attention to individuals not matched to a stable match partner, differences become considerable, ranging from 24% and 64% per person (between 65¢ and 171¢).

A second measure of market stability pertains to the number of blocking pairs.[22] Figure 1 presents empirical cumulative distribution functions (CDFs) corresponding to the number of blocking pairs in our different treatments. The figure contains the distribution of blocking pairs pertaining to *markets in which outcomes were not fully stable*; including all markets would imply a big spike at 0, which corresponds to a stable matching. The message echoes the

[21] An alternative gauge for the distance from stability is the difference between payoffs in unstable market outcomes and the payoffs from realizing existing blocking pairs (a comparison of the payoffs of the realized matching with those generated by a matching that is not necessarily stable). The conclusion is the same; existing blocking pairs do not result in substantial payoff gains.

[22] In our baseline treatments, when no agents are matched, the number of blocking pairs is 64, the number of potential pairs. When a stable matching is implemented, there are 0 blocking pairs.

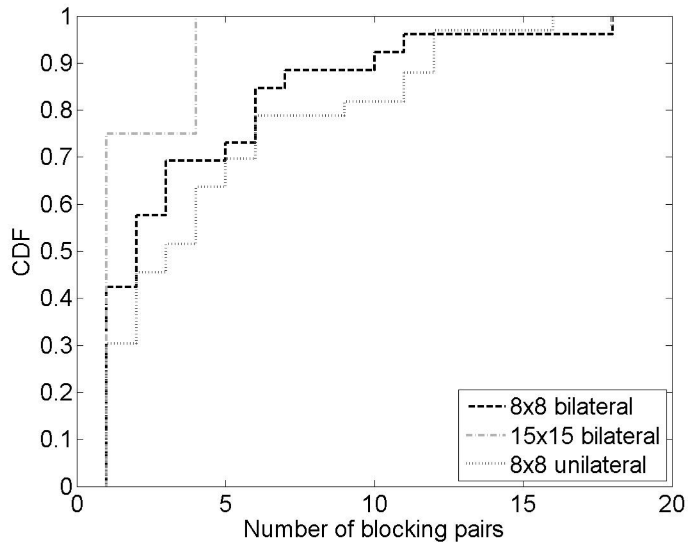

Figure 1: Cumulative Distributions of Blocking Pairs for Unstable Markets

conclusion we drew from payoffs: In our baseline treatment it appears that most outcomes are close to stable. Indeed, most of the markets with unstable outcomes have only a few blocking pairs.

Regarding time, convergence to the stable matching was rather rapid. On average, it took subjects 169 seconds (± 7) in our baseline treatment to achieve a matching. We return to the dynamics underlying these observations in Section 7.

4.2. The Emergence of Median Stable Matches. Table 3 contains the distribution of individual matches in our baseline markets entailing three stable match partners for each agent. Recall that the crucial wedge between the different 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 colors and y cents for foods. In the $20 - 20$ marginals market with 100 color shift, a 100¢ (or \$1) were added to the color payoffs in the $20 - 20$ marginals market.

The first observation to glean out of the table is that *the median stable matching is the modal outcome*: 73% of individual matches correspond to the median stable match. This observation holds for all of the experimental markets with three stable match partners per

Matching	70-70	70-20	20-70	20-20	20-20 w/100 shift	Total
Food optimal	8	0	15	4	0	27
Color optimal	11	8	0	9	13	41
Median	80	32	20	25	24	181

Table 3: Cardinal Effects

person. The table focuses on the individual outcomes since those are the ones concerning the participating agents in our markets (agents are concerned with partners, not with market-wide outcomes).

Shifting attention to market-wide outcomes, we establish similar insights. In none of our baseline treatments were the extreme stable matchings (food-optimal or color-optimal) selected. Furthermore, in 44% of the markets that ended up in a fully stable matching, the market-wide median matching, entailing *all* participants being matched to their median stable match partner, was achieved. Recall that most markets with three stable match partners per person have five market-wide stable matchings. In these markets, some non-extremal matchings entail some subjects being matched to their favorite stable match partners, and only a sub-group being matched to their median stable match partner. Therefore, the design was such that made the selection of the median stable matching, in which *all* individual matches are to the median stable match partner, arguably more difficult to achieve. We should emphasize that a median matching had to be "discovered" by the subjects: it is practically impossible to read the set of stable matchings from looking at the payoff tables (an 8×8 table with pairs of numbers as entries).

