Optimal Dynamic Matching

Mariagiovanna Baccara* SangMok Lee† Leeat Yariv‡§

January 28, 2018

Abstract

We study a dynamic matching environment where individuals arrive sequentially. There is a tradeoff between waiting for a thicker market, allowing for higher quality matches, and minimizing agents' waiting costs. The optimal mechanism cumulates a stock of incongruent pairs up to a threshold and matches all others in an assortative fashion instantaneously. In discretionary settings, a similar protocol ensues in equilibrium, but expected queues are inefficiently long. We quantify the welfare gain from centralization, which can be substantial, even for low waiting costs. We also evaluate welfare improvements generated by transfer schemes, and alternative priority protocols.

Keywords: Dynamic Matching, Mechanism Design, Organ Donation, Market Design.

*Olin School of Business, Washington University in St. Louis. E-mail: mbaccara@wustl.edu

†Department of Economics, Washington University in St. Louis. E-mail: sangmoklee@wustl.edu

‡Department of Economics, Princeton University. Email: lyariv@princeton.edu

§We thank Johannes Horner, George Mailath, Andy Postlewaite, and Larry Samuelson for very helpful comments. We gratefully acknowledge financial support from the National Science Foundation (SES 0963583).

1 Introduction

1.1 Overview

Many matching processes are inherently dynamic, with participants arriving and matches being created over time. For instance, in the child-adoption process, parents and children arrive steadily – data from one US adoption facilitator who links adoptive parents and children indicates a rate of about 11 new potential adoptive parents and 13 new children entering the facilitator's operation each month.[1] While the overall statistics on the entry of parents and children into the U.S. adoption process are not well-documented, adoption touches upon many lives: The Census 2010 indicates that about 1.5 million or 2.4 percent of all children have been adopted. Likewise, many labor markets entail unemployed workers and job openings that become available at different periods – the U.S. Bureau of Labor Statistics reports approximately five million new job openings and slightly fewer than five million newly unemployed workers each month this year. A similar picture emerges when considering organ donation. According to the Organ Donation and Transplantation Statistics, a new patient is added to the kidney transplant list every 14 minutes and about 3000 patients are added to the kidney transplant list each month. A significant fraction of transplants are carried out using live donors – in 2013, about a third of nearly $17,000$ kidney transplants that took place in the U.S. involved such donors.

Nonetheless, by and large, the extant matching literature has taken a static approach to market design – participants all enter at the same time and the market's operations are restricted in their horizon (see the literature review below for several important exceptions). In the current paper, we offer techniques for extending that approach to dynamic settings.

All of the examples we mentioned above share two important features. First, the quality of matches varies. In other words, agents care about the agents with whom they match. Second, waiting for a match is costly, be it for financial costs of keeping lawyers on retainer for potential adoptive parents, children's hardship from growing older in the care of social services, the lack of wages and needed employees in labor markets, or medical risks for organ patients and psychological waiting costs for donors. These two features introduce a crucial trade-off – on the one hand, a thick market can help generate a greater match surplus; on

[1] This adoption facilitator is one of 25 registered in its state of operation. See Baccara, Collard-Wexler, Felli, and Yariv (2014) for details.

the other hand, a thin market allows for quicker matching and cuts down on waiting costs. The goal of this paper is to characterize the resolution of this trade-off in both centralized and relatively more decentralized environments. Namely, we identify the optimal protocol by which a social planner would match agents over time. We also identify the conditions under which discretionary matching processes would especially benefit from centralized intervention using the optimal protocol.

Specifically, we consider a market that evolves dynamically. There are two classes of agents, which we refer to as "squares" and "rounds." At each period, a pair consisting of a square and a round enters the market. Squares and rounds each have two types, one type more desirable to the other side than the other. For instance, if we think of squares and rounds as children relinquished for adoption and potential adoptive parents, types can stand for gender of children and wealth levels of potential adoptive parents, respectively (see Baccara, Collard-Wexler, Felli, and Yariv, 2014). Alternatively, if we think of the two classes of agents as workers and firms, types can represent skills of workers and benefit packages offered by firms. We assume that preferences are super-modular so that the (market-wide) assortative matching maximizes joint welfare. We also assume that, once agents arrive at the market, waiting before being matched comes at a per-period cost.

