Timing in Dynamic Matching Markets: Theory and Evidence*

Erdem Yenerdag†

December 29, 2021

Click here for the latest version

Abstract

We study a two sided one-to-one matching model with two periods in which agents decide to match early and exit in the first period, or to wait and search for more agents to find a match in the second period, where a stable matching is implemented for those who remain in the market. The distribution of the quality of a potential match varies over time as some agents who have found mutually agreeable matches exit the market. We show that in equilibrium: (i) similar and relatively high type pairs match early, (ii) the probability of matching early is a non-monotonic function of type, and (iii) markets do not necessarily unravel even if each meeting is costly and agents have the option to make exploding offers. We also designed experiments with real time interactions to test our theoretical predictions and provide an extensive analysis of early matching incentives in a dynamic matching environment. In the experiments, we turn on and off the possibility of matching with partners from previous periods and vary the cost of meeting in each period and the number of periods. We find that the results are in line with our theoretical predictions.

Keywords: Matching, incomplete information, stability, unraveling, experiments.
JEL Classification: C72, C78, C90, D82.

*I am deeply indebted to my advisor Brian Rogers for numerous insightful discussions regarding this work, and for his continuing guidance and support. I would like to thank SangMok Lee for useful discussions, and Ana Babus for many fruitful conversations. I also wish to thank Yann Bramoullé for valuable discussions and guidance. For helpful feedback, I would also like to thank Yan Chen, Edward Honda and participants of WashU's theory and research groups. For helping me run the experiments in the lab, I thank Laura Hurtado-Moreno, Stella Papadokonstantaki and Shih-Wen Liu. I gratefully acknowledge that this study was funded by Weidenbaum Center Impact Grant. This study was approved by WashU's committee on human subjects as IRB 202107189 on July 29, 2021.

†Department of Economics, Washington University in St. Louis; erdemyenerdag@wustl.edu

1 Introduction

Individuals' decisions on when to match play a critical role in many matching markets' ability to function, and yet it is a little-studied and often disregarded problem in the market design literature. For instance, in entry-level labor markets for professionals it is known that participants might have incentives to arrange an early match, just a little earlier than their competitors. The main concern with the timing of matching, or contracting, relies in the potential loss of efficiency due to market-wide mismatches, generated because some participants halt searching for other opportunities and contract too early. A recent September 2021 announcement from the American Economic Association (AEA), for instance, reveals the concerns of market-wide implications of participants' timing decisions in the market for new Ph.D. economists. In the announcement, the AEA attempts to establish a uniform timeline before which interviews and fly-outs should not be conducted. Similar concerns also exist in centralized matching markets where participants arrange the interviews in a decentralized way before submitting their rank-order lists to a centralized clearing house, which determines the matching outcome.[1] [2]

Markets in which agents have incentive to arrange an early match share two important features. First, there is an accepted normal timeline for matching, but firms and workers have opportunities to contract earlier than this normal timeline. Second, market participants search over time and conduct interviews sequentially. Particularly, workers and firms can conduct additional interviews before they reach an agreement. Therefore, early matching decisions must take into account the distribution of options available at later dates, which are themselves determined by participants' early matching decisions. In other words, workers and firms are essentially facing a non-stationary problem in making decisions on when to match.

Nevertheless, the literature studying these markets has thus far simplified their analysis by imposing at least one of the following assumptions: (*i*) *stationary distribution of individual qualities* and (*ii*) *prohibition of matching in a later date, if an early matching opportunity is missed* (see the literature review below for more detailed discussion)[3]. We call the latter the *no recall* assumption.

In this paper, we study individuals' early matching incentives in a simple dynamic matching model, without imposing the above-mentioned simplifying assumptions. We theoretically and experimentally show that each of these assumptions has profound impact on

[1]See, for instance, the National Resident Matching Program, (NRMP), and a descriptive model of NRMP by Echenique et al. (2020)

[2]Roth and Xing (1994), Roth (2010) show that regardless of the market structure, centralized or decentralized, almost all matching markets suffer from early matching inefficiencies.

[3]These assumptions are not specific to matching literature that studies labor markets. In general, dynamic matching literature has focused on the long-run steady state analysis from a central planner's point of view, Akbarpour, Li and Gharan (2020) and Baccara, Lee and Yariv (2020)

individuals' timing decisions as well as on market-wide matching outcomes. Our theoretical model is based on Bramoullé, Rogers and Yenerdag (2021). In this paper, we generalize the model and provide a more detailed analysis of incentives of early matching.

Specifically, our baseline model is a one-to-one two sided matching market over two periods. We assume that agents in both sides have privately known types drawn from a common distribution over a compact interval. Matching with higher type always yields higher utility, and agents are risk neutral. Agents randomly meet each other over two periods. In the first period, each agent on one side meets an agent on the other side. Upon meeting, they observe each other's type and decide to match and exit the market or wait to meet more agents and match in the second period. A pair of agents who meet in the first period can match and exit the market if, and only if, there is a mutual agreement. Otherwise, they both remain in the market. We assume that a stable matching is formed for those who remained in the second period market. In our model, matching in the second period corresponds to the normal timeline of the market, whereas any matching realized in the first period corresponds to *early matching*.

For the second period matching, we assume that a pair of agents who did not meet in either periods cannot form a blocking pair. That is, the stability in the second period is constrained by the meetings. We assume that agents who did not match in the first period, meet at most $k-1$ more agents in the second period. In particular, we assume $k \geq 2$ and it is binding in the sense that no agent can meet with all the agents from the other side. In other words, second period matching is stable with respect to k[4]. Common preferences and continuous type distribution imply that there exists a unique stable matching in the second period with probability 1. Note that a pair who did not match early in the first period can still be matched in the second period provided that it is a stable pair. As discussed above, this feature is ruled out in the existing literature by imposing the no recall assumption. Because of this, we call our model "matching with perfect recall" and show that this has a profound impact on individuals' early matching incentives.

We show that the existence and characterization results of Bramoullé, Rogers and Yenerdag (2021) hold in our setting. In particular, there exists a symmetric threshold equilibrium that is non-trivial, strictly increasing and continuous. According to our partial characterization, any such equilibrium induces some agents to match early. Moreover, we find that very low types and the highest type do not match early, and the probability of matching early as a function of type is non-monotonic. Also, based on our numerically computed threshold equilibrium with uniform type distribution, the probability of matching early is an inverse U-shaped function of type.

We also provide a detailed analysis of *unraveling*, which refers to a complete collapse of markets causing all matchings to be extremely inefficient as all participants in the market

[4]Similarly, Echenique et al. (2020) defines a stable matching constrained by the interviews.

are matching in a completely random fashion. Empirically, it is known that unraveling is a rare event when markets are able to produce stable matching (McKinney, Niederle and Roth (2005)). However, some theoretical studies have shown that markets can unravel even under stable matching due to a very interesting channel similar to strategic complementarity. According to this, if some agents match early, it becomes more attractive for other agents to match early, which makes it more attractive for more agents to match early – and so on and so forth (see,Echenique and Pereyra (2016))[5]. In our model, however, we show that implications of perfect recall and equilibrium type distribution in the second period halt this kind of strategic unraveling result. On the other hand, unraveling is also an equilibrium in our model. However, it simply derives from a coordination failure similar to the empty graph problem in network formation models. Thus, in this paper we are able to provide an explanation of why unraveling is a rare event.

In addition, we show that our non-trivial threshold equilibrium is robust. As it is shown in Bramoullé, Rogers and Yenerdag (2021), when there is a participation cost, the equilibrium threshold strategy is similar to the equilibrium with no cost. This result sharply contrasts with Damiano, Li and Suen (2005), which studies our model with no recall assumption. They show that when the participation cost is too small, the only equilibrium is unraveling. The no recall assumption can also be interpreted as "exploding offers," which is considered an important cause of unraveling (see, e.g., Roth and Xing (1994) and McKinney, Niederle and Roth (2005)). By extending our model, we show that our non-trivial threshold equilibrium survives even if firms can make exploding offers.

In this paper, we also design an experiment to test our theoretical predictions and provide an extensive analysis of agents' early matching incentives in a dynamic matching environment. In our experimental sessions, participants engage in a sequence of decentralized dynamic matching markets with incomplete information varying in three dimensions: agents' ability to recall, which is either no recall or perfect recall; per period participation cost, from no cost to a small cost; and the number of periods, which is either two or three periods.

In our experiments, the size of the markets is fixed to 20 subjects (10 each side). The baseline treatment is our model, matching with perfect recall. The experimental matching markets are decentralized: subjects can make offers to one another. The only restriction in the design is that only one offer can be made or accepted at a time. Subjects can also reject, withdraw and withhold offers without any restriction. We also designed treatments with no recall, which corresponds to the model in Damiano, Li and Suen (2005). Thus, our design allows us to capture the implications of the no recall assumption on individuals' early matching incentives.

[5] This channel was first mentioned by Roth and Xing (1994) (see, page 1030) and later systematically studied by Echenique and Pereyra (2016)

According to the results obtained in the experiments, matchings are predominantly stable. 74% of our markets ended up in stable outcomes. Moreover, markets that did not end up being stable were actually very close to the stable matching in terms of the number of blocking pairs and payoffs. This result is very important for our analysis, as we use a strategy method to directly elicit agents' threshold strategies in the first period. Thus, for this to be an effective method, it must be the case that subjects have a good estimate of their expected outcome in the second period, which is determined by a stable matching. Therefore, with this result we have strong evidence that subjects in our experiments understood the implications of stability in the second period when deciding whether to match early or wait.

In line with our theoretical predictions, agents' threshold strategy is increasing in type and inducing some types to match early. The empirical frequency of matching early as a function of type closely follows the theoretical prediction: it is a non-monotonic inverse U-shaped function. On the other hand, we also observed several systematic deviations from the theoretical benchmark. First, in the experiment more people are matching early compared to the theory. Second, few but a non-negligible number of low type agents are also matching early. We explain these patterns by risk aversion, since our theoretical predictions are derived under the assumption of risk neutrality.

Our experimental design allows us to make within subjects and between subjects comparisons of treatment effects. By comparing treatments with recall and no recall, we empirically confirm that the possibility of matching with partners from previous periods has profound impact on agents' early matching incentives. In line with our theoretical predictions, the number of agents who matched early in our baseline treatments is strictly less than in the treatments with no recall (around 50% less). In addition, our statistical analysis confirms that early matching is more assortative in our baseline treatments with recall.

In treatments with perfect recall, we empirically confirm that agents' behavior does not change dramatically when a participation cost is introduced. However, different from what Damiano, Li and Suen (2005) predicted, varying participation cost did not lead to a statistically significant change in behavior in treatments with no recall. In other words, our experimental markets did not unravel. On the other hand, in the no recall treatments with costly participation, we recorded some markets in which more than 50% of the agents matched early, which is not observed in any other treatments.

In what follows, we present a review of the related literature. In section 2, we describe our baseline model. In section 3, we present existence and characterization theorems and provide a theoretical explanation of why unraveling is a rare event in our model. We then describe our experimental design and present our empirical results in section 4. Section 5 concludes.

1.1 Literature Review

Our paper contributes to the literature of dynamic matching in general. A recent but growing literature in dynamic matching takes a stationary approach to market design, and studies implications of timing from the planner's perspective (Akbarpour, Li and Gharan (2020); Baccara, Lee and Yariv (2020); Ashlagi, Nikzad and Strack (2019); Leshno (2019)). Our paper, however, takes a non-stationary approach and studies individual incentives of timing in matching markets.

Markets can unravel for multiple reasons. Li and Rosen (1998) and Li and Suen (2000) study a two period matching market with transfers and with risk averse agents. In their model, agents learn their type in the second period. Under these assumptions, they show that early matching is an insurance device. In addition to this, a recent paper, Vohra (2020), provides an explanation of unraveling by the transparency of secondary markets.

When risk aversion and other reasons are not too much of a concern, Roth (1991) and Kagel and Roth (2000) provide empirical evidences that markets that can generate a stable matching are more resilient against unraveling. In line with this argument, for example, McKinney, Niederle and Roth (2005) show that unraveling is a rare event if the market is able to generate a stable matching. They show that the market for gastroenterologists unraveled due to a rare event: an unusual, one-time shock to the demand and supply.

On the other hand, there is a sequence of theoretical papers that show that markets can unravel under a stable matching even if agents are risk neutral, and supply and demand of the market is perfectly balanced. In a seminal paper, Roth and Xing (1994) provide an explanation of *strategic unraveling* under these conditions. According to the strategic unraveling, agents match early because of their concerns that others match early. Roth and Xing (1994) study a two period matching market with incomplete information and show that markets can still unravel under a stable matching. However, they explicitly impose the no recall assumption in their model. More recently, Echenique and Pereyra (2016) also study a two period matching market with incomplete information in which agents learn their type only in the second period. They explicitly quantify the strategic unraveling effect on agents' early matching incentive and show that unraveling is the unique equilibrium when the market size is not large. However, their assumptions imply that agents' early matching decisions do not affect the type distribution in the second period.

In our model, on the other hand, agents decide to match early or wait with perfect recall. Moreover, agents' decisions on matching early induce an endogenous type distribution in the second period. We show that, in equilibrium, the good-but-not-best agents will assortatively match early, however the market does not unravel due to the strategic unraveling. Unraveling is a trivial equilibrium in our model, and it simply derives from a coordination failure similar to the empty graph problem in network formation models.

In this sense, our paper is the first theoretical and experimental study that shows that unraveling is a rare event when agents are risk neutral.

Damiano, Li and Suen (2005) also study a two period matching market with incomplete information. As in our model, agents' decisions on matching early induce an endogenous type distribution in the second period. However, they impose the no recall assumption and show that unraveling is the unique equilibrium when participation is costly. In their model, with the no recall assumption, each agent has the exact same continuation value for participating in the second period. When agents have perfect recall, on the other hand, we show the continuation value of each agent is strictly increasing in type and markets do not unravel even if there is a participation cost. As we show below, the strategic unraveling channel also exists in Damiano, Li and Suen (2005).

Du and Livne (2016) consider a two period matching model with incomplete information. In their model, new ex-ante identical agents arrive in the second period. In the first period, agents have complete information as everyone observes the type of the other agents. Thus, agents' early matching decisions do not alter the type distribution of the agents that arrive in the second period. They study the role of transfers in unraveling, and show that under a flexible-transfer regime, the market will not unravel.

Halaburda (2010) shows that similarity of preferences plays a key role in unraveling. In Halaburda (2010), workers have commonly known identical preferences. If firms' preferences are sufficiently similar then some firms have incentives to match early, which can lead to unraveling. In our model, however, agents also have common preferences. Some agents are matching early in equilibrium, but the market does not unravel.

Our paper also contributes to the experimental matching literature. There are fairly few experimental studies of decentralized matching. Echenique and Yariv (2012), Pais, Pintér and Veszteg (2012) and some other recent studies (He, Wu and Zhang (2020) and Agranov et al. (2021)) investigate decentralized matching experimentally under different settings. These papers do not study agents' early matching incentives. Instead, the focus is how a stable match emerges as a result of decentralized interactions.