These results are relevant when considering centralized systems. Indeed, absent incentive compatibility concerns, most centralized clearinghouses are based on the Gale-Shapley algorithm and implement one of the extremal stable matchings (in our case, the matching optimal for food or for color). Our results suggest that decentralized markets may be more likely to generate an intermediate stable matching. From a welfare perspective, while all stable matchings are Pareto efficient, we can compare the median and the optimal stable matchings according to the sum of agents' payoffs. In our markets, the median stable matchings always generate a utilitarian welfare level that is (weakly) in between that generated by the extremal matchings. However, which extremal matching generates the maximal utilitarian welfare (food-optimal or color-optimal) depends on the market. Therefore, the utilitarian welfare merit of a centralized system over a decentralized system may depend on which extremal matching is being implemented.[23]

We note that decentralized matching markets appear to comprise another incidence in which a computationally hard problem (see Cheng, 2008 and Gusfield and Irving, 1989) is solved quickly in the laboratory. Granted, the set of markets presented in our experiments is a rather selected one, but it is intuitively hard for a single person to choose a median outcome out of the set of stable matchings in our experiments (for a similar observation regarding the NP-complete four-coloring problem, see Kearns, Suri, and Montfort, 2006).

Table 3 also suggests the impact of cardinal representations of preferences. Indeed, participants on the side of the market (foods or colors) having "more to lose" by forgoing their most preferred match 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 6.1.

5. Robustness to Bargaining Protocol and Market Size

There are two natural dimensions that are important to consider when contemplating the robustness of our results: market size and the bargaining power afforded to each side of the market.

Market Size. Our treatments involving markets with 30 participants (15 on each side) suggest that our observations from the baseline treatments are not special to the volume of 16 individuals engaged in the market. Indeed, our larger markets generated very similar results to those yielded by our baseline treatments. In terms of markets created in the lab, a size of 30 is quite large. It does not come near the much larger actual matching markets, but it is a comfort that the duplication of the size of our baseline markets does not upset the results.

Regarding matching outcomes, in the large markets, 89% of individual matches were stable. In terms of market-wide outcomes, 67% of matchings were stable, and the number of blocking pairs within unstable markets was very small, all entailing less than 5 blocking pairs (see Figure 1). Interestingly, the seemingly more complex markets with three stable matchings always

[23] We note that in some markets the median stable matching generates a higher utilitarian welfare than the average generated by the extremal matchings.

culminated in a stable matching in the 15×15 markets. As one might expect, convergence was slower in these large markets: The average duration of a round was 321 seconds (± 53).

As for the selected stable match partners, results are similar to those in our baseline treatments, though arguably more extreme. In the large market treatments, 93% of individual matches were median matches (the rest evenly split between food-optimal and color-optimal). The market-wide stable matchings were exclusively median stable matchings.[24]

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. Intuitively, there is a wedge between our experimental design and the setting of the Gale-Shapley algorithm in which only one side makes offers while the other side decides which offers to accept. The Gale-Shapley algorithm 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. 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 test this, we ran markets in which only one side could make offers: in our unilateral treatment, only foods could propose matches. We find very similar results to our baseline treatment, namely that most outcomes are stable and correspond to median outcomes.

In our unilateral treatments, 98% of subjects were matched. 87% of these individual matches were stable, without great variation across the types of markets.[25] At the market level, 43% of the markets culminated in a stable matching. However, much as in our baseline treatments, resulting unstable matchings were very close to stable ones. In fact, 73% of the unstable matchings corresponded to a stable matching in which one pair of participants was unmatched. The convergence in these markets was somewhat faster than in our baseline treatments: rounds lasted on average 145 seconds (± 10).

Regarding selection, two important observations emerge. First, in markets with only

[24] In a few cases, some agents happen to have the same partner in the median as in an extremal matching. Such matches were not counted as medians in the individual-level statistics.

[25] The percentages are 83%, 96%,and 82% for markets with one, two, and three stable match partners per person, respectively.

two stable matchings, the stable matchings favored by foods (the proposers) are much more frequent. In fact, 91% of the matches are food-optimal, which is significantly greater than the observed 61% in our baseline treatments (at any conventional confidence levels). For these markets, results are consistent with the Gale-Shapley intuition. Nonetheless, the results for markets with a median stable matching (those corresponding to markets with three stable match partners for each individual) are very similar to those in our baseline treatments. 61% of the individual stable matches are median stable matches, compared to 73% when both sides make offers. There is an increase in the number of matches that are optimal to the proposing side, from 11% to 27%, but this comes mainly at the expense of matches that are optimal for the receiving side.

6. The Effects of Cardinal Incentives

The notion of stability depends on agents' ordinal preferences, not on specific cardinal utility values. In fact, in our data the likelihood a stable matching emerges does not seem to depend on the cardinal representation of preferences. In principle, however, *which* stable matching gets selected may depend on the cardinal utility representations. Existing theories provide little guidance as to which stable matchings are to be expected in different (decentralized) contexts. In that respect, our experiments are the first to suggest some market aspects that affect the selection of stable matchings.

Our design entailed several markets composed of identical ordinal preference profiles, but different cardinal utility representations of these preferences. Focusing on those markets, there are two consistent patterns we observe. First, higher cardinal incentives for one side of the market make that side's optimal matching more likely to emerge. Second, equality of payoffs across match partners makes a matching more likely, as long as it is stable.