We start by analyzing the optimal matching mechanism in such settings, the mechanism that maximizes the expected per-period payoffs for market participants. We show that the optimal mechanism takes a simple form. Whenever congruent pairs of agents – a square and round that are both of the more desirable type or the less appealing type – are present in the market, they are matched instantaneously. When only incongruent agents are present in the market, they are held in a queue. When the stock of incongruent pairs in the queue exceeds a certain threshold, they are matched in sequence, until the queue length falls back within the threshold. Such thresholds induce a Markov process, where states correspond to the length of queues of incongruent pairs of agents. Any threshold yields a different steady-state distribution over possible queue lengths. We evaluate the expected welfare of the mechanism in the steady-state; The optimal mechanism utilizes the threshold that maximizes welfare. When waiting costs are vanishingly small, the welfare under the optimal mechanism approaches the maximum feasible, that generated by no matches of incongruent pairs. As time frictions, or waiting costs increase, the welfare generated by the optimal mechanism decreases.

This welfare decrease raises the question of the value of dynamic clearinghouses for non-trivial waiting costs in different environments, identified by type distributions and preferences. We therefore study the performance of a simple discretionary matching process in our setting. As before, we consider agents arriving at the market in sequence. At each period in which they are on the market, agents declare their willingness to match with partners of either type. After these demands have been made, the maximal number of pairs of willing agents are matched in order of arrival (*first-in-first-out*, or FIFO protocol).[2] Those who prefer to stay in the market, or have to stay for lack of willing partners, form a queue.[3] In our environment, individuals waiting in the market impose a negative externality on those who follow them in the queue, as they force them to wait a longer period and potentially miss desirable matches. On the other hand, waiting in the market generates a positive externality on agents on the other side of the market, as they are more likely to get a quicker match with a partner they prefer. As it turns out, the negative externality of waiting, which is not internalized by individuals, overwhelms this positive externality and leads to excessive waiting in the discretionary setting. In fact, the matching protocol induced by equilibrium in the discretionary matching process ends up resembling the protocol corresponding to the optimal mechanism, but with higher thresholds for the queues' lengths.

We evaluate the difference in welfare generated by a centralized and a discretionary market as a function of the underlying primitives of the environment, namely the agents' type distribution and the cost of waiting.

With respect to the type distribution, as the frequency of desirable types on either side of the market increases, the option value of waiting becomes higher and the wedge between the performance of the centralized and discretionary processes grows.

The comparative statics with respect to costs of waiting are more subtle. An increase in the cost of waiting has a direct and indirect effect. The direct effect is due to the longer expected queues in the discretionary setting. Fixing the expected queue lengths corresponding to the optimal and discretionary processes, an increase in per-period waiting costs on the welfare differential is effectively a multiplier effect – it is the difference between the expected time agents wait in queue under these two processes, multiplied by the change in costs. The indirect effect is that both the optimal threshold as well as the equilibrium threshold in the

[2] This process is reminiscent of a double auction, as each agent submits a "demand function" specifying which types of agents she would be interested in matching with immediately.

[3] We provide a characterization of preferences that assure that the process is individually rational for all participants.

discretionary process decrease as a function of waiting costs. The difference between these two thresholds therefore narrows as costs increase, which works to mute the welfare gap between the two processes. We show that the combination of these effects leads to a welfare wedge that is locally increasing in costs (formally, it is piece-wise increasing), but exhibits a general decreasing trend. Ultimately, when costs are prohibitively high, both processes lead to instantaneous matches and identical welfare levels.

Finally, we ask in which ways one can improve upon a discretionary setting with interventions that are simpler than the full-fledged optimal mechanism. Indeed, centralization may be hard to implement for two main reasons: first, it requires the central planner to be able to force matches upon individuals and, second, it requires the central planner to monitor arrivals and to possibly create matches at every period, yielding potentially high administrative costs. To address the first issue, we consider a discretionary setting in which per-period taxes are introduced for the agents that decide to wait. Our characterization of the optimal mechanism allows us to identify a budget-balanced tax scheme that implements the optimal welfare levels. Such a tax scheme can be tailored so that it does not distort agents' incentives to enter the market to begin with. Nonetheless, even such a scheme may be viewed as cumbersome administratively since it requires continuous monitoring of agents' location in the queue. To address this problem, we consider a discretionaly setting governed by a *last-in-first-out* (LIFO) priority protocol. We show that this alternative protocol still generates excessive waiting, but reduces it with respect to the FIFO protocol. Moreover, we show that the equilibrium under the LIFO protocol is asymptotycally efficient as waiting costs disappear. In the Online Appendix we analyze two additional matching mechanisms: first, we look at a simple mechanism in which all matches occur at fixed time intervals. The length of these time intervals can be chosen to balance the costs of waiting and the quality of the resulting matches, which we characterize. We show that such a simple procedure, while still inferior to the fully optimal mechanism, can improve welfare substantially relative to a fully discretionary market, and it becomes asymptotically efficient as waiting costs go to zero. Second, we consider a discretionary market in which individuals are matched according to a uniformly random priority rule. The results we obtain by means of simulations suggest that both in terms of excessive waiting and welfare performance this mechanism turns out to be intermediate between the FIFO and the LIFO mechanisms.