Kagel and Roth (2000) analyze the transition from unraveled decentralized matching markets to centralized clearinghouses. More related to ours, Niederle and Roth (2009) study a matching market with incomplete information. In their experiments, firms make exploding or open offers to workers over several periods. They analyze the implications of offer structure on the final matching and market efficiency.

2 Model

We study a two period model of a two sided one to one matching market. Agents have privately known types. Matching with higher type is more desirable. In the first period, each agent meets an agent from the other side. Upon meeting, agents observe each other's type, and can match and exit the market under mutual agreement. In the second period, each remaining agent meets multiple agents from the other side who remain in the market. We assume that a stable matching is formed for those who remained in the second period. Agents can keep track of previous meetings and can later match with those partners provided that the matching is stable. In our model, matching in the second period corresponds to the normal timeline of the market, whereas any matching realized in the first period corresponds to *early matching*.

We adopt the language of labor markets and call the two sides *workers* and *firms*. Let $\mathcal{W}$ and $\mathcal{F}$ be two disjoint *finite* sets representing the set of workers and firms respectively. Let $\mathcal{N} = \mathcal{W} \cup \mathcal{F}$ denote the set of all agents. Throughout the paper we assume that each firm has only one position to fill and that there are as many workers as firms, and the total number of agents is $\#\mathcal{N} = N \geq 4$. Each worker and firm has a privately known type drawn independently and identically from a probability distribution H. We assume that H is distributed on the interval $[a, b] \subseteq \mathbb{R}_+$, absolutely continuous with respect to the Lebesgue measure and admits a strictly positive and smooth probability density function h over the entire support. We let $w_i \in [a, b]$ denote a realized type of worker i. Analogously, $f_j \in [a, b]$ denotes a realized type of firm j. We refer to agents with their types. For example, we call a worker i with type w_i "worker w_i". We use capital letters W_i and F_j to denote the random type of worker i and firm j respectively. The subscripts will be superseded when there is no ambiguity.

Let $u_i(f_j|w_i)$ denote the utility of a worker $i \in \mathcal{W}$ of type $w_i \in [a, b]$ from *matching* with a firm of type $f_j \in [a, b]$. Analogously, let $u_j(w_i|f_j)$ denote the utility of a firm j of type $f_j \in [a, b]$ from *matching* with a worker w_i. That is, a matching between firm f_j and a worker w_i generates total surplus $u_i(f_j|w_i) + u_j(w_i|f_j)$. We assume that each agent has a type dependent outside option. We let $\mathcal{O}(x_i)$ be the utility of agent $i \in \mathcal{N}$ of type $x_i \in [a, b]$ from being unmatched. We impose the following assumptions on preferences:

Assumption 1. *For all agent $i \in N$ of type $x_i \in [a, b]$*

(*i*) *$u_i(y|x_i)$ is strictly increasing and linear in $y \in [a, b]$*

(*ii*) *$u_i(a|x_i) = \mathcal{O}(x_i)$.*

The first part of Assumption 1(i) is common in matching literature. It says that agents have **strict preferences** over the types of agents on the other side and matching with a higher type is more desirable. The last part of Assumption 1(i) implies that agents are **risk neutral**. Assumption 1(ii) says that agents are indifferent between matching with the lowest type and being unmatched. Notice that this implies that being unmatched is strictly worse than being matched with any agent with probability 1, since the type distribution is continuous. We impose this assumption for simplicity and tractability purposes. For instance, it allows us to normalize a, the lower bound of the type space, to be equal to zero without loss of generality.

Agents can match in period $t \in \{1, 2\}$. We assume that agents do not discount the time. In the first period ($t = 1$), a worker w *randomly* meets a firm f. Upon meeting, each agent learns the other agent's type, and each agent decides to "match early" or "wait" and seek a match in the next period. If both agents decide to match early, then they *match* and *exit* the market. If at least one agent decides to "wait", then they do not match in period $t = 1$ and remain in the market for the next period. The matching in the first period is called "early matching".

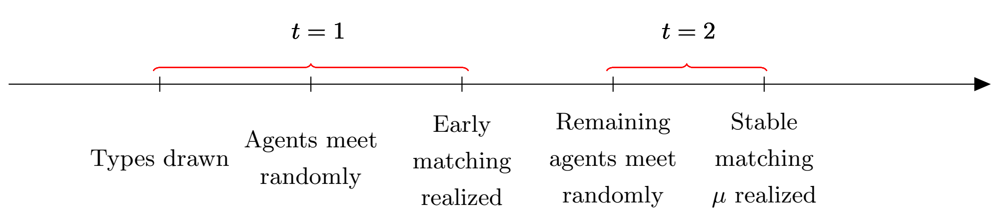
In the second period ($t = 2$), every *remaining* worker w *randomly meets* a new *remaining* firm. In the Appendix, we extend this model and prove the results with at most $k \in [2, \frac{N}{2} - 1]$ meetings over two periods. We call a pair (w, f) a "period t pair" if they meet in period $t \in \{1, 2\}$. A *matching* in the second period is denoted by μ. If a pair (w, f) is matched under μ then it is either a period $t = 1$ or $t = 2$ pair. We write $\mu(i) = j$ and/or $\mu(j) = i$ if a pair (w_i, f_j) is matched under μ. With slight abuse of notation, we let $\mu(x)$ denote the type of agent x's match under matching μ. We write $\mu(x) = \mathcal{O}(x)$ if agent x is unmatched under μ.

Definition 1 (Blocking Pair)**.** *A pair (w_i, f_j) is a blocking pair if*

(i) *(w_i, f_j) is a period $t \in \{1, 2\}$ pair, and*

(ii) *$u_j(\mu(f_j)|f_j) < u_j(w_i|f_j)$ and $u_i(\mu(w_i)|w_i) < u_i(f_j|w_i)$.*

A matching μ is *stable* if there does not exist a blocking pair[6]. The following timeline describes how events unfold:



Note that the stability in the second period is constrained by realized two period meetings, since a pair who did not meet cannot block a matching in the second period. Hence, a stable matching with respect to agents' full preferences is not guaranteed. This assumption

[6] Equivalently, we could define our second period matching as follows: At the end of the second period meetings, each agent only knows their own type and the type of the agents that they met over two periods. Based on this information, each agent who has not matched in the first period submits a rank order list only including the partners that he/she met in either period to a revelation mechanism that produces a stable match with respect to the submitted rank order lists. Instead of doing this, we introduced a constraint in the definition of blocking pair that we think it is more realistic for decentralized matching markets.

is realistic both for centralized and decentralized matching markets. In the US medical match, for example, each doctor has a limit on the number of interviews they can attend, while each hospital offers a limited number of interview slots. In the end, hospitals and doctors submit their rank order lists to the NRMP only for those they interviewed with. Hence, the matching implemented by the NRMP is not always stable with respect to agents' full preferences, even if everyone reports their rank order lists truthfully. Similarly, in decentralized matching markets like the Econ PhD job market, as schools cannot interview all the candidates, it is unlikely that the resulting match is stable with respect to agents' full preferences.

Recall that Assumption 1 implies that agents have common and strict preferences. Since the type distribution is also atomless, it is trivial to show that there exists a unique stable matching in the second period with probability 1. We state this result as a remark:

Remark 1. *For every early matching realization in period $t = 1$, there exists a unique stable matching μ in period $t = 2$ with probability 1.*

2.1 First Period Strategic Interactions: Early Matching

Upon meeting in the first period, each agent decides to "match early" or "wait" based on the expected utility from the second period stable matching μ. We note two important features of our model. First, agents who did not match in the first period can still match with each other in the second period, provided that the matching is stable. Second, since some agents might match early and exit the market, the strategic interactions in the first period yield an endogenous type distribution in the second period market. Thus, to calculate the expected utility in the second period, one needs to pin down the resulting type distribution first, then with respect to that, the distribution of all possible stable matchings.

In this paper we focus on symmetric threshold strategies[7]. Let $\mathcal{X}$ denote the set of all mappings on $[a, b]$ to itself. And, let $\mathcal{Y}$ be the set of all monotone functions in $\mathcal{X}$. That is,

$$\mathcal{Y} = \{s : [a, b] \to [a, b] | s(x) \geq s(x') \quad \forall x, x' \in [a, b] \quad s.t. \quad x \geq x'\}. \tag{1}$$

That is, $\mathcal{Y}$ is the set of all monotone functions on $[a, b]$. Formally, a threshold strategy is defined as follows:

Definition 2 (Threshold Strategy)**.** *A threshold strategy is a mapping $s \in \mathcal{X}$ such that an agent of type $x \in [a, b]$ decides to match early with type $y \in [a, b]$ if, and only if, $s(x) \leq y$.*

[7] For more details on the existence of threshold strategies as a best response to any strategy, see Bramoullé, Rogers and Yenerdag (2021)

To understand the strategic interactions in the first period, consider a worker i of type $w_i \in [a, b]$. Suppose that w_i meets with a firm $f \in [a, b]$ in the first period, and all other agents except i follows a threshold strategy s. Conditional on not matching early, the expected type of worker w_i's stable partner in the second period is denoted by

$$\mathbb{E}_s[\mu(w_i)|w_i, f]. \tag{2}$$

Let N_2 denote the number of agents remaining in the second period. Note that symmetric threshold strategy s induces a type distribution in the second period and a distribution of N_2. Thus, conditional on w_i's own type and the type of first period partner f, the expectation given in eq. (2) is calculated with respect to the distribution of all possible stable matchings.

The following notation is useful for analysis. For any $c \in [a, b]$, we let s^c denote the constant threshold strategy at c. That is, under s^c, agents decide to match early for all types above c, and decide to wait otherwise. One trivial constant threshold strategy is s^a. Note that under s^a everyone matches early with probability 1. If everyone follows trivial strategy s^a, worker w_i would be indifferent between "match early" and "wait." To see this, note that in that case w_i would end up matching with f in the second period even if w_i decides to wait. For this reason, we let[8]

$$\mathbb{E}_s[\mu(w_i) \mid w_i, f] \equiv f \quad \forall f \in [a, b] \quad \forall s \in \mathcal{X} \quad s.t. \quad s = s^a \ \lambda - a.e. \tag{3}$$

Also, note that s^a is a Nash equilibrium. This is a special equilibrium known in the literature as "unraveling".

Now, fix any arbitrary weakly increasing symmetric threshold strategy profile $s \in \mathcal{Y}$. Suppose that every agent follows s. For worker i of type w_i, consider the following set:

$$B_s(w_i) = \{f \in [a, b] : \mathbb{E}_s[\mu(w_i) \mid w_i, f] \leq f\} \quad \forall w_i \in [a, b]. \tag{4}$$

Note that for all type $w_i \in [a, b]$ of worker i, "match early with all types in $B_s(w_i)$, and wait otherwise" is a best response. More formally, fix any strategy profile of all other agents except i, and except the first period partner of i, say agent j. Denote this by $\boldsymbol{\sigma}_{-i,j}$. Then, for any fixed $\boldsymbol{\sigma}_{-i,j}$, $B_{\boldsymbol{\sigma}_{-i,j}}$ is always optimal regardless of agent j's strategy. For this reason, we call $B_{\boldsymbol{\sigma}_{-i,j}}$ a locally weakly dominant strategy.

Bramoullé, Rogers and Yenerdag (2021) proves that for all $s \in \mathcal{Y}$, $B_s(x)$ is a closed upper set of $[a, b]$. Moreover, $\min B_s(x)$ is weakly increasing and continuous in $x \in [a, b]$. Note that $\min B_s(x)$ is a "threshold type" of agent x under strategy s. That is, agent

[8] $\lambda-$a.e. stands for "almost everywhere with respect to measure λ." Here, we use λ to denote the underlying Lebesgue measure.

x is indifferent between matching early with $\min B_s(x)$ and waiting. Next, we give our equilibrium definition based on this local dominance notion:

Definition 3 (Symmetric Threshold Equilibrium)**.** *A threshold strategy $s : [a, b] \to [a, b]$ is a symmetric equilibrium if, and only if,*

$$s(x_i) = \min B_s(x_i) \quad \forall x_i \in [a, b], \tag{5}$$

where $B_s(w_i) = \{f \in [a, b] : \mathbb{E}_s[\mu(w_i) \mid w_i, f] \leq f\}$.

3 Equilibrium Analysis

Consider the trivial threshold strategy s^a. Note that this trivial threshold strategy satisfies our equilibrium definition given in Definition 3, which is known in the literature as the "unraveling" equilibrium[9]. Since it is a trivial strategy, we also call it the "trivial equilibrium." In this section, we show that there exists a non-trivial threshold equilibrium. All the proofs are presented in the appendix.

Theorem 1. *There exists an $\bar{N} \in \mathbb{N}$ such that every market size of $N \geq \bar{N}$ has a strictly increasing and continuous non-trivial symmetric threshold equilibrium.*

For our model described in Section 2, the proof of Theorem 1 follows directly from Bramoullé, Rogers and Yenerdag (2021). However, we are going to prove the above existence theorem for a model where workers meet at most $k-1$ firms in the second period with $k \in [2, \frac{N}{2} - 1]$. The intuition of the proof remains the same. Suppose that every agent follows a weakly increasing symmetric threshold strategy s with $s(x) > a$ for all $x \in [a, b]$ in the first period. After some agents match early and exit the market, every remaining worker meets $k-1$ remaining firms in the second period. Note that the second period meetings are basically $k-1$ independent draws from an endogenous type distribution induced by s, and it has full support since s is weakly increasing and bounded away from a. Moreover, meeting with $k \leq \frac{N}{2} - 1$ agents in total implies that no one is meeting with everyone in the second period with probability 1. Thus, under the stable matching in the second period, there exists a strictly positive probability of being unmatched for all types $x < b$. In turn, this means that even for the lowest type worker, there exists a strictly positive probability of matching with a firm in the second period whose type is drawn from an endogenous type distribution induced by s.

Building on this intuition, we prove that agents' option value from waiting is bounded away from their outside option for all symmetric non-trivial weakly increasing threshold strategies. For all ex-ante type distributions, we characterize this lower bound and establish

[9] This is due to Equation 3.

a best response mapping on a set of all weakly increasing threshold functions that are greater everywhere than this lower bound. Then, we show that the best response mapping is a self map and complete the proof of Theorem 1 by applying Schauder fixed-point theorem. One subtlety in the proof is that when the lowest type worker meets the lowest type firm in the first period, the stability of second period matching implies that they match with probability 1 if and only if $N_2 \leq 2k$. In particular, when threshold strategy s is too close to s^a, the probability of having at most $2k$ agents becomes high. To get around this problem, we require a sufficiently large number of agents $\bar{N}$ in period $t = 1$ so that the probability of $N_2 > 2k$ is high enough.

We now give our equilibrium characterization.