6.1. Selection of Extremal Matchings. Table 3 reports the distribution of stable matches across the different cardinal representations of preferences in markets in which each participant had three stable match partners.

The table makes transparent the importance of cardinal incentives, keeping ordinal incentives fixed. The side that benefits more from getting a more preferred match, in terms of marginal increases or absolute payoffs, tends to generate its preferred stable matches at significantly greater frequencies (the asymmetry between, say, the $70-20$ marginals market and the $20-70$ marginals market is due to the fact that our underlying ordinal market is not symmetric across the two market sides). It is only when marginal incentives are balanced that we see some extreme stable matches of both types (food-optimal and color-optimal).

We note that while cardinal incentives affect the distribution of extremal stable matches, they do not appear to affect the generated fraction of median stable matches.

6.2. Social Preferences. One may suspect that inherent attributes of our markets' cardinal representations make the median stable matches more appealing to subjects. In particular, there is a long legacy in experimental economics identifying subjects' tendency to satisfy some taste for egalitarian or "fair" outcomes (the literature has suggested different types of social preferences of the sort, for surveys see, e.g., Chapter 4 in Kagel and Roth, 1995 and Fehr and Gächter, 2000). Such considerations may be important since different cardinal representations of preferences may make one stable matching more or less appealing when having equality of payoffs in mind.

We analyzed these effects in two ways. We first explored egalitarian concerns by introducing some markets with a salient egalitarian but unstable matching (see the description in Section 3). These are markets in which one unstable matching entailed identical payoffs to all market participants, and comparable utilitarian welfare to that generated by the stable matchings.[26] Subjects consistently failed to select the egalitarian matching. In our baseline treatments, none of these markets ended up in an unstable (and potentially more equal) matching outcome. Our larger markets (in which more payoff variation could be introduced), generated 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, but which was much more unequal.[27] While the final outcome in these markets was not always fully stable (in some instances a few blocking pairs remained), subjects clearly avoided the egalitarian unstable matching.

[26]By specifying identical payoffs to all market participants, we avoided the challenge of identifying the individuals that are relevant for a subject's social preferences.

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

We used the markets with two stable matchings to assess the effect of cardinal utilities on the distribution of matchings between those optimal for foods and those optimal for colors (see Section 3).[28] We find that when the dispersion of payoffs is relatively high in the color-optimal matching, markets tend to achieve the food-optimal matching; when the dispersion is relatively low, markets tend to settle on the color-optimal matching. Specifically, we computed the *coefficient of variation* of agents' payoffs in a given matching: this is the standard deviation of payoffs divided by the mean payoff. The coefficient is a "scale free" measure of the dispersion in payoffs across agents. We computed the ratio of the coefficient of variation at the food optimal matching and the coefficient at the color-optimal matching. When the ratio is above the median ratio in our data, the color-optimal matching obtains 36.7% of the time (±8.8%); when it is below the median it obtains 75.0% of the time (±6.3%). The difference between these values is significant at any conventional levels 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 while some egalitarian or fairness considerations play a role in selecting outcomes, they are not so strong as to trump stability.

It should be noted that while distributional concerns are a prevalent concern in experimental economics, our results cannot be directly compared with the extant experimental bargaining literature. Indeed, 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, it is far from obvious how to find the median matchings in the markets we tested.[29]

7. Market dynamics

We now turn to the dynamics that led to the outcomes we have been describing. Overall, Figure 2 presents the evolution of the number of blocking pairs over the duration of a round,

[28] Recall that each of these markets was constructed to embed two smaller 4×4 embedded markets, that allowed us to gain more payoff variations. Two notes are in order: 1. The embedding of the "submarkets" was effective - indeed, there were only 7% cross-offers and only 1% cross-matches; 2. In these markets, 96% of matches were stable, providing additional supporting evidence to the robustness of our results with respect to market size.

[29] Our markets are complex enough that eyeballing a stable matching is arguably impossible (and as mentioned above, from a computational perspective, finding the set of stable matchings in generally a difficult problem). Furthermore, there are no simple algorithms for finding the median stable matching.