Theorem 2. *Suppose that β is a non-trivial threshold equilibrium. Then, β satisfies the following conditions:*

(i) $\beta(a) > a$ *and* $\beta(b) = b$

(ii) *There exists an* $x^* \in (a, b)$ *with* $\beta(x^*) = x^*$ *such that* $\beta(x) < x$ *for every* $x \in (x^*, b)$.

Theorem 2 is a partial characterization of any non-trivial threshold equilibrium β. It implies that, in equilibrium, the probability of no one is matching early equals to 0. To see this, note that a first period pair (x, y) match early under β if, and only if, $\beta(x) \leq y$ and $\beta(y) \leq x$. Equivalently, we can write this as $y \in (\beta(x), \beta^{-1}(x))$. Thus, by Theorem 2, we can conclude that $\mathbb{P}[Y \in (\beta(x), \beta^{-1}(x))] > 0$. We can also conclude from Theorem 2 that the probability of matching early as a function of type is non-monotonic. We show this in Corollary 1.

Corollary 1. *Under any non-trivial threshold equilibrium, some agents match early and the probability of matching early conditional on type is non-monotonic.*

In Section 4, we numerically show that the equilibrium probability of matching early as a function of type is an inverse $U-$ shaped function.

3.1 Recall vs No Recall: The strategic unraveling argument

Other related papers that study agents' early matching incentives characterize a very interesting channel called *strategic unraveling* that leads the market to unraveling. According to this channel, markets can end up being unraveled because when some agents match early, it becomes more attractive for other agents to match early, which makes it more attractive for more agents to match early – and so on and so forth (Echenique and Pereyra (2016)). To our knowledge, this was first explained in Roth and Xing (1994), then explicitly quantified by Echenique and Pereyra (2016). In this section, we show that this strategic unraveling channel does not exist in our model.

Roth and Xing (1994) and Echenique and Pereyra (2016) impose additional assumptions that simplify the expectation in Equation 2 and show that unraveling is an equilibrium. For example, Echenique and Pereyra (2016) assume that agents learn their type only in the second period. Note that with this assumption, agents' early matching decisions are type independent and do not induce an endogenous type distribution in the second period. Thus, the expectation in Equation (2) becomes just a constant for all types. Similarly, in Roth and Xing (1994), Equation (2) is simplified by the no recall assumption.

In our model, on the other hand, the expectation in Equation 2 is obtained with respect to the endogenous type distribution conditional on the agent's own type, the type of the first period partner, symmetric threshold strategy, and the stability of the second period matching. The strategic unraveling channel is halted because of the implications of perfect recall and endogenous type distribution in the second period. In order to understand the effect of the strategic unraveling, we need to specify who matches early in the first period as agents know their own type in our model.

For an illustrative example, let $[0,1]$ be the type space. First, assume that no agents match early in the first period. That is, all agents follow the threshold strategy $s^1(x) = 1$ for all $x \in [0,1]$. From the proof of Theorem 2, it is easy to show that there exists an $\epsilon > 0$ such that

$$\min B_{s^1}(1-\epsilon) = \min\{f \in [0,1] : \mathbb{E}_{s^1}[\mu(w)|w = 1-\epsilon, f] \leq f\} < 1-\epsilon. \tag{6}$$

That is, if no one matches early, worker $w = 1-\epsilon$ is willing to match early with a firm of type strictly less than $1-\epsilon$.

Now, assume that all agents follow a strategy s^* such that[10]

$$s^*(x) = \begin{cases} 1-\epsilon & \text{if} \quad x \in [0, 1-\epsilon], \\ 0 & \text{if} \quad x \in (1-\epsilon, 1]. \end{cases}$$

Note that under s^*, all agents with type $1-\epsilon$ and above match early. Moreover, any agent meeting a type greater than $1-\epsilon$ in the first period also matches early. In this case, it is straightforward to conclude that

$$\min B_{s^*}(1-\epsilon) = \min\{f \in [0,1] : \mathbb{E}_{s^*}[\mu(w)|w = 1-\epsilon, f] \leq f\} = 1-\epsilon. \tag{7}$$

In words, under s^* some agents match early, and the worker $w = 1-\epsilon$ is not willing to match early with any firm whose type is below $1-\epsilon$. This is because the strategy s^* induces a type distribution in the second period such that the worker $w = 1-\epsilon$ is the highest type worker in the second period with probability 1. Hence, the stability of μ implies that if

[10] Note that s^* is a decreasing threshold strategy.

$w = 1 - \epsilon$ meets a firm $f < 1 - \epsilon$ in the first period, by waiting, $w = 1 - \epsilon$ can meet with more firms in the second and can match with the highest one the second period.

Therefore, we can conclude that because agents' early matching decisions affect the type distribution in the second period, when some agents match early, matching early does not necessarily become more attractive for all agents in our model.

Damiano, Li and Suen (2005) study two period matching markets with incomplete information. In their model, agents learn their type first, then each agent meets an agent from the other side in each period. However, Damiano, Li and Suen (2005) impose the no recall assumption: if a pair of agents do not match in the first period, then they cannot match in the second period. Note that for $k = 2$ in our model, the only difference between Damiano, Li and Suen (2005) and our model is the no recall assumption. With this assumption, the expectation in Equation (2) is dramatically simplified.

Note that in Damiano, Li and Suen (2005), agents' early matching decisions induce an endogenous type distribution. However, due to the no recall assumption, the expectation in Equation (2) is constant for all types. That is, given a strategy s in the first period, every type expect the same outcome in average from the second period market[11]. As a consequence, the strategic unraveling channel also exists in Damiano, Li and Suen (2005).

To see this, suppose that each agent's type is independently distributed from a uniform distribution over $[0, 1]$. Suppose that every agent follows the threshold strategy s^1. That is, no agent matches early with probability 1. In that case, the expected match in the second period market is 0.5 for all types in Damiano, Li and Suen (2005). Now, suppose that every agent follows a threshold strategy $s^{0.75}(x) = 0.75$ for all $x \in [0, 1]$. That is, a period 1 pair (x, y) matches early if and only if $x \geq 0.75$ and $y \geq 0.75$. The joint distribution of a pair in the second period under $s^{0.75}$ is depicted in Figure 1. An agent who stays in the market meets and matches in the second period with an agent whose type is drawn from the marginalization of this joint distribution. Note that the marginal distribution of the green colored region over the unit square is dominated by the uniform $[0, 1]$ distribution in the sense of First Order Stochastic Dominance (FOSD). This implies that agents' expected stable match in the second period is strictly less under $s^{0.75}$ than under s^1, which is the case in which no agents match early.

[11]This essentially means that the stars of the market and the low quality candidates expects the same outcome from the market. This feature does not suit well most of the entry level job markets.

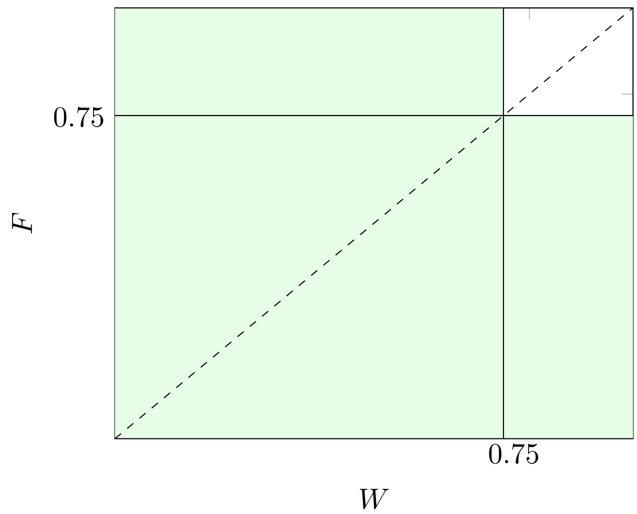


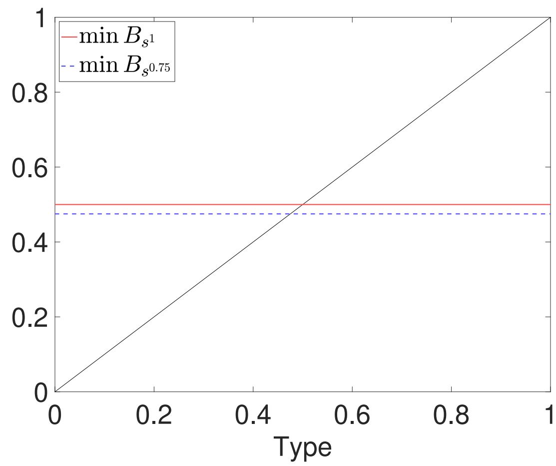
Figure 1: The joint distribution of a pair in the second period market under the strategy $s^{0.75}(x) = 0.75$ for all $x \in [0, 1]$. The support of the joint distribution is shown in green.

In general, more agents match early in the first period under a lower threshold strategy $s^{c'} < s^c$. In Damiano, Li and Suen (2005), the expected match in the second period is decreasing when c decreases uniformly. This is because the marginal type distribution in the second period is decreasing in the FOSD sense as s^c decreases uniformly, and in Damiano, Li and Suen (2005) agents match with whomever they meet in the second period. The left panel of Figure 2 depicts the expected match in the second period under the strategies s^1 and $s^{0.75}$. Hence, when some agents match early, it is more attractive for everyone in the market to match early in Damiano, Li and Suen (2005) as well.

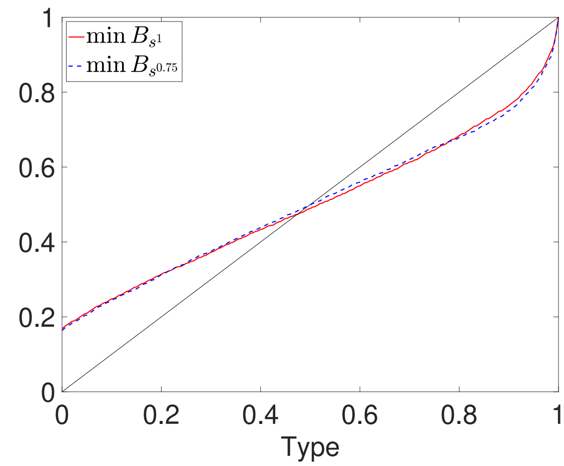
In our model, on the other hand, a uniformly lower threshold strategy does not necessarily decrease the expectation in Equation (2) for all agents. The right panel of Figure 2 depicts the numerically calculated best responses to the strategies s^1 and $s^{0.75}$ in our model.[12] According to the figure, when some agents match early under the strategy $s^{0.75}$, an agent of type $x \in [0.25, 0.75]$ is less willing to match early compared to the case in which no agents match early under s^1. Therefore, in our model the market does not suffer from strategic unraveling.

As a consequence of strategic unraveling, Damiano, Li and Suen (2005) show that when participating in each period is costly, unraveling is an always equilibrium. Moreover, when participation cost is sufficiently small, they show that unraveling is the unique equilibrium. The model in Damiano, Li and Suen (2005) is the no recall version of our model with $k = 2$.

[12] In the numerical computations, we fixed the market size to 20.



(a) No Recall



(b) Perfect Recall

Figure 2: Lowest acceptable type to match early in the first period under the threshold strategy s^1 and $s^{0.75}$. Left panel corresponds to the model with no recall, Damiano, Li and Suen (2005). Right panel corresponds to our model with $k = 2$.

Bramoullé, Rogers and Yenerdag (2021) prove that with perfect recall even if participating in each period is costly, unraveling is never a unique equilibrium. Moreover, equilibrium with costly participation with perfect recall is similar to the non-trivial equilibrium without participation cost[13]. In this paper, we experimentally test this result in Section 5.

The no recall assumption also has a behavioral extent. For example, it can also be interpreted as one side of the market making "exploding offers." Note that the no recall assumption implies that if a pair of agents miss an early matching opportunity in the first period, then it is impossible for the same pair to match in the second period. Essentially, this is a main feature of an exploding offer, and as it is shown in McKinney, Niederle and Roth (2005) and Niederle and Roth (2009), that it can lead the market to unraveling.

We now show that if firms have an option to make exploding offers, the non-trivial equilibrium, shown in Theorem 1-2, is still an equilibrium. We prove this result in the following proposition.

Proposition 1. *There exists a non-trivial threshold equilibrium β such that no firm makes an exploding offer.*

4 Experimental Design

We designed a decentralized matching experiment to study agents' early matching incentives in two sided one-to-one dynamic matching markets. The environment is designed to

[13]See, Theorem 3 in Bramoullé, Rogers and Yenerdag (2021).

test our theoretical results, stated in Theorem 2, as well as to capture the implications of the no recall assumption on agents' behavior in dynamic matching problems. In our experimental sessions, participants engage in a sequence of decentralized dynamic matching markets with *incomplete information* varying in three dimensions: agents' ability to recall, which is either no recall or perfect recall, per period participation cost, from no cost to a small cost, and the number of periods, which is either two or three periods. We describe each of these dimensions in turn. Sample experimental instructions are presented in the Appendix.

Our experimental two sided matching markets consist of 20 participants (10 on each side). The two sides are neutrally labeled as *Circles* and *Squares*. These are just semantics for firms and workers in labor markets, men and women in the marriage market, doctors and hospitals in the NRMP, etc. Hereinafter, we will refer to the two sides of the market as workers and firms. At the beginning of each round, participants privately observe their own *type*, phrased as "productivity level" in the experiment, which is an independent draw from uniform distribution over the interval $[1, 10]$.[14] In each period, participants randomly meet with a participant from the opposite side. Each meeting is *uniformly random* and *independent* of types. Upon meeting, two participants observe the type of each other. Participants could match with at most one participant from the opposite side, deriving a payoff in points that is equal to the product of the types. That is, if a worker of type $w \in [1, 10]$ is matched with a firm of type $f \in [1, 10]$, they each earn $w \times f$ points. Remaining unmatched resulted in a payoff equal to one's own type.

The main design variable in our dynamic matching markets is the agents' ability to recall. We say that participants have *perfect recall* if, and only if, their potential match partners in period t are the unmatched participants from the opposite side who they met with in every period $t' \leq t$. Similarly, we say that participants have *no recall* if, and only if, their potential match partner in period t is the participant from the opposite side who they met with in period t.

Markets with *perfect recall* were designed to mimic our theoretical matching model described in Section 2. In particular, participants were asked to input their *threshold type* after observing their own types and before the first period meeting realization. The threshold type is the lowest acceptable type of a participant for matching early with his/her first period potential match partner. A pair of participants match early and exit the market if, and only if, the type of each participant is greater than or equal to the lowest acceptable threshold type of the other participant. Participants who did not match early and remained in the market proceed to the second period. In the second period, each participant meets with a new remaining participant from the opposite side. At the end of the second period, there is a *decentralized matching protocol* in which participants can send match proposals

[14] In the experiment, we rounded the types up to two decimals points by preserving the realized ranking.

to their *potential match partners*. In three period matching markets with *perfect recall*, participants who are unmatched at the end of the second period randomly meet with a new participant from the opposite side who remains in the market. Then, at the end of the third period, participants decide to match with each other under the same decentralized matching protocol.

Our theoretical model, given in Section 2, does not restrict the final period matching to be centralized or decentralized. However, for the experimental design, we opted for a decentralized matching protocol in the final period. The main reason for this design choice is to improve further the elicitation method that we used for the threshold strategies of the participants. Note that after learning their own types, participants were asked to input their threshold type before the first period meeting realization. That is, we are using *strategy method* to elucidate that the equilibria that are actually played are of the form of threshold strategies. Therefore, to input their threshold types, participants should understand the implications of their own type and early matching decisions of others on the stable match in the final period. For example, for a worker i of type $w_i \in [1, 10]$ to determine its own threshold type, the worker must have a good estimate of its expected stable match partner's type in the second period,

$$\mathbb{E}_{s_{-i}}[\mu(w_i)|w_i, f], \tag{8}$$

for each type $f \in [1, 10]$ of the firm that the worker is meeting in the first period. To that end, we designed a simple and intuitive decentralized matching protocol in the final period so that each participant understands as much as possible the implications of stability in the final period matching.

We now describe the rules of the decentralized matching protocol that we implemented in periods $t > 1$ of markets with *perfect recall*. In each round, participants who did not match early start off unmatched in the decentralized matching protocol. Participants observe the types of their potential match partners and their current availabilities. Note that this is sufficient information about the match payoffs. Each participant is free to make at most one match proposal to any of his/her potential match partners at any given time. Proposers are allowed to cancel unanswered proposals. Also, participants are free to accept or reject match proposals they receive from other individuals. If a responder accepts a proposal, then the proposer and responder match and exit the round. That is, the matching decision is irreversible. Participants do not need to reject or accept match proposals immediately, and there isn't a time limit for withholding a proposal received from other individuals. However, there is a time limit for the entire decentralized matching period, which is set to two minutes. This design choice was made for practical reasons, two minutes' time is sufficient since no individual has more than 3 potential match partners in any treatment. In fact, there isn't any session in which the two minutes timing was binding for participants.

Markets with *no recall* were designed to mimic the matching model given by Damiano, Li and Suen (2005). We do not vary the time period in these markets. The design is exactly the same with two period *perfect recall* markets, except that participants' potential match partner in the second period is the participant that he/she met in the second period. Thus, second period matching is trivial: participants who remain in the market match with anyone they meet with in the second period.

Moreover, we also vary per period participation costs in the markets. In these markets, meeting in every period is optional and costly. Participants were given options to meet with a new participant in every period. The per period cost is fixed to 0.2 in terms of payoff points. Thus, no individual can end up earning a negative payoff since the lowest possible type is 1. When participation is costly in every period, there is an important difference between markets with perfect recall and no recall. In markets with perfect recall and participation cost, individuals can choose to participate in the first period and not in the second period. In that case, those individuals will have only one potential partner in the decentralized matching protocol, and they can still match in the second period. However, the same individual would be unmatched in markets with no recall. Note that this difference is actually implied by the definitions of the perfect recall and no recall notions, and it is in line with our theoretical model and with the model of Damiano, Li and Suen (2005). A detailed summary of the markets used in the experimental sessions with their characteristics appears in Table 1.

Perfect Recall		No Recall	
Market PR0	No early matching	Market NR	Two periods
Market PR	Two periods	Market NRC	Two periods with cost
Market PRC	Two periods with cost	–	–
Market PR3	Three periods	–	–

Table 1: Dynamic Matching Market Designs: Four different markets with perfect recall: Market with perfect recall (PR), market with perfect recall and costly participation (PRC), three period market with perfect recall (PR3), and market PR0, which is a two period dynamic matching market with perfect recall, and no early matching opportunity. And, two different markets with no recall: Market with no recall (NR) and market with no recall and costly participation (NRC).

Obtaining good estimates of threshold strategies requires running a large number of rounds for each market. For this reason, except market PR0, participants engaged in 15 real rounds in each market, each corresponding to a new market of the same sort. The number of rounds for market PR0 was set to 2[15]. Consequently, since running all the

[15] The treatment PR0 was essentially designed for participants to practice the decentralized matching protocol.

markets in one experimental session would be time consuming, we designed four different sessions. A detailed structure of each session is shown in Table 2. In particular, we designed our treatments so that *perfect recall* and *no recall* are both within subject and between subject design variables, and each individual experimental session is relatively less time consuming.

Session A	Session B	Session C	Session D
Market PR0	Market PR0	Market NR	Market PR0
Market PR	Market PR	Market NRC	Market PR
Market NR	Marker PRC		Market PR3

Table 2: Summary of Treatments

The experimental sessions were run at the Missouri Social Science Experimental Laboratory (MISSEL) at Washington University in St. Louis between August 2021 and September 2021. Each session shown in Table 2 was run twice, each consisting of 20 participants. This generated 8 experimental sessions overall and 152 participants[16]. Each session lasted approximately 60 minutes and paid an average of \$16, combined with a \$5 show-up fee. The experiment was programmed and conducted with oTree software Chen, Schonger and Wickens (2016).

4.1 Theoretical Predictions

To guide our experimental investigation, we numerically calculate a non-trivial threshold equilibrium for our baseline model, and also for the model with participation cost. These theoretical benchmarks are obtained by fixing $N = 20$ with uniform type distribution over $[1, 10]$. Since the treatments with no recall are designed based on Damiano, Li and Suen (2005), we explicitly calculate the equilibrium for each of their models. Figure 3 shows the corresponding equilibrium threshold strategies. The left panel of Figure 3 corresponds to our model, matching with perfect recall, and the right panel corresponds to Damiano, Li and Suen (2005).

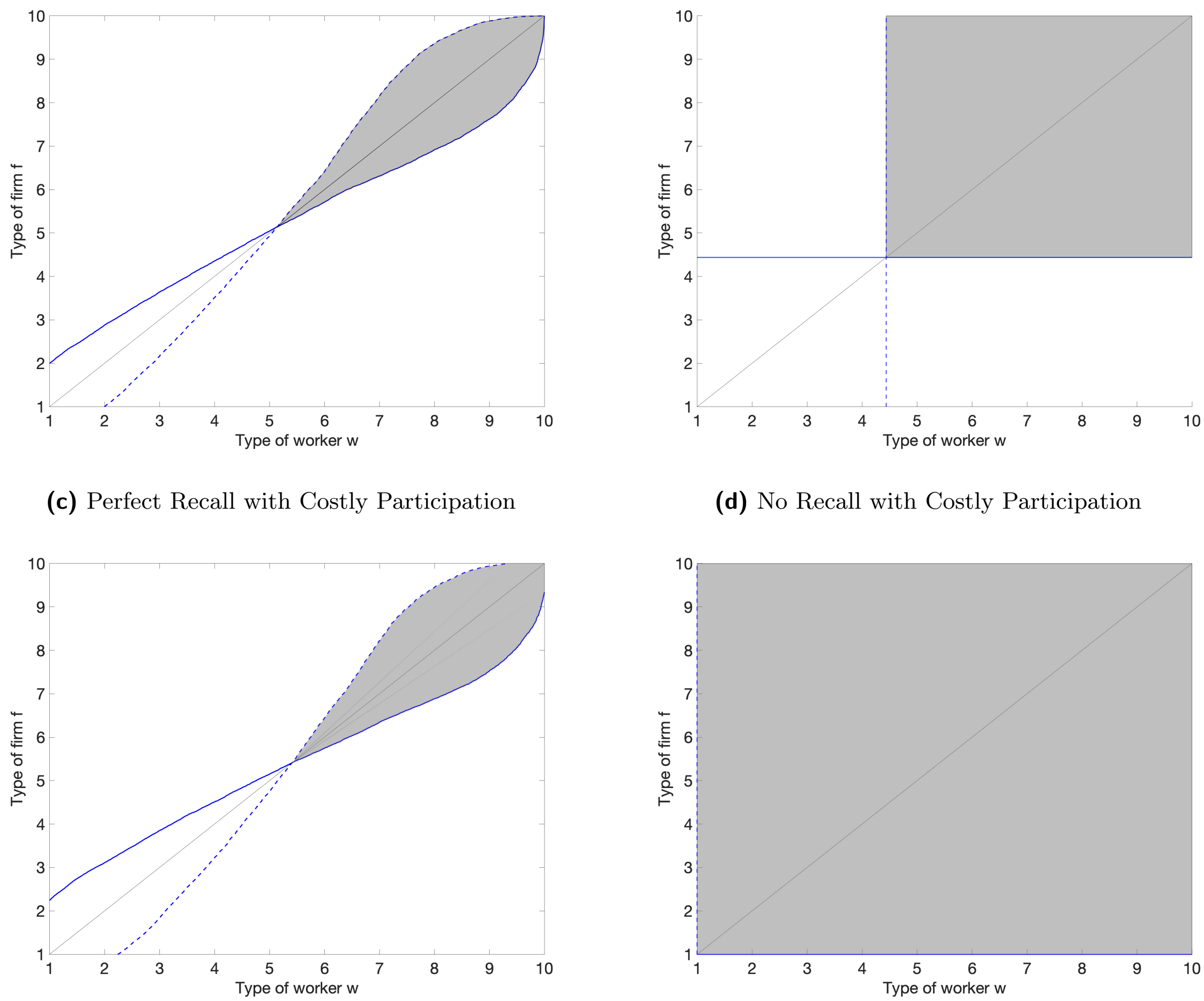
Based on the figure, we can summarize our analysis as a series of hypotheses. First, note that each shaded region in Figure 3 corresponds to the set of first period pairs that match early. Therefore, according to Figure 3(a) and (b), it is straightforward to conclude that fewer agents are expected to match early in the treatment PR compared to the treatment NR. This is because the shaded region that corresponds to treatment PR is a strict subset of the shaded region in the treatment NR.

[16]One of our sessions (Session D) was run with 12 participants.

Figure 3: Equilibrium Threshold Strategy, by treatment

(a) Perfect Recall **(b)** No Recall

(c) Perfect Recall with Costly Participation **(d)** No Recall with Costly Participation



Note: The shaded regions correspond to pairs of types that match early in the first period. A solid curve represents the equilibrium threshold strategy of a worker. A dashed curve represents the equilibrium threshold strategy of a firm.

Hypothesis 1. *More subjects match early in the treatment NR compared to the treatment PR.*

From Figure 3(a) and (b), we can directly make the following observation: Only relatively high type agents match early in both treatments PR and NR.

Hypothesis 2. *Only relatively high type agents match early in both treatments PR and NR.*

Additionally, in the treatment PR agents who match early are more similar to each other compared to the treatment NR. To see this, note that the shaded region in Figure 3(a) is around the 45-degree line, whereas the shaded region in Figure 3(b) is relatively far from the 45-degree line.

Hypothesis 3. *Early matching in the treatment PR is more assortative than in the treatment NR.*

Given a non-trivial threshold equilibrium β, note that the equilibrium probability of matching early conditional on being a type $x \in [1, 10]$ in our baseline model is given by

$$\frac{(\beta^{-1}(x) - \beta(x))\mathbb{1}_{(\beta^{-1}(x)-\beta(x)\geq 0)}(x)}{9}. \tag{9}$$

In words, fix a type of the worker $x \in [0, 1]$, and draw a vertical line at x. Then, the probability of matching early conditional on being type x is the segment of the vertical line that lies in the shaded region of Figure 3(a) divided by total length, which is 9. For the treatment NR, the probability of matching early conditional on being a type x can be calculated analogously from Figure 3(b). Thus, we can conclude that the probability of matching early for any type is always higher in treatment NR compared to the treatment PR.

Hypothesis 4. *Probability of matching early conditional on any type is higher in the treatment NR than in the treatment PR.*

Fix a type of worker $x \in [0, 1]$. Then, from Figure 3(a) we can conclude that the probability of matching early as a function of x should be first increasing then decreasing in the treatment PR. In Figure 6(a) this is depicted as an inverse U-shaped function. On the other hand, in the treatment NR, this probability is constant across all types $x \in [0, 1]$ that lie in the shaded region of Figure 3(b).

Hypothesis 5. *In the treatment NR, the probability of matching early for all types above 4.5 is predicted to be constant. In the treatment PR, the probability of matching early as a function of type is first increasing then decreasing for all types above 5.5.*

Another theoretical finding for the treatment PR is regarding the equilibrium of threshold strategies for the model with and without costly participation. Bramoullé, Rogers and Yenerdag (2021) show that the two threshold equilibria are qualitatively similar except for the highest type agent. This is confirmed by Figure 3(a) and (*c*).

Hypothesis 6. *The equilibrium threshold strategy in the treatment PR is expected to be similar to the equilibrium threshold strategy in the treatment PRC.*

In terms of unraveling, we showed that with perfect recall the market does not unravel. However, for the treatments with no recall, Damiano, Li and Suen (2005) show that the market unravels when there is a participation cost[17].

Hypothesis 7. *Markets do not unravel in the treatment PRC.*

Hypothesis 8. *Markets unravel in the treatment NRC.*

In Theorem 2, it is shown that any non-trivial threshold equilibrium strategy is strictly increasing in type. On the other hand, the threshold equilibrium in Damiano, Li and Suen (2005) is constant for all types. These results are also depicted in Figure 3.

Hypothesis 9. *The equilibrium threshold strategy in the treatment PR is strictly increasing. On the other hand, the equilibrium threshold strategy in the treatment NR is constant.*

4.2 Aggregate Outcomes of Baseline Treatments' Decentralized Matching

In this section we summarize the results from our baseline treatments at the aggregate market level. First, we discuss one of the main motivations for using the strategy method to elicit agents' threshold strategies and validate it by analyzing the matching outcomes observed in our treatments with perfect recall. One of the advantages of using the strategy method is that we are able to collect rich individual level data corresponding to subjects' threshold type as a function of their own type. Table 3 contains the number of observations collected in our baseline treatments.

Treatment	N. of Sessions	Total N. of Rounds	N. of Observations
Market PR	5	75	1500
Market PRC	2	30	600

Table 3: Treatments with Perfect Recall: Baseline treatment, market with perfect recall (PR), and market with perfect recall with costly participation (PRC).

[17] We note that Damiano, Li and Suen (2005) show that for any small participation cost, unraveling might not be the unique equilibrium. It is unique only in the limit.

In addition, we designed a simple and intuitive decentralized matching protocol in the final period conjecturing that participants would reach the stable match by their own decisions. Our results demonstrate that stable matchings are the typical outcome. Moreover, we will show that markets that end up being unstable are actually close to stable. We measure this both in terms of the number of blocking pairs and in terms of unrealized payoff gains. That is, markets on their own culminate in matchings that are stable or close to stable. Thus, we have strong evidence that subjects in our experimental sessions understand the implications of the stable match in the final period and input their threshold types accordingly. Finally, we study some tangible outcomes experienced by subjects in our baseline treatments, namely time spent under the decentralized matching protocol, number of offers made, and payoffs earned.