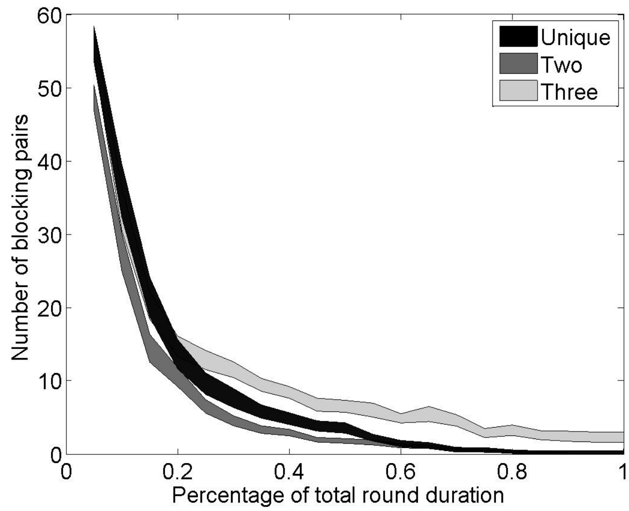

Figure 2: Volume of Blocking Pairs over Time within Rounds

where the thickness of the curves corresponds to one standard error below and above the mean at each point. The figure suggests that blocking pairs vanish over time, and they do so at a relatively fast pace. When 20% of a round had elapsed, markets with a unique stable matching had on average roughly 18 blocking pairs, similar to those in which each agent had three stable match partners. When 95% of the duration of the round had elapsed, only the markets with three stable matches per person still exhibit some blocking pairs (consistent with the percentage of such markets converging to a matching that is close to, but not precisely, stable). This is in line with the rapid convergence documented in Section 4. Regardless of the number of stable match partners available for each participant, most markets end with very few blocking pairs; and after a relatively short interval of time, most blocking pairs are traded away.[30]

Another important feature is the relatively small volume of offers apparent in our experimental sessions. The first column of Table 4 presents the volume of offers by the number of stable match partners each market participant has. The number of offers seems intuitively small – the number of distinct offers (from both sides) in each market is 128, while the Gale-Shapley algorithm would entail between $12 - 16$ steps in most of the markets.[31] The volume

[30] A similar image emerges for our additional treatments entailing unilateral offers or larger markets.

[31] In two markets with a unique stable matching the number of steps the Gale-Shapley algorithm requires

Number of Stable Partners	Number of Offers	% of Repeat Offers	% of Repeat Matches
One	45 ± 6	29%	14%
Two	39 ± 3	32%	20%
Three	61 ± 4	35%	28%
Total	48 ± 3	32%	21%

Table 4: Market Dynamics – Aggregate Statistics

of offers depends to some extent on the cardinal incentives. Focusing on the markets discussed above with three stable match partners per person, in which the overall average number of offers per market is 61, the markets described as 70-70 and 20-20 have on average 65 and 69 offers. The other, unbalanced, markets have between 45 and 49 offers on average.[32]

Another important feature is the relatively small volume of offers apparent in our experimental sessions. The first column of Table 4 presents the volume of offers by the number of stable match partners each market participant has. The number of offers seems intuitively small – the number of distinct offers (from both sides) in each market is 128, while the Gale-Shapley algorithm would entail between $12-16$ steps in most of the markets.[33] The volume of offers depends to some extent on the cardinal incentives. Focusing on the markets discussed above with three stable match partners per person, in which the overall average number of offers per market is 61, the markets described as 70-70 and 20-20 have on average 65 and 69 offers. The other, unbalanced, markets have between 45 and 49 offers on average.[34]

The overall statistics of the dynamics show some degree of "exploration" on the part of our subjects. An offer from agent i to j is a *repeat offer* if i offered matching to j at least once before in the round. A match between i and j is a *repeat match* if i and j have been matched before and the the match was broken. The right-most columns of Table 4 reveal

is either 22 or 36, depending on the proposing side.

[32]Interestingly, the different extremal matchings are not associated with a higher volume of offers on either side. Our markets with two stable match partners for each participant are a good testing ground: in markets in which the food-optimal matching was realized, $48\% \pm 2\%$ of the offers were by foods, whereas in markets in which the color-optimal matching was realized, $56\% \pm 3\%$ of the offers were by foods.

[33]In two markets with a unique stable matching the number of steps the Gale-Shapley algorithm requires is either 22 or 36, depending on the proposing side.

[34]Interestingly, the different extremal matchings are not associated with a higher volume of offers on either side. Our markets with two stable match partners for each participant are a good testing ground: in markets in which the food-optimal matching was realized, $48\% \pm 2\%$ of the offers were by foods, whereas in markets in which the color-optimal matching was realized, $56\% \pm 3\%$ of the offers were by foods.

fairly low numbers of repeated offers and matches, suggesting that situations where subjects get matched, then try matching with others, only to later re-match with previous partners, are relatively infrequent.

We analyze the underpinnings of the dynamics that ensued, trying to match the scope of market activity with the insights on the likelihood of specific stable matchings, particularly the median stable ones. We start by assessing some existing theoretical models that allow for unravelling of matching markets sequentially. We then proceed to analyze the individual choices, pertaining to making as well as accepting offers, over time.