Table 4: Summary of decentralized matching outcomes

	Mean (se)	95% Confidence Interval
% of Stable Markets	73.3% (5%)	[59.45%, 87.20%]
% of Stable Pairs	91.14% (2.12%)	[85.25%, 93.04%]
% of Matched	86.19% (0.86%)	[83.79%, 88.58%]

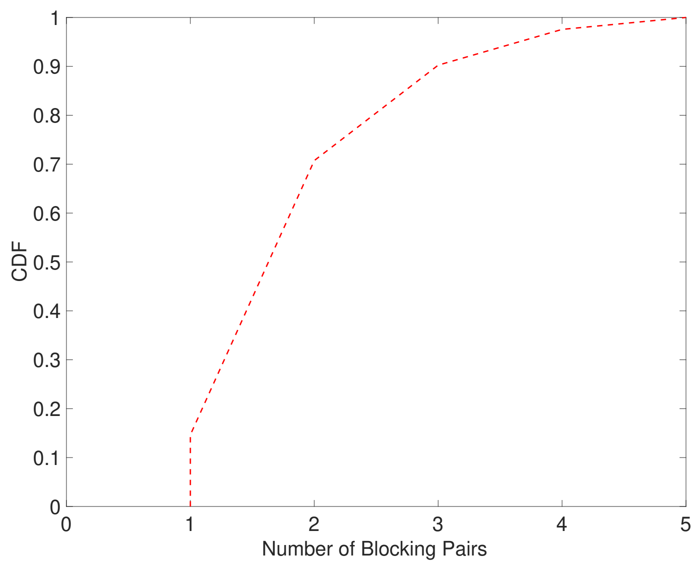
Note: Robust standard errors (reported in parentheses) are obtained by clustering at the session level to account for interdependencies of observations that come from the same session.

4.2.1 Stability and Efficiency

In our baseline treatments, subjects who did not match early were matched with each other in the final period under the decentralized matching protocol[18].Table 4 summarizes the overall outcomes of the decentralized matching in our baseline treatments. Across all markets and all rounds, 86.19% of the agents (1478 of 1716) were matched. For each matched pair, including unmatched pairs, we can check whether that pair is the *stable pair*, under the unique stable matching. According to the table, 91.14% of matched pairs, including unmatched, are stable. Note that, in our decentralized matching protocol, matching decisions are irreversible. That is, we do not allow blocking pairs to be formed in the decentralized matching protocol. Even though it is harder to reach the market-wide stable match under this design choice, 73% of all the markets were stable, which means there were no blocking pairs most of the time.

[18]Note that our baseline treatments are markets with perfect recall and markets with perfect recall and costly participation.

Figure 4: Cumulative Distribution of Blocking Pairs for Unstable Markets



Moreover, markets that were not stable were actually close to the stable matching. To see this, we check the number of blocking pairs in the unstable markets. Figure 4 depicts the empirical cumulative distribution function (CDF) of the number of blocking pairs across all the markets with unstable matching outcome. Based on the figure, most of the markets with unstable matching have only few blocking pairs. Thus, we can conclude that most of the unstable outcomes are close to stable.

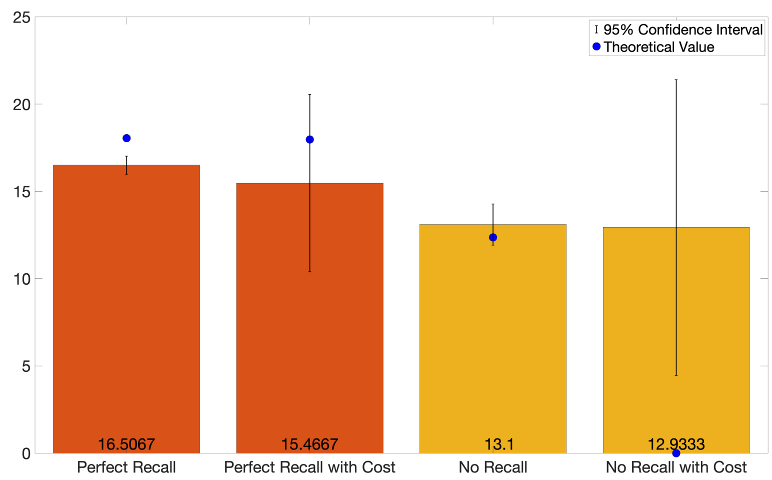
It is crucial for our experiment that agents have good estimates of their expected stable match, given in Equation (21), as we directly elicit agents' threshold strategies in the first period using the strategy method. Indeed, the aggregate decentralized matching outcomes given above constitute strong evidence that subjects in our baseline treatments were making good estimates of their expected stable match in the final period.

4.3 Market Dynamics: Matching Early vs Late

We start by looking at the aggregate frequencies of early matching outcomes across treatments. Recall that we use N_2 to denote the population size in the final period. Figure 5 presents the average N_2 across treatments. Figure 5 also shows the theoretical average N_2 corresponding to each treatment[19].

[19]For treatments with perfect recall, we obtained the average N_2 under the numerically computed equilibrium. For treatments with no recall, we simulated markets under the theoretical equilibrium to obtain the average N_2.

Figure 5: Average population size of the final period matching market, N_2, by treatment



Note: Error bars are the 95% confidence intervals based on robust standard errors obtained by clustering at the session level to account for interdependencies of observations that come from the same session. Each blue dot represents the theoretical average population size in the final period, N_2, of the corresponding treatment.

According to the figure, fewer participants matched early in perfect recall treatments compared to the no recall treatments. In line with our theoretical predictions, statistical analysis confirms that average population size remaining in the final period market in treatment PR is strictly greater than in treatment NR with p value almost equal to zero[20]. Thus, our data on the number of early matched agents is consistent with Hypothesis 1, which asserts that on average fewer people match early in the treatment PR compared to the treatment NR.

For our baseline treatments, markets with perfect recall and perfect recall with costly participation, in line with our theoretical equilibria Figure 5 shows that on average slightly fewer people matched early in treatment PR compared to the treatment PRC. However, this difference is not statistically significant due to the sample size. Hence, this result shows clear support for Hypothesis 6, according to which the equilibrium of the treatment PR should not be too different than the equilibrium of the treatment PRC. For treatments with no recall, on the other hand, there is not a significant difference between the treatments NR and NRC. This result across treatments NR and NRC sharply contrasts the theoretical result of Damiano, Li and Suen (2005) as they show that introduction of per period cost dramatically changes the matching equilibrium and results in the unraveling of markets[21].

[20] That is, we reject the null hypothesis $\mathbb{H}_0$: Average N_2 in No Recall Treatment $\geq$ Average N_2 in Perfect Recall Treatment, in favor of $\mathbb{H}_1$: Average N_2 in No Recall Treatment $<$ Average N_2 in Perfect Recall Treatment.

[21] We note that the p value for $\mathbb{H}_0 : N_2 = 0$ in treatment PRC is 0.033. In Damiano, Li and Suen (2005), unraveling is an equilibrium for any strictly positive participation cost. Moreover, unraveling equilibrium

Therefore, our data refutes Hypothesis 8[22].

Figure 5 also shows the comparison between each treatment and its theoretical benchmark. For each treatment, we cast a null hypothesis that states the average N_2 obtained from the treatment is no different than its theoretical benchmark. According to the figure, relatively more agents matched early in our baseline treatment (Market PR), 3.75 agents, compared to its theoretical prediction, 1.95 agents. Statistical analysis also reveals that we reject the null hypothesis. In treatment PRC, the theory predicts 2.03 agents match early on average and we fail to reject the null hypothesis that states the outcome in our experiment is equal to its theoretical prediction. Note that these results are consistent with Hypothesis 6 and 7 that are outlined in Section 4.1.

Additionally, the outcome in the treatment NR is also similar to its theoretical benchmark. Theory predicts that 7.64 agents match early on average, and we fail to reject the null hypothesis. On the other hand, for the treatment NRC, there is an unraveling equilibrium which means theoretical average N_2 is zero. However, in the treatment NRC only 8.07 agents matched early on average. As stated earlier, we reject the null hypothesis with p value equal to 0.033. Therefore, our data provides additional evidence against Hypothesis 8, which asserts that markets are expected to unravel in the treatment NRC.

To analyze the types of agents who match early, Figure 6 depicts the probability of matching early as a function of type across treatments, along with the corresponding theoretical values. Figure 6 suggests several insights.

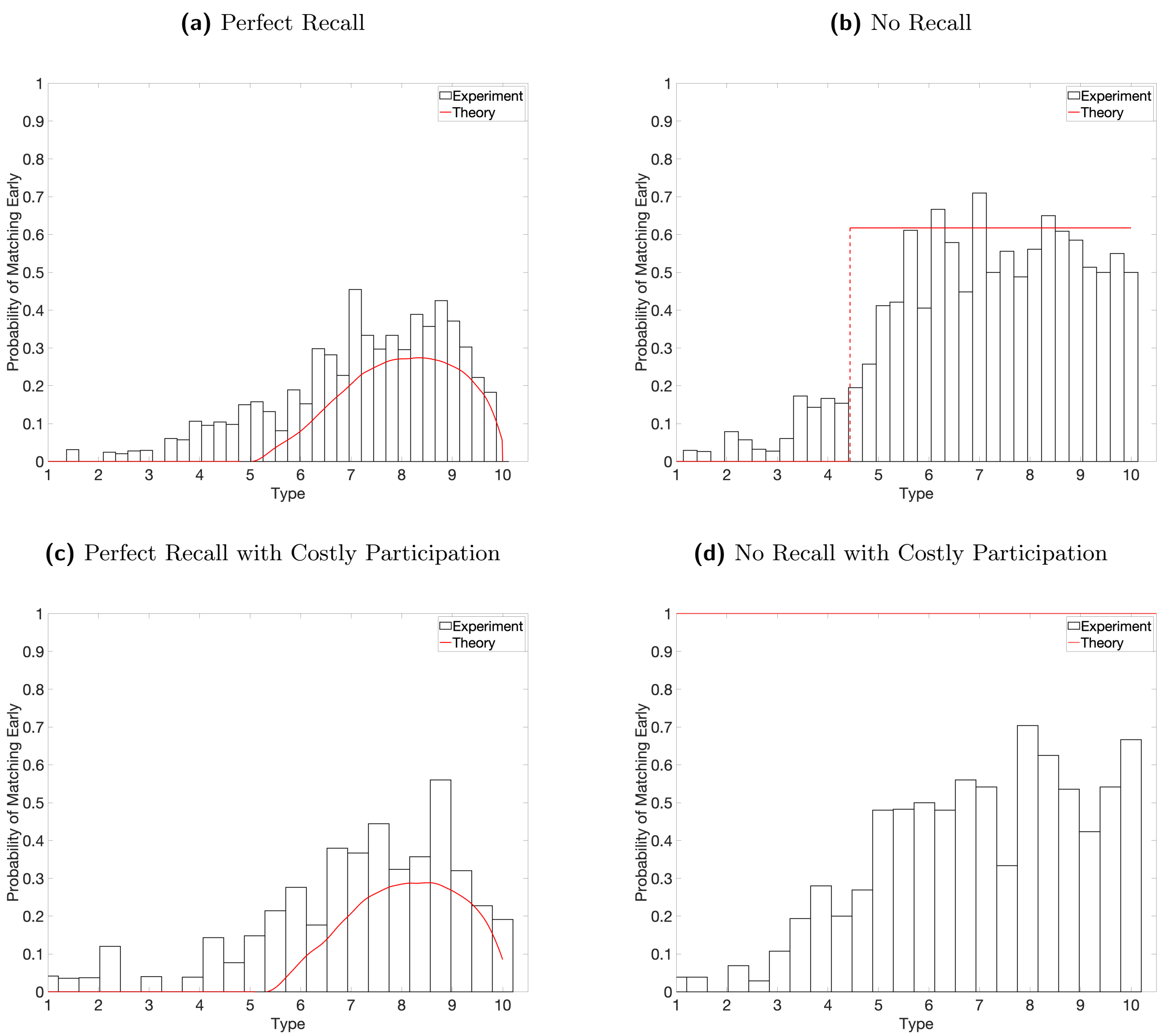
First, the probability of matching early as a function of type in each treatment, except in the treatment NRC, is qualitatively in line with its corresponding theoretical benchmark. For example, in the treatment PR, the probability of matching early is increasing in type at first, then decreasing down to zero at the highest possible type 10. Its theoretical benchmark has a very similar pattern: "the inverse U$-$ shaped curve", also shown in panel (b), which is a direct consequence of threshold equilibrium shown in Figure 3(a). Note that in the treatment NR, early matching probabilities are also in line with its theoretical benchmark. However, our data in the treatment NRC clearly deviates from the theory. While the theory predicts that every type matches early with probability 1, the data shows similarity to the treatment NR with a slightly thicker left tail. Therefore, we have a strong evidence to support Hypothesis 5. However, our data is inconsistent with Hypothesis 8.

Second, we observe a clear behavioral difference between the treatments PR and NR. According to the figure, early matching probability of each type is higher in the treatment NR compared to the treatment PR. Note that this outcome is in line with Hypothesis 4. On the other hand, when we compare treatment PR and PRC, we do not observe a

is the unique equilibrium when the participation cost is arbitrarily small.

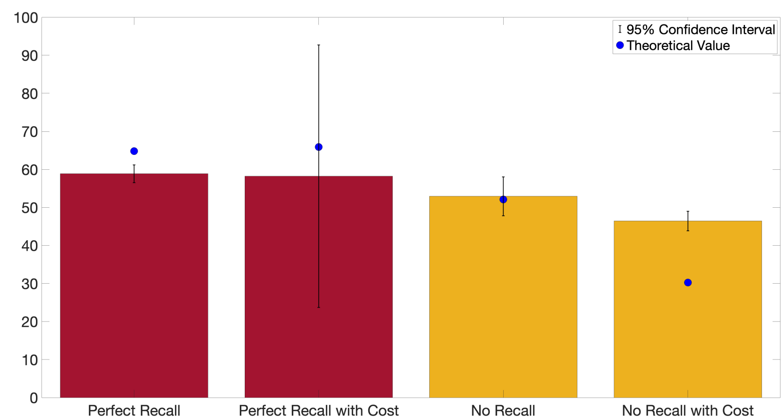
[22] We note that there could be another equilibrium different than the unraveling one when participation is costly. However, Damiano, Li and Suen (2005) do not characterize any equilibrium other than the unraveling one.

Figure 6: Probability of Matching Early as a function of type, by treatment.



significant difference in early matching probabilities. Thus, we empirically confirm that a small participation cost does not change the results of our baseline model, supporting Hypothesis 6.

Figure 7: Average utility of an early matched agent



Note: Error bars are the 95% confidence intervals based on robust standard errors obtained by clustering at the session level to account for interdependencies of observations that come from same session. Each blue dot represents average utility of an early matched agent under the numerically computed equilibrium of the corresponding treatment.