7.1. Dynamic Models and Simulations. The theoretical literature on dynamics in matching markets is not very developed.[35] The leading model for achieving a stable matching is the *random paths to stability* model of Roth and Vande Vate (1990). That model specifies a process in which, starting from some matching at time t, say μ_t, the set of all blocking pairs is calculated. One of these blocking pairs is chosen at random and created. That is, the corresponding color and food in that blocking pair get matched and their partners in μ_t (if they exist) become single. The resulting matching is defined to be μ_{t+1} and the process continues iteratively. Roth and Vande Vate (1990) proved that this dynamic converges to a stable matching with probability one.[36]

The random paths to stability process is inherently independent of the cardinal representation of preferences. In order to allow for such dependence we consider some variations of the basic model in which the probability a blocking pair forms depends on the welfare gain for the agents participating. Specifically, we consider two variations of the model:

- *Uniform.* As in the original Roth and Vande Vate (1990) setting, given a matching μ_t, all blocking pairs are equally likely to match and generate μ_{t+1}.
- *Exponential.* Given a matching μ_t, let $g_{f,c}$ be the gain in the sum of payoffs for food f and color c if they match. So $g_{f,c}$ is the sum of monetary payments that f and c get if

[35] Recently, there has been some work on decentralized matching processes that generate stable matchings with some imposed market structure, say allowing only one side of the market to make offers in sequence. See the literature review above for references.

[36] Blum, Roth, and Rothblum (1997) study a related process leading to stability for which the initial matching is one in which all blocking pairs are comprised of unmatched agents.

they match, after subtracting the payments f and c would get from their matches under μ_t. We assume the blocking pair (f, c) forms with probability proportional to $\exp(\lambda g_{f,c})$, where λ is a sensitivity parameter.[37] When λ is very low, the model approximates the uniform setting. As λ grows, blocking pairs that generate the greatest pairwise welfare gain become increasingly likely.[38]

In order to assess these models vis-à-vis our experimental results, we simulated these variations of the random paths to stability model using some of our experimental markets. The markets entailing three stable matches for each subject are particularly useful – they correspond to a unique ordinal preference profile, and generated persistent selection patterns of the stable matchings (in terms of the predominance of the median stable matching and the comparative statics regarding cardinal representations).

Ideally, the model would conform to the data in three ways. First, outcomes should be stable matchings. Second, the selection of stable matchings should be similar to the selection observed in our experiments, including a high incidence of median matchings. Third, the main features of the dynamic paths taken over time should be similar to those observed in our data.

We ran 10,000 simulations of each model using the underlying preferences of our markets entailing three stable match partners per participant. The main results are in Table 5. We present the results for the "$20-20$" and "$70-20$" treatments; these are two instances of the same (ordinal) market with three stable match partners (the models' predictions for the $20-20$ markets would be identical to those corresponding to the $70-70$ markets). Note that the last row of the table has the experimental results. The three dynamics are guaranteed, by the result of Roth and Vande Vate (1990), to converge to a stable matching, hence we replicate the first aspect of the data by design.

The results on selection depend on the sensitivity of the dynamics to the cardinal gains in payoffs. The uniform model predicts the selection in the $20-20$ market reasonably well, but gets the $70-20$ market wrong. In the exponential model, when we choose a relatively large value of λ, we see convergence to the median matching; so we can infer that when the

[37] This process is reminiscent of the network formation process studied by Jackson and Watts (2002).

[38] We also ran "power" models, where (f, c) forms with probability proportional to $g^k_{f,c}$. The results are similar.

	20-20 Marginals				70-20 Marginals			
	C-Optimal	F-Optimal	Median	Offer Volume	C-Optimal	F-Optimal	Median	Offer Volume
Uniform	30.9%	10.7%	58.4%	682	31.8%	10.1%	58.1%	700
	(0.5%)	(0.3%)	(0.5%)		(0.5%)	(0.3%)	(0.5%)	
Proportional	38.5%	3.8%	57.8%	182	47.6%	0.6%	51.8%	117
	(0.5%)	(0.2%)	(0.5%)		(0.5%)	(0.1%)	(0.5%)	
λ=0.0005	32.3%	7.7%	60.0%	450	36.1%	5.2%	58.7%	417
	(0.5%)	(0.3%)	(0.5%)		(0.5%)	(0.2%)	(0.5%)	
λ=0.001	33.0%	6.2%	60.8%	319	39.6%	3.3%	57.1%	265
	(0.5%)	(0.2%)	(0.5%)		(0.5%)	(0.2%)	(0.5%)	
λ=0.0025	35.4%	3.1%	61.6%	143	47.9%	0.7%	51.4%	98
	(0.5%)	(0.2%)	(0.5%)		(0.5%)	(0.1%)	(0.5%)	
λ=0.005	33.0%	1.5%	65.5%	78	56.9%	0.0%	43.1%	42
	(0.5%)	(0.1%)	(0.5%)		(0.5%)	(0.0%)	(0.5%)	
λ=0.01	22.9%	0.8%	76.3%	57	64.4%	0.0%	35.6%	24
	(0.4%)	(0.1%)	(0.4%)		(0.5%)	(0.0%)	(0.5%)	
λ=0.025	18.1%	0.8%	74.5%	42	74.4%	0.0%	25.6%	15
	(0.4%)	(0.1%)	(0.4%)		(0.4%)	(0.0%)	(0.4%)	
λ=0.05	12.3%	0.7%	87.0%	33	87.9%	0.0%	12.1%	10
	(0.3%)	(0.1%)	(0.3%)		(0.3%)	(0.0%)	(0.3%)	
Experiment	23.7%	10.5%	65.8%	68	16.7%	0.0%	83.3%	49

Standard deviations appear in parentheses.