Third, since each meeting is independent, Figure 6 also shows that early matching is more assortative in treatments with recall compared to the treatments with no recall. To support this observation, Figure 7 depicts the average utility of an agent who matched early, along with 95% confidence intervals and corresponding theoretical values. Recall that utility of an agent of type f matching with type w is just the product of types, $f \times w$, and it is equal to the points earned from a match in the experiment. That is, higher average utility corresponds to more assortative matching.

The statistical analysis also confirms that average utility of an early matched agent is higher in treatment PR compared to treatment NR, with $p = 0.02$. This result suggests that early matching is more assortative in treatments with recall as outlined in Hypothesis 3. At the same time, there is not a significant difference for pairwise comparison between the treatments PR and PRC, which is in line with Hypothesis 6. On the other hand, the results from the treatment NRC show a clear deviation from the theoretical benchmark, as Damiano, Li and Suen (2005) show that an introduction of a small participation cost yields complete unraveling in markets with no recall. That is, the theory predicts complete random matching in the treatment NRC. It can be seen from Figure 6 and Figure 7 that this is not the case, and there is not a significant difference between treatments NR and NRC. Hence, our data consistently refutes Hypothesis 8.

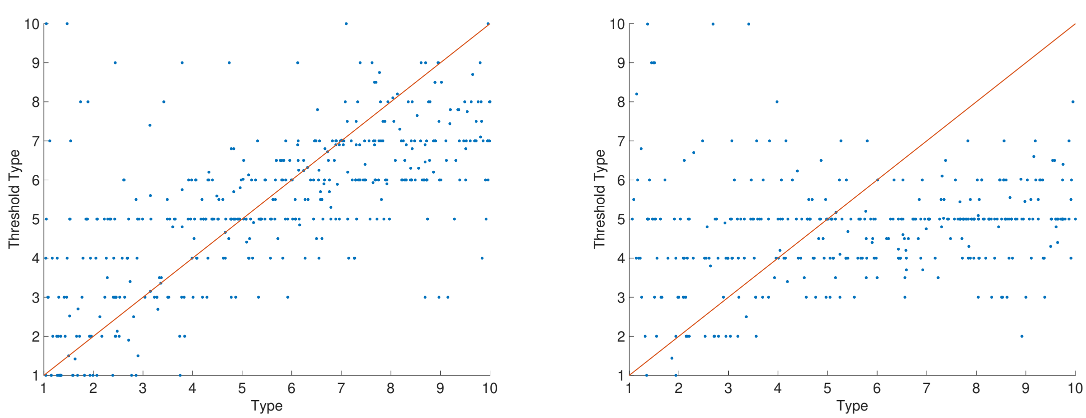
Finally, another observation from Figure 6 is that the early matching probability in each treatment systematically deviates from its theoretical benchmark. Note that, for any given type, participants matched early more often than its theoretical benchmark, except for treatment NRC. In addition, we observe that a significant fraction of agents with type less than 5 matched early in all treatments. However, theoretical models, corresponding to treatments PR, PRC and NR, predict that agents with type below 5 are matching early with probability zero[23] Hence, this result is partially inconsistent with Hypothesis 2. We suspect that this outcome is due to risk aversion. Note that we derived the equilibrium threshold strategies for all models under the assumption of risk neutrality. On the other hand, it is a general observation that agents in the lab are risk averse.

4.3.1 Threshold Strategy: Individual analysis

In this section, we mainly focus on treatments PR, and investigate individual level responses on threshold types. In our data, we observe that the threshold strategy of almost all subjects is increasing in type in the baseline treatment PR[24]. In addition, we observe a clear treatment effect between treatments PR and NR.

Figure 8: Threshold Strategies of each type, within subject comparison

(a) Perfect Recall **(b)** No Recall



Note: The solid line depicts the 45 degree line.

To analyze the effect of recall, Figure 8 depicts threshold strategies of subjects who participated in both treatments PR and NR[25]. According to the figure, the threshold

[23] Also, it can be seen in Figure 6 that for all types less than 5 the probability of matching early is decreasing as type decreases across all treatments.
[24] We note that only 3 participants out of 152 had decreasing threshold function in type.
[25] In this figure, we excluded the first 5 rounds in order to eliminate the noise in the data.

strategy is increasing in type in treatment PR, and almost constant in treatment NR, which shows a clear support for Hypothesis 9.

Table 5: Fixed Effect Regression
Dependent Variable: Threshold Type

	PR	**NR**
type	0.625***	0.075**
	(0.21)	(0.0365)
$type^2$	0.015	
	(0.41)	
$type^3$	-0.002	
	(0.002)	
cons	2.481***	3.853***
	(0.345)	(0.197)
N	1395	1200
R^2	0.7338	0.3275

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Note: Standard errors are given in parentheses below the estimates, and they are obtained by clustering at the individual level to account for interdependencies of observations that come from same individual.

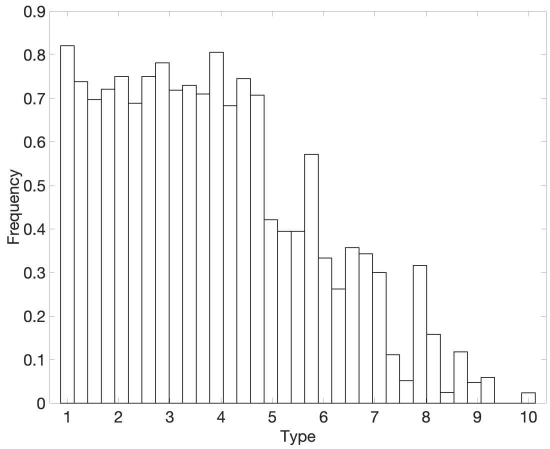
Table 5 reports the impacts of type on agents' threshold strategies in the treatments with recall and no recall. According to the regression analysis, agents' threshold strategies are strictly increasing in type in the perfect recall treatments. Note that when there is no recall, the equilibrium threshold strategy is independent of type. However, Table 5 reports a small but statistically significant positive association between types and threshold strategy in the treatments with no recall. Even though this is contrary to the theoretical benchmark, Table 5 shows that there is a clear treatment effect of recall since the positive association between type and threshold strategy is much stronger in the treatments with recall compared to the treatments with no recall.

Another distinct feature of the non-trivial threshold equilibrium in our model is that for low types, the equilibrium threshold is above the 45 degree line, and for high types it is below the 45 degree line[26]. Based on this, Figure 9 depicts the frequency of subjects reporting thresholds below and above the 45 degree line as a function of type, across all sessions with treatment PR.

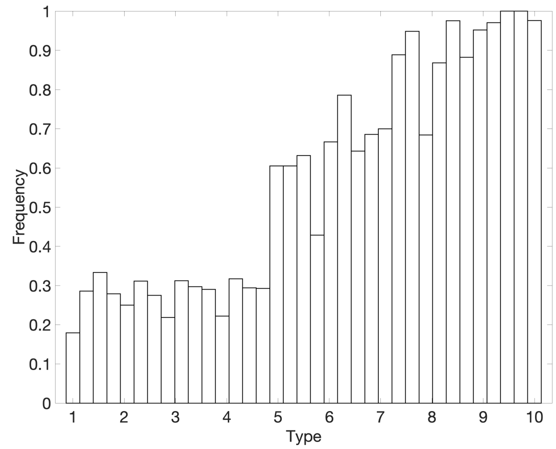
Figure 9 suggests several insights. First, the majority of subjects reported threshold types in line with the theoretical benchmark in the sense that low types had a threshold

[26]See, for instance, Figure 3(a)

above the 45 degree line, and for high types it was below the 45 degree line. However, a significant frequency of subjects behaved in the other direction. Around 25% of the subjects with low type reported a threshold type above the 45 degree line, and around 15% percent of the subjects with high type reported a threshold below the 45 degree line.



(c) Above



(d) Below

Figure 9: Perfect Recall: Above vs Below.

This result suggests that around 20 percent of all the subjects behaved with extreme risk aversion and wanted to match early with their first period partner.

5 Conclusion

In this paper, we theoretically and experimentally study agents' early matching incentives in a dynamic matching environment. We find several important results. Our first result is on unraveling. In equilibrium, some agents match early, however markets do not completely unravel. In our model, unraveling is also an equilibrium but it derives simply from a coordination failure similar to the empty graph problem in network formation models. Hence, in this sense, unraveling is a rare event. Second, we find that agents who are matching early are good-but-not-best type agents. That is, low type agents do not match early, and, except the highest type, relatively high type agents match early with positive probability. In particular, the probability of matching early as a function of type is non-monotonic in equilibrium. Moreover, early matching is highly assortative. These results are important from a market design perspective. When too many agents match early in a random fashion, markets can face a significant loss of efficiency due to resulting market-wide mismatches.

We designed a decentralized matching experiment with real time interaction to provide an extensive analysis of agents' early matching incentives and to test our theoretical predictions. According to our data, decentralized markets predominantly culminate on their own in a stable matching. By collecting rich individual level data, we are able to confirm most of our theoretical predictions. In the experiments, agents' threshold strategy is increasing, and induces highly assortative early matchings. Overall, our data shows that when the markets are able to produce a stable matching, even in an inefficient manner, early matchings do not generate market-wide mismatches, and do not lead to unraveling.

Appendix A Omitted Proofs

Proof of Theorem 1

Before turning to the proof, we briefly discuss the steps of the proof of Theorem 1. To that end, the following notations will be useful. We let $\tilde{\beta}$ be a mapping on $\mathcal{Y}$ defined as

$$\tilde{\beta}[f](x) \equiv \min B_f(x) \quad \forall x \in [a, b], \tag{10}$$

where $B_f(x) = \{y \in [a, b] : \mathbb{E}_f[\mu(x)|x, y] \leq y\}$[27].We call the mapping $\tilde{\beta}$ best response mapping.

The first step of the proof of Theorem 1 is to prove the following lemma:

Lemma 1. *Best response mapping $\tilde{\beta}$ is well defined, and $\tilde{\beta}[f] \in \mathcal{Y}$ for all $f \in \mathcal{Y}$.*

Lemma 1 says if f is a weakly increasing threshold function, then the best response $\tilde{\beta}[f]$ is a weakly increasing threshold function as well.

Next, for any fixed $\epsilon > a$, we let $\mathcal{Y}_\epsilon$ be a subset of $\mathcal{Y}$ defined as

$$\mathcal{Y}_\epsilon = \{f \in \mathcal{Y} | f(x) \geq \epsilon \quad \forall x \in [a, b]\}.$$

The second step of the proof of Theorem 1 is the following key lemma:

Lemma 2. *There exists an $\epsilon > a$ and $\bar{N} \in \mathbb{N}$ such that for all market size $N \geq \bar{N}$ the best response mapping $\tilde{\beta}$ is a self-map on $\mathcal{Y}_\epsilon$. That is,*

$$\tilde{\beta} : \mathcal{Y}_\epsilon \to \mathcal{Y}_\epsilon.$$

After establishing Lemma 2, we prove the following lemmas:

Lemma 3. *For all $\epsilon \geq a$, $\mathcal{Y}_\epsilon$ is a convex and compact subset of Banach space $L^1[a, b]$.*

Lemma 4. *The best response mapping $\tilde{\beta} : \mathcal{Y}_\epsilon \to \mathcal{Y}_\epsilon$ is continuous in L^1.*

[27] y denotes a type of the first period partner.

Our existence theorem is now obvious from the above lemmas. We restate Theorem 1 and give its proof:

Theorem 1. *There exists an $\bar{N} \in \mathbb{N}$ such that every market size of $N \geq \bar{N}$ has a strictly increasing and continuous non-trivial symmetric threshold equilibrium.*

Proof. By Lemma 1 and Lemma 2, there exists an $\epsilon > a$ and $\bar{N} \in \mathbb{N}$ such that for all market size $N \geq \bar{N}$ such that

$$\tilde{\beta} : \mathcal{Y}_\epsilon \to \mathcal{Y}_\epsilon.$$

Note that non-emptiness of $\mathcal{Y}_\epsilon$ is obvious. By Lemma 3, it is also convex and compact. Moreover, $\tilde{\beta}$ is continuous by Lemma 4. Therefore, by Schauder fixed-point theorem $\tilde{\beta}$ admits a fixed point in $\mathcal{Y}_\epsilon$. Note that, any such fixed point satisfies our equilibrium definition given in Definition 3. And, it is a non trivial threshold equilibrium by construction.

Next, for all $s \in \mathcal{Y}$, Lemma 1 proves that $\mathbb{E}_s[\mu(w_1)|w_1, f_1]$. it is weakly increasing in w_1. Since the ex-ante type distribution is continuous and the fixed point is in $\mathcal{Y}_\epsilon$ for some $\epsilon > 0$, then it implies that $\mathbb{E}_s[\mu(w_1)|w_1, f_1]$ is strictly increasing. Also note that $\mathbb{E}_s[\mu(w_1)|w_1, f_1]$ is continuous in both arguments, since w_1, f_1 only enters as a limit and integrand of the corresponding well defined integrals. Hence, the fixed point must be strictly increasing and continuous by Equation 10. ■

The proofs of Lemma 3, and Lemma 4 directly follows from Bramoullé, Rogers and Yenerdag (2021). We provide the proofs of Lemma 1 and Lemma 2 when agents have at most $k \in [1, \frac{N}{2} - 2]$ meetings in the second period. That is, each agent meets with $k + 1$ agents over two periods whenever it is possible.

Proof of Lemma 1

Proof. First, note that for all $s \in \mathcal{Y}$ and $w_1 \in [a, b]$, the function given in Equation 11 is a continuous function of f_1 defined on $[a, b]$.

$$\mathbb{E}_s[\mu(w_1)|w_1, f_1] - f_1. \tag{11}$$

To see this, recall that if s is almost everywhere equal to a, then $\mathbb{E}_s[\mu(w_1)|w_1, f_1] = w_1$. Hence, Equation 11 would be equal to 0 for all $w_1 \in [a, b]$. And, as long as s is strictly greater than 0 for a positive measure set, then the properties of h, the pdf of ex-ante type distribution, are sufficient to conclude that Equation 11 is a continuous function of w_1, since w_1 only enters as a limit and integrand of corresponding well defined integrals. Next, Bramoullé, Rogers and Yenerdag (2021) proves the following:

$$B_s(x) = \{w_1 \in [a, b] : \mathbb{E}_f[\mu(x)|x, f_1] \leq f_1\}, \tag{12}$$

is a closed upper set of $[a, b]$ for all $x \in [a, b]$ when $k = 1$. This result can be extended trivially to any k meetings in the second period as long as $k + 1 \leq N/2 - 1$. Hence, best response mapping $\tilde{\beta}$ is well defined.

To prove the second part of the lemma, take any weakly increasing threshold function s. Take any agent on one side with type x who met with type y from other side in period $t = 1$. By definition of $\tilde{\beta}$ given in Equation 10, it is sufficient to show that $\mathbb{E}_f[\mu(x)|x, y]$ is weakly increasing as a function of $x \in [a, b]$. To see this, first note that all agents strictly prefer higher types. Then, for any realizations of all other agents' types and meetings, increase in type x always improves the agent's match in period $t = 2$, since μ is a stable matching. Therefore, $\mathbb{E}_f[\mu(x)|x, y]$ must be weakly increasing as a function of $x \in [a, b]$. Hence, this completes the proof. ■

Proof of Lemma 2

Proof. First, note that by Assumption 1(ii), we can normalize the type space to $[0, 1]$. Bramoullé, Rogers and Yenerdag (2021) proves the lemma for $k = 1$, that is, one meeting per period. So, we need to prove that the lemma holds for any $k \in [1, \frac{N}{2} - 2]$.

We introduce additional useful notations for the proof. We use capital letters W_j and F_j to denote random type of worker j and firm j, respectively. Recall that we say that (w, f) is a "period t pair" if worker w and firm f meets in period $t \in \{1, 2\}$. Without loss of generality, we relabel all agents such that (w_j, f_j) is a period 1 pair for all $j = 1, 2, \ldots, \frac{N}{2}$.

First note that $k \in [1, \frac{N}{2} - 2]$ implies that workers cannot meet with all firms, and vice versa. Since meetings are random and independent, it implies that there exists a strictly positive probability for all agent to be unmatched in the second period. To see this, first suppose that the highest worker meets with the highest firm, and the second highest worker meets with the second highest firm, and so on so forth, up to kth highest worker and firm. Now, fix j and j' such that $k + j \leq N_2/2 - 1$ and $k + j' \leq N_2/2 - 1$. Next, assume that the $k + j$th highest worker does not meet with any firm whose rank is lower than k, and the $k + j'$th highest firm does not meet with any worker whose rank is lower than k. In this case, under the stable match, the first highest k workers match with the first highest firm assortatively. In addition, the highest $k + j$th worker and the highest $k + j'$th highest firm are unmatched under the stable matching since they did not meet with anyone else and cannot form a blocking pair. Thus, since $N < \infty$ and meetings are random, this can happen with strictly positive probability.

Now, fix a worker w_1 who meets f_1 in the first period. Suppose that all other agents except f_1 follows a weakly increasing threshold strategy s. Consider the following

$$\mathbb{E}_s[\mu(w_1)|w_1 = 0, f_1 = 0]. \tag{13}$$

Following the arguments in Bramoullé, Rogers and Yenerdag (2021), it is sufficient to prove that there exists an $\epsilon > 0$ such that

$$\mathbb{E}_s[\mu(w_1)|w_1 = 0, f_1 = 0] > \epsilon \quad \forall s \geq \mathcal{Y}_\epsilon. \tag{14}$$

To that end, let $(W_j, F_j)|s$ denote the joint distribution of a first period pair who do not match early in the first period. Also, let $F^L(w_i)$ denote the set of all firms that worker i meets over two periods. At the end of the first period, suppose that $F(w_1) = \{f_1, F_2|s, F_3|s, \ldots, F_{k+1}|s\}$, where $F_j|s$ denotes the random variable with a distribution function equals to the marginal of $(W_j, F_j)|s$. Now, from the argument above, for all non trivial s, there is a strictly positive probability of a firm F that w_1 meets in the second period and all other workers that F meets are matched with some other firm. We let R^* denote the set of all rankings of all agents' types in the second period conditional on $w_1 = f_1 = 0$ such that for all ranking $R \in R^*$ induces a stable matching in which w_1 is matched with F. Take any $R \in R^*$. Then, we can write the following inequality

$$\mathbb{E}_s[\mu(w_1)|w_1 = 0, f_1 = 0] > \mathbb{P}_s(N_2 \geq k)\mathbb{P}_s(R)\mathbb{E}_s[F|R]. \tag{15}$$

Now, note that since the ex-ante type distribution is continuous, we can conclude the following the facts:

i. Fix any $c \in (0, 1]$. There exists a $s^* \in \mathcal{Y}_c$ such that $\mathbb{P}_{s^*}(R) = \inf_{s \in \mathcal{Y}_c} \mathbb{P}_s(R)$

ii. Fix any $c \in (0, 1]$. Then,

$$\lim_{c \to 0} \mathbb{E}_{s^c}[F|R] = m > 0$$

iii. Fix any $c \in (0, 1]$.

$$\mathbb{E}_{s^c}[F|R] \leq \mathbb{E}_s[F|R] \quad \forall s \in \mathcal{Y}_c.$$

First, note that (*ii*) is trivial because the endogenous marginal type distribution is continuous. For (*i*), note that under each $s \in \mathcal{Y}_c$, R is just a ranking among a finite number of continuous random variables. Lastly, (*iii*) is implied by FOSD. Note that the marginal type distribution under s FOSD the marginal type distribution under s^c for all $s \geq s^c$[28].

Therefore, by setting $\epsilon = \mathbb{P}_{s^*}(R)m$, we obtain

$$\mathbb{E}_s[\mu(w_1)|w_1 = 0, f_1 = 0] > \mathbb{P}_m(N_2 \geq k)\epsilon \quad \forall s \in \mathcal{Y}_\epsilon. \tag{16}$$

Now, note that $\mathbb{P}_m(N_2 \geq k)$ is approaching to 1 as N grows. Hence, we can find $\bar{N} < \infty$

[28] A detailed proof can be found in Bramoullé, Rogers and Yenerdag (2021)

such that for all $N \geq \bar{N}$ we can conclude

$$\mathbb{E}_s[\mu(w_1)|w_1 = 0, f_1 = 0] > \epsilon \quad \forall s \in \mathcal{Y}_\epsilon, \tag{17}$$

which was to be shown. ■

Theorem 2. *Suppose that β is a non-trivial threshold equilibrium. Then, β satisfies the following conditions:*

(i) $\beta(a) > a$ *and* $\beta(b) = b$

(ii) *There exists an $x^* \in (a, b)$ with $\beta(x^*) = x^*$ such that $\beta(x) < x$ for every $x \in (x^*, b)$.*

Proof. (*i*) Take any non-trivial threshold equilibrium β. Note that $\beta(a) > a$ immediately follows from the fact that $\beta \in \mathcal{Y}_\epsilon$ for some $\epsilon > a$. For $\beta(b) = b$, take a worker $w_1 = b$ who meets with a firm f_1 in the first period. Recall that all agents strictly prefer to be matched with higher type. Thus, stability of μ implies that

$$\mathbb{E}_\beta[\mu(b)|b, f_1] = \mathbb{E}_\beta[\max\{F^L(1)\}] \quad \forall f_1 \in [a, b], \tag{18}$$

where $F^L(1) = \{f_1, F_2|\beta, F_3|\beta, \ldots, F_k|\beta\}$ is the set of all firms that $w_1 = b$ meets over two periods.

Moreover, since $F_j|\beta$ admits a strictly positive density over $[a, b]$, we have

$$\mathbb{E}_\beta[\mu(b)|b, f_1] > f_1 \quad \forall f_1 \in [a, b), \tag{19}$$

$$\mathbb{E}_\beta[\mu(b)|b, f_1] = f_1 \quad f_1 = b, \tag{20}$$

Since β is a fixed point of $\tilde{\beta}$, then $\beta(b) = \min\{f_1 \in [a, b]|\mathbb{E}_\beta[\mu(b)|b, f_1] \leq f_1\}$. Hence, by Equation (19) and Equation (20) we conclude that $\beta(b) = b$. □

(*ii*) Fix a worker $w_1 \in [a, b)$ who meets with a firm f_1 in period 1. From w_1's point of view, $\mu(m_1)$ is a random variable. Given a threshold function β, w_1, f_1 and market size N; $\mu(m_1)$ can yield three different outcomes: Matching with first period partner, matching with a second period partner, and being unmatched. Let, $\mathcal{E}_1(\beta, N, w_1, f_1)$, $\mathcal{E}_2(\beta, N, w_1, f_1)$, $\mathcal{E}_\emptyset(\beta, N, w_1, f_1)$ denote the events under which w_1 is matched with first period partner, second period partner and being unmatched, respectively. Therefore, we have

(a) $[\mu(w_1)|\mathcal{E}_1(\beta, N, w_1, f_1)] = f_1$

(b) $[\mu(w_1)|\mathcal{E}_2(\beta, N, w_1, f_1)] = [F|\mathcal{E}_2(\beta, N, w_1, f_1)]$

(c) $[\mu(w_1)|\mathcal{E}_\emptyset(\beta, N, w_1, f_1)] = a$.

Next, let $p_1(\beta, N, w_1, f_1)$, $p_2(\beta, N, w_1, f_1)$ and $p_\emptyset(\beta, N, w_1, f_1)$ denote the probabilities of the corresponding events. Hence, by the law of total probabilities:

$$\mathbb{E}_\beta[\mu(w_1)|w_1, f_1] = p_1(\beta, N, w_1, f_1)f_1 + p_2(\beta, N, w_1, f_1)\mathbb{E}[F|\beta, \mathcal{E}_2(\beta, N, w_1, f_1)] + p_\emptyset(\beta, N, w_1, f_1)a. \tag{21}$$

Now, note that since $\beta \in \mathcal{Y}_\epsilon$, $F|\beta$ admits strictly positive probability density function. Then, it is easy to verify that $\mathbb{E}[F|\beta, \mu(w_1) = F, w_1, f_1]$ is a continuous function of f_1 and w_1. Also, we have $\mathbb{E}[F|\beta, \mathcal{E}_2(\beta, N, w_1, f_1)] \in (a, b)$ for all $w_1, f_1 \in [a, b]$. Therefore,

$$\max_{w_1, f_1 \in [a,b]} \mathbb{E}[F|\mathcal{E}_2(\beta, N, w_1, f_1)] = y < b. \tag{22}$$

Hence, for all $x \in [y, b)$ we can conclude:

$$\mathbb{E}_\beta[\mu(x)|x, x] \leq p_1(\beta, N, x, x)x + p_2(\beta, N, x, x)y + p_\emptyset(\beta, N, x, x)a < x, \tag{23}$$

since we have $p_1(\beta, N, x, x) > 0$, $p_2(\beta, N, x, x) > 0$ and $p_\emptyset(\beta, N, x, x) > 0$ for all $x \in [y, b)$ due to $k \in [1, \frac{N}{2} - 2]$. Now, since β is an equilibrium threshold strategy (fixed point of $\tilde{\beta}$), then Equation (23) implies that $\beta(x) < x$ for all $x \in [y, b)$.

Finally, by (i) we have $\beta(a) > a$ and β is continuous and strictly increasing. Therefore, by intermediate value theorem, there exists an $x^* < y$ such that $\beta(x^*) = x^*$ and $\beta(x) < x$ for all $x \in (x^*, b)$. This completes the proof.

■

Corollary 1. *Under any non-trivial threshold equilibrium, some agents match early and the probability of matching early conditional on type is non-monotonic.*

Proof. The first part of the corollary is proven in the text. We provide the proof of the second part. To that end, without loss of generality, fix a type space $[0, 1]$. Take any non-trivial threshold equilibrium β. Then, the probability of matching early for an agent of type $x \in [0, 1]$ is $(\beta^{-1}(x) - \beta(x))\mathbb{1}_{(\beta^{-1}(x)-\beta(x)\geq 0)}(x)$. To conclude the proof, we need to show that there exists $x, x', y, y' \in [0, 1]$ with $x > x'$, $y > y'$ such that

$$(\beta^{-1}(x) - \beta(x))\mathbb{1}_{(\beta^{-1}(x)-\beta(x)\geq 0)}(x) < (\beta^{-1}(x') - \beta(x'))\mathbb{1}_{(\beta^{-1}(x')-\beta(x')\geq 0)}(x'), \tag{24}$$

$$(\beta^{-1}(y) - \beta(y))\mathbb{1}_{(\beta^{-1}(y)-\beta(x)\geq 0)}(y) > (\beta^{-1}(y') - \beta(y'))\mathbb{1}_{(\beta^{-1}(y')-\beta(y')\geq 0)}(y'). \tag{25}$$

Note that β is strictly increasing, continuous and $\beta(1) = 1$. In addition, there exists $x^* \in (0, 1)$ such that β crosses the 45-degree line from above at x^*, and $\beta(x) < x$ for all

$x \in (x*, 1)$. Hence, these are sufficient to conclude that there exists $x, x', y, y' \in [0, 1]$ such that Equation (24) and Equation (25) hold. ■

Proposition 1. *There exists a non-trivial threshold equilibrium β such that no firm makes an exploding offer.*

Proof. Consider the model described in Section 2. Take any non-trivial threshold equilibrium β without exploding offers. Now, suppose that firms have an option to make exploding offers with full commitment. To prove the proposition, it is sufficient to show that no agent has incentive to deviate from β.

To that end, except firm f, suppose that no firm makes an exploding offer and every agent follows the threshold strategy β. Without loss of generality, assume that f meets worker w_1 in the first period. Let $W_f = \{w_1, w_2, \ldots, w_k\}$ a list of workers that f meets over two periods. There are 3 cases to consider:

Case 1: If $\beta(w_1) \leq f$ and $\beta(f) \leq w_1$. Since β is an equilibrium threshold strategy, then when f does not make an exploding offer we have

$$\mathbb{E}_\beta[\mu(f)|f, w_1] \leq w_1. \tag{26}$$

Note that in this case, f can match early with w_1. Hence, we need to show that f cannot improve in the second period matching by making an exploding offer to w_1.

Suppose that f makes an exploding offer to w_1 in the first period. This equivalently means that the firm f listing worker w_1 as unacceptable in his/her rank order list in the second period. However, we have $w_1 \geq a$, and by Assumption 1 utility of being unmatched is equal to utility of matching with the lowest type. Thus, w_1 is acceptable. And, listing w_1 as not acceptable is equivalent to submit a rank order list equal to $\{a\} \cup (W_f \setminus w_1)$. Let $[\mu(f)|W_f \setminus w_1]$ denote a matching realization for f in the second period with a rank order list W_f. Therefore, since second period matching is unique with probability 1, every matching realization satisfies

$$[\mu(f)|a, W_f \setminus w_1] \leq [\mu(f)|W_f]. \tag{27}$$

This is because when there exists a unique stable matching it is always optimal to submit the true rank order list. Note that Equation (27) implies

$$\mathbb{E}_\beta[\mu(f)|f, a] < \mathbb{E}_\beta[\mu(f)|f, w_1]. \tag{28}$$

Hence, firm f does not have incentive to make an exploding offer since by doing that f cannot improve the second period matching.

Case 2: If $\beta(w_1) > f$. In this case, we have

$$\mathbb{E}_{\beta}[\mu(f)|f, w_1] > w_1.$$

Hence, firm f is not willing to match with w_1 in any case. Therefore, there isn't incentive for f to make an exploding offer to w_1

Case 3: If $\beta(w_1) > f$ and $\beta(f) \leq w_1$. This case is similar to Case 1. If f makes an exploding offer to w_1, we obtain the inequality given Equation (28). Hence, firm f does not have incentive to make an exploding offer.