Table 5: Simulation Results of Random Blocking Pairs Dynamics

probability of a blocking pair is sensitive to the cardinal gain in the pair, we tend to generate many median matchings. Nonetheless, for all values of λ, the models have problems predicting the large observed number of median matches in the $70-20$ market. In addition, the simulations underestimate the frequency of food-optimal matches. Only for the $20-20$ market does the uniform distribution over blocking pairs get close to the number of food-optimal matches, but generates too few median matches.

As a description of the dynamic path taken to stability, we focused on the volume of offers made. The uniform distribution over blocking pairs gets the volume of offers wrong by an order of magnitude. For the exponential model, however, there are values of λ that would replicate the mean volume of offers. These values of λ also predict the selection of matchings in the $20-20$, but not the $70-20$ market. Furthermore, while the mean value of offers is right, the variation in the number of offers in the simulations is substantially higher than that observed in our experiments (the standard deviation is consistently of the order of magnitude of the mean number of offers). There are frequently realized dynamics with much larger, and much smaller, numbers of offers than what we see in the data, even when the parameters are calibrated so that on average the number of offers corresponds to what we see in the data.[39]

The simulations also display more repeat offers (f makes an offer to c more than once, or vice versa) than those we document experimentally. For example, when blocking pairs are generated uniformly, 91% of matches in our simulations are repeat matches. In order to address this feature of our experimental markets, we also simulated the above versions of the model without allowing for repeat offers. Such an exercise leads to very infrequent ultimate stable matchings (appearing in less than 5% of the simulated markets). As a result, we rule these alternative models as not fitting the basic attributes of the data.

In addition, we considered a variation of the Gale-Shapley algorithm in which the proposer is selected at random and makes offers in order of their preferences, as in the original algorithm. Responders hold on to their best available offer (and make an offer to a superior potential partner that has not rejected them yet if chosen as proposers at a later stage). By construction, such a (random) procedure involves no rematches. Intuitively, since the original Gale-Shapley algorithm generates the stable matching preferred by the proposing side, providing both sides with equal shots of making offers seems to potentially shift outcomes toward intermediate solutions. Unfortunately, this dynamic process also generates very few stable matchings (4% of $10,000$ simulations of our three stable-match partners market), and a lower incidence of the median matching than our "uniform" dynamic described above (only 37% of simulations culminating in a stable matching corresponded to the median matching).

7.2. Individual Dynamics. We now proceed to the main determinants of subjects' behavior over time. We consider the factors driving market participants to make and accept offers. There are three layers to our analysis: 1. What makes offers likely? 2. Conditional on an offer made, to whom is it directed? 3. Conditional on receiving an offer, when is it likely to be accepted?

Considering the generation of offers, we start by noting that offers appear to be driven by the volume of potential blocking pairs in the market. This could be deduced indirectly from the evidence showing how blocking pairs vanish over time. Figure 3 illustrates directly the correlation between volume of offers and number of blocking pairs (that will be echoed in the

[39] This variation causes the mean number of offers in our simulations to differ slightly across the cardinal treatments when blocking pairs are determined uniformly, independent of payoffs.

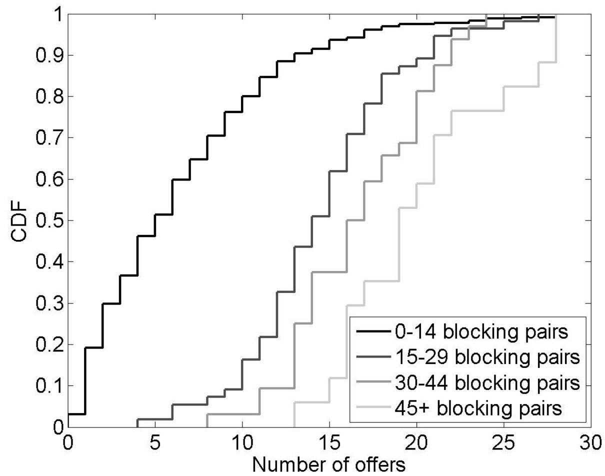

Figure 3: Cumulative Distributions of Offers for Different Volumes of Blocking Pairs

regression analysis below).

We now turn to the target of an offer, addressing the question: conditional on an agent making an offer, who does the offer go to? We use a discrete choice model to identify the variables that induce a subject to make a particular offer. The first column of Table 6 contains the results of a conditional logit estimation explaining offer targets (where errors are clustered by subject and marginal effects evaluated at the mean values of the regressors are reported). We use the following regressors, some of them are continuous variables while others are dummies. *Payoff advantage* is the increase in payoff to the subject if the offer is accepted; the same variable followed by "d" is a dummy that takes the value of 1 when *Payoff advantage* is positive. *Blocking Pair* is a dummy for when the pair formed by the proposer and the receiver is a blocking pair. *Number of rejections* counts the number of times that the recipient has rejected the proposer. *Fairness* is the absolute value of the difference between the proposer's rank in the recipient's preference list, and the recipient's rank in the proposer's preference list;[40] and *Recipient matched* is a dummy variable for whether the recipient is matched. We interacted all the variables with a dummy for the second half of the round of the market, to detect changes in behavior late in the market.