■

Appendix B Experiment: Sample Instructions

We provide the sample instructions of Session A shown in Table 2.

Welcome

Welcome to MISSEL and thank you for participating in today's experiment. Remember that your participation is voluntary. You are free to leave this experiment at any time.

Guidelines

You will be paid in private and in cash at the end of the experiment. You have each earned a \$5 payment for showing up. The final amount that you will earn in the experiment depends on your decisions, the decisions of others, and random chance. Please use the PCs as instructed and do not attempt to browse the web, and please do not socialize and talk so that we can have your complete attention.

Overview

This experiment consists of 3 different "blocks." In each block, you'll be making bunch of decisions. Block 1 consists of 2 rounds, and Block 2 and 3 consist of 15 rounds each. Your decisions and the outcome in each round will not affect your decisions and outcomes in the other rounds and other blocks. The rounds within a block are completely independent. Moreover, three blocks are also completely independent. We'll go over instructions at the start of each block.

At the beginning of the experiment, you will have a randomly assigned ID number. Also, you will be randomly assigned to a role: either a "circle" or a"square". Total number of participants in this experiment is 20. There will be 10 circles and 10 squares. **Your role will remain fixed** across all rounds and all blocks of the experiment.

At the beginning of each round, you will be **privately** and **randomly** assigned a **productivity level**, which can be any number between 1 and 10. For example: 2, 3.14, 7.44 etc. Productivity level of each participant is **independent** from each other.

Each round of every block consists of two main stages. **In each stage**, you will be **randomly** paired with a participant from the opposite role. We call these pairings "meetings". In each stage, you will meet a new participant from the opposite role. For example, if you are a Circle, you will meet a Square. Random meetings are completely independent from productivity levels. Regardless of your productivity, **meeting anyone is with equal probability.**

Paired participants will observe the productivity of each other. Your task is to try to match with a participant from the opposite role or to remain unmatched. You **CANNOT** match with a participant that you did not meet. So, each participant that you meet is your **potential match partner**.

Match Payoff

Suppose that you are a Square and your productivity is 3.2. If you are matched with a Circle with productivity of 2.5, then your payoff is $3.2 \times 2.5 = 8$ points. If you are unmatched, then your payoff is equal to your productivity. In this case, you will earn 3.2 points from being unmatched.

Note that regardless of your productivity level, being unmatched yields the lowest point. Regardless of your productivity level, matching with participant with higher productivity yields higher points.

Total Points of a Block

At the end of all rounds in a block, we will choose one round at random. And, the randomly chosen round will be your total points of the block.

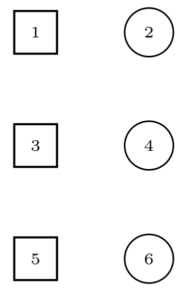
Total points of the Experiment

The sum of all points from each block will be your total points earned in the experiment. We will convert your points to a dollar amount. Higher points yield higher cash amount.

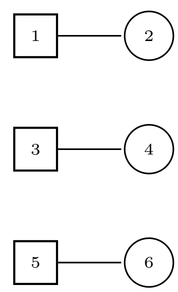
Instructions of Block 1

This block consists of 2 rounds.

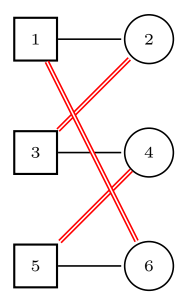
- Round n: 3 circles and 3 squares. Productivity levels are privately and independently drawn. Any number between 1 and 10 equally likely.



- **First Stage:** Random Pairing (Meeting)



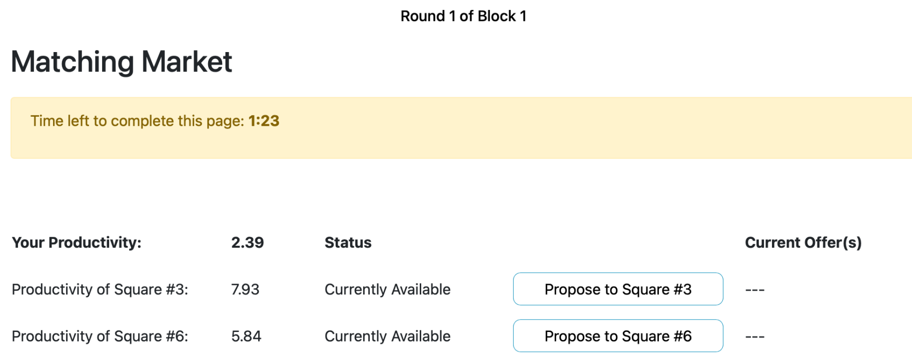
- **Second Stage** Random Pairing (Meeting)

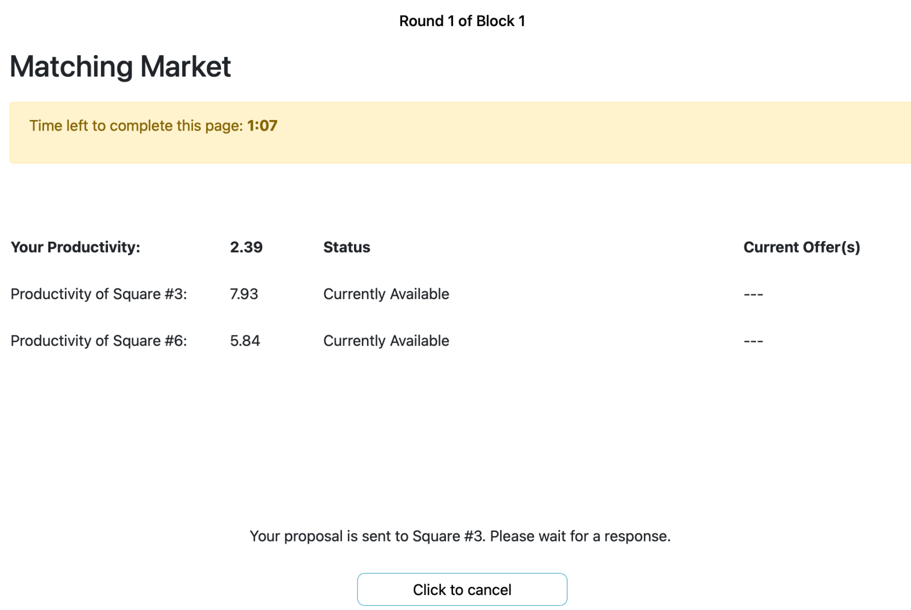


- After the two stages, there is a decentralized matching market.

Rules of the Decentralized Matching Market

You will start off unmatched. You have two potential match partners from stages 1 and 2. You have 120 seconds in total to arrange a match. You are free to make at most one match proposal to any of your potential match partners at any given time. You can cancel your match proposal if it is unresponded. If you receive a match proposal, you can accept, reject or withhold. As long as you are not matched, you can receive match proposals. If you accept a match proposal, or your proposal is accepted, then you will get the match points and exit the round. If all of your potential match partners are matched with someone else, then you are unmatched and receive points equal to your productivity level.





Round 1 of Block 1

Matching Market

Time left to complete this page: **0:57**

Your Productivity:	**7.93**	**Status**		**Current Offer(s)**		
Productivity of Circle #1:	2.39	Currently Available		Circle 1 is willing to match	Click to accept Circle #1	Click to reject Circle #1
Productivity of Circle #2:	3.75	Currently Available	Propose to Circle #2	---		

Round 1 of Block 1

Matching Market

Time left to complete this page: **0:41**

Your Productivity:	**2.39**	**Status**	**Current Offer(s)**		
Productivity of Square #3:	7.93	Currently Available	---		
Productivity of Square #6:	5.84	Currently Available	Square 6 is willing to match	Click to accept Square #6	Click to reject Square #6

Your proposal is sent to Square #3. Please wait for a response.

Click to cancel

Round 1 of Block 1

Matching Market

Time left to complete this page: **0:16**

Your Productivity:	**2.39**	**Status**	**Current Offer(s)**		
Productivity of Square #3:	7.93	Not Available. Matched with someone else.	---		
Productivity of Square #6:	5.84	Currently Available	Square 6 is willing to match	Click to accept Square #6	Click to reject Square #6

Round 1 of Block 1

Matching Market

Time left to complete this page: **0:02**

Your Productivity:	**2.39**	**Status**	**Current Offer(s)**
Productivity of Square #3:	7.93	----	---
Productivity of Square #6:	5.84	Matched with you.	---

You are matched with Square #6.

Your payoff is 13.96 points.

Go to next page

Round 2 of Block 1

Matching Market

Time left to complete this page: **1:20**

Your Productivity:	**1.1**	**Status**	**Current Offer(s)**
Productivity of Square #6:	3.35	Not Available. Matched with someone else.	---
Productivity of Square #3:	4.02	Not Available. Matched with someone else.	---

Square #6 and Square #3 are matched with other participants. So, you are unmatched.

Your payoff is 1.1 points.

Go to next page

Instructions of Block 2

In this Block, you can match early in the first stage and exit the round, and do not go to the Matching Market at the end of the second stage.

Rules of early matching in the first stage

Before you randomly meet in the first stage, you need to input your lowest acceptable productivity level to match early and exit the round. You will be matched with your first stage partner only if **both of you are acceptable each other.** If you are matched with your first stage partner, then you will earn the match points.

If you do not match in the first stage with your first stage partner, you can still send proposals to each other in the Matching Market. In that case, in the second stage you will meet with a new participant from the opposite role who did not match early and exit the round. Then, there will be a decentralized matching market as in Block 1.

Instructions of Block 3

In this Block, you can match early in the first stage and exit the round, and do not go to the Matching Market at the end of the second stage.

Rules of early matching in the first stage

Before you randomly meet in the first stage, you need to input your lowest acceptable productivity level to match early and exit the round. You will be matched with your first stage partner only if **both of you are acceptable each other.** If you are matched with your first stage partner, then you will earn the match points.

If you do not match in the first stage with your first stage partner, then you cannot match with your first stage partner again. That is, you can match with your first stage partner only in the first stage.

If you do not match in the first stage, then you will meet with a new participant from the opposite role who did not match early and exit the round in the second stage. Since being unmatched yields the lowest possible point, and you have no other potential match partner, then you will be matched with whoever you meet in the second stage and earn the match points.

References

Adachi, Hiroyuki. 2003. "A search model of two-sided matching under nontransferable utility." *Journal of Economic Theory*, 113.

Agranov, Marina, Ahrash Dianat, Larry Samuelson, and Leeat Yariv. 2021. "Paying to Match: Decentralized Markets with Information Frictions."

Akbarpour, Mohammad, Shengwu Li, and Shayan Oveis Gharan. 2020. "Thickness and Information in Dynamic Matching Markets." *Journal of Political Economy*, 128(3).

Ambuehl, Sandro, and Vivienne Groves. 2020. "Unraveling over time." *Games and Economic Behavior*, 121.

Ashlagi, Itai, Afshin Nikzad, and Philipp Strack. 2019. "Matching in dynamic imbalanced markets." *Available at SSRN 3251632*.

Baccara, Mariagiovanna, SangMok Lee, and Leeat Yariv. 2020. "Optimal Dynamic Matching." *Theoretical Economics*, 15(3).

Bramoullé, Yann, Brian W. Rogers, and Erdem Yenerdag. 2021. "Matching with Recall." *Working Paper*.

Chade, Hector, Jan Eeckhout, and Lones Smith. 2017. "Sorting through Search and Matching Models in Economics." *Journal of Economic Literature*, 55(2).

Chen, Daniel L., Martin Schonger, and Chris Wickens. 2016. "oTree—An open-source platform for laboratory, online, and field experiments." *Journal of Behavioral and Experimental Finance*, 9(C): 88–97.

Damiano, Ettore, Hao Li, and Wing Suen. 2005. "Unravelling of Dynamic Sorting." *The Review of Economic Studies*, 72(4).

Du, Songzi, and Yair Livne. 2016. "Rigidity of Transfers and Unraveling in Matching Markets." *mimeo*.

Echenique, Federico, and Juan Sebastián Pereyra. 2016. "Strategic complementarities and unraveling in matching markets." *Theoretical Economics*, 11.

Echenique, Federico, and Leeat Yariv. 2012. "An experimental study of decentralized matching."

Echenique, Federico, Ruy Gonzalez, Alistair Wilson, and Leeat Yariv. 2020. "Top of the Batch: Interviews and the Match." *arXiv preprint arXiv:2002.05323*.

Fainmesser, Itay P. 2013. "Social networks and unraveling in labor markets." *Journal of Economic Theory*, 148.

Halaburda, Hanna. 2010. "Unravelling in two-sided matching markets and similarity of preferences." *Games and Economic Behavior*, 69.

He, Simin, Jiabin Wu, and Hanzhe Zhang. 2020. "Decentralized matching with transfers: experimental and noncooperative analyses." Working paper.

Hoppe, Heidrun C., Benny Moldovanu, and Aner Sela. 2009. "The Theory of Assortative Matching Based on Costly Signals." *The Review of Economic Studies*, 76.

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.

Lauermann, Stephan. 2013. "Dynamic Matching and Bargaining Games: A General Approach." *The American Economic Review*, 103(2).

Leshno, Jacob. 2019. "Dynamic matching in overloaded waiting lists." *Available at SSRN 2967011*.

Li, Hao, and Sherwin Rosen. 1998. "Unraveling in matching markets." *American Economic Review*, 371–387.

Li, Hao, and Wing Suen. 2000. "Risk Sharing, Sorting, and Early Contracting." *Journal of Political Economy*, 108(5).

McKinney, C. Nicholas, Muriel Niederle, and Alvin E. Roth. 2005. "The Collapse of a Medical Labor Clearinghouse (and Why Such Failures Are Rare)." *The American Economic Review*, 95(3).

Niederle, Muriel, and Alvin E. Roth. 2003. "Unraveling Reduces Mobility in a Labor Market: Gastroenterology with and without a Centralized Match." *Journal of Political Economy*, 111(6).

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

Ostrovsky, Micheal, and Micheal Schwarz. 2010. "Information Disclosure and Unraveling in Matching Markets." *American Economic Journal: Microeconomics*, 2.

Pais, Joana, Agnes Pintér, and Robert F Veszteg. 2012. "Decentralized matching markets: a laboratory experiment."

Roth, Alvin E. 1991. "A natural experiment in the organization of entry-level labor markets: Regional markets for new physicians and surgeons in the United Kingdom." *The American economic review*, 415–440.

Roth, Alvin E. 2010. "Marketplace Institutions Related to the Timing of Transactions." National Bureau of Economic Research.

Roth, Alvin E., and Xiaolin Xing. 1994. "Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions." *The American Economic Review*, 84(4).

Smith, Lones. 2006. "The Marriage Model with Search Frictions." *Journal of Political Economy*, 114(6).

Vohra, Akhil. 2020. "Unraveling and Inefficient Matching." *mimeo.*