The findings are very intuitive. As remarked above, *whether* an offer is made is affected

[40] A specification using the difference in payoffs obtains similar results

Dependent Variable	Offer Target	Offer Acceptance
Payoff advantage	0.0890*	0.0009*
	(0.0053)	(0.0001)
Number of rejections	0.0288*	-0.0788*
	(0.0123)	(0.0138)
Fairness (rank)	-0.0196*	-0.0288*
	(0.0021)	(0.0049)
Recipient matched (d)	-0.0274	0.3481*
	(0.0142)	(0.0424)
Blocking pair (d)	0.1276*	-0.0125
	(0.0197)	(0.0413)
Payoff advantage (d)	0.1291*	0.3434*
	(0.0134)	(0.0514)
Payoff advantage * second half (d)	-0.0512*	0.000
	(0.0067)	(0.002)
Number of rejections * second half (d)	-0.0199	0.0089
	(0.0131)	(0.0211)
Fairness (rank) * second half (d)	-0.0009	-0.0069
	(0.0028)	(0.0102)
Recipient matched (d) * second half (d)	-0.0503*	-0.0669
	(0.0139)	(0.0345)
Blocking pair (d) * second half (d)	0.0119	0.0771
	(0.0202)	(0.0821)
Payoff advantage (d) * second half (d)	0.0768*	-0.0396
	(0.0330)	(0.0690)
$\chi 2$	1727.89	781.22
Log-Likelihood	-7072.59	-2297.88
Observations	41064	5134

(d) denotes discrete change of variable from 0 to 1. (*) indicates significance with 95% confidence. Standard errors appear in parentheses and are clustered by subject.

Table 6: Conditional Logit Results for Offer Targets and Offer Acceptances

by the presence of blocking pairs. In Table 6 we see that blocking pairs explain who receives offers as well; the dummy variable *Blocking Pair* is significant. The results are evidence of a certain degree of sophistication by the subjects: both the own payoff advantage from forming a pair (as captured by *Payoff advantage*), and the partner's payoff advantage (captured by the pair forming a blocking pair) are taken into account. The number of past rejections, a fairness wedge, and the potential partner being matched all reduce the likelihood of an offer in a significant, but more limited way. In addition, for the most part, interaction terms corresponding to behavior further along within a round are not significant.

The second column in Table 6 considers the factors behind the acceptance of an offer. If an agent receives an offer, when does he or she accept it? We use a binary logit model (errors clustered by subject, marginal effects evaluated at the mean of the regressors reported). *Recipient matched* in this regression is a dummy for whether the recipient of the offer is matched before receiving the offer. The results are in line with those seen for proposers. Offers are likely to be accepted if they result in higher payoffs. One somewhat puzzling finding is that subjects are more likely to accept an offer when they are matched than when they are not (as captured by the significance of the *Recipient matched* variable). This could possibly reflect some form of selection: desirable partners tend to get many offers.[41] Echoing proposers's behavior, fairness has a positive impact on receivers' responses. In addition, acceptance patterns appear stable throughout the round.

8. Conclusions

We reported results from an array of experiments looking at decentralized matching markets. With respect to outcomes, there are three main insights that shine through. First, decentralized markets very often culminate in stable matchings. Second, there is a persistent pattern in the selection of a stable matching: the median stable matching has very strong drawing power. Third, cardinal representations of preferences impact the distribution of selected matchings that are not the median one. Roughly speaking, the side of the market that has "more to lose" from forgoing their favorite stable matching, is more likely to implement it.

These results are important from a market design perspective. On the one hand, they reinforce the appeal to create institutions that implement a stable matching. Indeed, the results 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, they 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 Gale-Shapley algorithm that, absent incentive compatibility concerns, generate extremal stable matchings (favored by one side of the market). Our

[41] An alternative explanation is that unmatched recipients are more often found at the start of a round, when they may be waiting on the response to an offer of their own, or hope to get better offers. This explanation is inconsistent, however, with the insignificance of the interaction term with the second half of the round.

decentralized processes often generated the median stable matching. The utilitarian welfare comparison of these matchings may therefore be important in determining when centralized clearinghouses could be particularly beneficial.

With respect to dynamics, the arrival and selection of stable matchings occurs rather rapidly (in terms of both time and volume of offers). The action in our experiments is mainly driven by an attempt to exploit blocking pairs: the timing of offers, and who they are made to, are to a large extent explained by the existence of blocking pairs. Nonetheless, existing models that generate stable matchings through the sequential formation of blocking pairs à la Roth and Vande Vate (1990) do not match basic features of our data. There is therefore room for future theoretical work on dynamic matching processes that would likely provide foundational guidance on the selection of stable matchings in decentralized markets.

Another interesting avenue for future research pertains to incomplete information. Our experiments were designed as a first step toward understanding how decentralized markets operate and were constructed to match the cooperative game theoretical setting which, under complete information, predicts outcomes to be in the core, i.e. comprise stable matchings. We do not have an equivalent theoretical benchmark for matching under incomplete information. Nonetheless, in many large markets such as those for medical residents, law clerks, or marriage partners, participants presumably have incomplete information on some participants' preferences. While the theoretical underpinnings of the core under incomplete information are still developing, we find this to be an important direction to pursue both theoretically and experimentally.

References

[1] Bergstrom, Ted and Rod Garratt (2010), "Choosing Partners," mimeo.

[2] Blum, Yosef, Alvin E. Roth, and Uriel G. Rothblum (1997), "Vacancy Chains and Equilibration in Senior-Level Labor Markets," *Journal of Economic Theory*, **76**, 362-411.

[3] Cheng, Christine T. (2008), "The Generalized Median Stable Matchings: finding them is not that easy," *LATIN 08 Proceedings of the 8th Latin American Conference on Theoretical Informatics*.

[4] Echenique, Federico, Alistair J. Wilson, and Leeat Yariv (2010), "Clearinghouses for Two-Sided Matching: An Experimental Study," Caltech Social Science Working Paper Number 1315.

[5] Featherstone, Clayton and Eric Mayefsky (2010), "Stability and Deferred Acceptance: Strategic Behavior in Two-Sided Matching," mimeo.

[6] Featherstone, Clayton and Muriel Niederle (2011), "School Choice Mechanisms under Incomplete Information: An Experimental Investigation," mimeo.

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

[8] Gale, David and Lloyd Shapley (1962), "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 9-15.

[9] Gusfield, Dan and Robert W. Irving (1989), *The Stable Marriage Problem: Structure and Algorithms*, MIT Press.

[10] Harrison, Glenn W. and Kevin A. McCabe (1992), "Testing Noncooperative Bargaining Theory in Experiments," *Research in Experimental Economics*, **6**, 53-129.

[11] Haruvy, Ernan and M. Utku Ünver (2007), "Equilibrium Selection and the Role of Information in Repeated Matching Markets, *Economics Letters*, **94**, 284-289.

[12] Irving, Robert W. and Paul Leather (1986), "The Complexity of Counting Stable Marriages," *SIAM Journal on Computing*, **15(3)**, 655-667.

[13] Jackson, Matthew O. and Alison Watts (2002), "The Evolution of Social and Economic Networks," *Journal of Economic Theory*, **106**, 265-295.

[14] Kagel, John H. and Alvin E. Roth (1995), *The Handbook of Experimental Economics*, Princeton University Press.

[15] Kagel, John and Alvin E. Roth (2000), "The Dynamics of Reorganization in Matching Markets: A Laboratory Experiment Motivated by A Natural Experiment," *Quarterly Journal of Economics*, **115(1)**, 201-235.

[16] Kearns Michael, Siddharth Suri, Nick Montfort (2006), "An Experimental Study of the Coloring Problem on Human Subject Networks," *Science*, **313**, 824-827.

[17] 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.

[18] Niederle, Muriel and Alvin E. Roth (2009), "Market Culture: How Rules Governing Exploding Offers Affect Market Performance," *American Economic Journal: Microeconomics*, **2(1)**, 199-219.

[19] Niederle, Muriel and Leeat Yariv (2011), "Decentralized Matching with Aligned Preferences," mimeo.

[20] Pais, Joana (2008), "Incentives in Decentralized Random Matching Markets," forthcoming in *Games and Economic Behavior*.

[21] Pais, Joana and Ágnes Pintér (2008), "School Choice and Information: An Experimental Study on Matching Mechanisms," *Games and Economic Behavior*, **64(1)**, 303-328.

[22] Plott, Charles R. and Vernon L. Smith (1978), "An Experimental Examination of Two Exchange Institutions," *The Review of Economic Studies*, **45(1)**, 133-153.

[23] Roth, Alvin E., and Marilda Sotomayor (1990), *Two-sided Matching: A Study in Game-Theoretic Modelling and Analysis*, Cambridge University Press, Cambridge, England.

[24] Roth, Alvin E. and John H. Vande Vate (1990), "Random Paths to Stability in Two-Sided Matching," *Econometrica*, **58(6)**, 1475-1480.

[25] Smith, Vernon L. (1962), "An Experimental Study of Competitive Market Behavior," *Journal of Political Economy*, **70**, 111-137.

[26] Teo, Chung-Piaw and Jay Sethuraman (1998), "The Geometry of Fractional Stable Matchings and Its Applications," *Mathematics of Operations Research*, **23(4)**, 874-891.