1.2 Related Literature

The interest surrounding dynamic matching is recent and the literature on this topic is still relatively limited. Much of this literature originally stemmed from the organ donation application. Zenios (1999) develops a queueing model to explain the differences between waiting times of different categories of patients anticipating a kidney transplant. In the context of kidney exchange, Ünver (2010) focuses on a market in which donors and recipients arrive stochastically, preferences are compatibility-based, and the goal of a central planner is to minimize total discounted waiting costs. Under some conditions, he shows that the efficient two-way matching mechanism always carries out compatible bilateral matches as soon as they become available. However, when multi-way matches are possible, some two-way matches could be withheld in order to allow future multi-way matches.[4]

Akbarpour, Li, and Oveis Gharan (2017) is possibly the closest to our study as they also inspect the benefits of different mechanisms in a dynamic matching environment. In their setting, however, preferences are compatibility-based according to a network mapping the set of exchange possibilities. Agents in the system (thought of as patient-donor pairs) become "critical" at random dates, and perish immediately after if they remain unmatched. Therefore, when waiting costs are negligible, the goal of the planner is to minimize the number of perished agents. Market thickness is beneficial in that it guarantees the availability of immediate matches for agents who become critical. Left to their own devices, agents in that setting would match quickly and useful mechanisms induce agents to wait. In contrast, in our setting, agents in a discretionary process wait too long and useful mechanisms induce agents to wait shorter times on the market. In addition, while the welfare benchmark in Akbarpour, Li, and Oveis Gharan (2016) is that of an omniscient planner, our different focus allows us to characterize the optimal mechanism, which serves as the benchmark for welfare comparisons.[5]

[4]Some recent models in inventory control have a similar flavor to the compatibility-based matching process considered by Ünver (2010), see e.g. Gurvich and Ward (2014) and Hu and Zhou (2016).

[5]Loertcher, Muir, and Taylor (2016) follow up on the current paper and consider a setting similar to ours. They focus on the optimal mechanism when the planner and participants use the same discount factor to assess future utilities. The interpretation of the objective function in their setting is subtle. In particular, when two agents of identical types who arrived at different times are matched at date t, the agents themselves experience different discounted utilities, but the planner's utility from the two matches is identical. Similarly, when assessing future matches, the planner cares only about when they are formed, rather than on how much wait they generated for participants. In other words, individual costs of delay do not enter the planner's objective function. Herbst and Schickner (2016) study markets in which agents are drawn from a unique pool (in contrast with our two-sided setting). They consider environments in which heterogeneous agents arrive randomly over time and need to be grouped in pairs. They characterize the optimal mechanism and analyze

In a somewhat different realm, Leshno (2015) studies a one-sided market in which objects (say, public houses) need to be allocated to agents who wait in a queue. Welfare maximization always requires agents to be matched to their preferred objects. However, if agents' preferences are unknown to the planner and their preferred item is in short supply, agents may prefer a mismatched item earlier to avoid costly waiting. Leshno (2015) shows that the welfare loss from mismatches can be reduced substantially through a policy under which all agents who decline a mismatched item face the same expected wait for their preferred item.[6] Anderson, Ashlagi, Gamarnik, and Kanoria (2015) study an environment in which each agent is endowed with an item that can be exchanged with an item owned by someone else. Compatibility is stochastic, and three classes of feasible exchanges are considered: two-way exchanges only, two- and three-way cycles, and any kind of chain. They find that a policy that maximizes immediate exchanges without withholding them in the interest of market thickness performs nearly optimally.[7,8]

There is also a recent theoretical literature that studies discretionary matching processes that are dynamic, considering both informational and time frictions (see, e.g., Haeringer and Wooders, 2011, Niederle and Yariv, 2009, and Pais, 2008). In that literature, the number of agents on each side of the market is fixed at the outset and agents on one side can make directed offers to agents on the other side. The main goal is the identification of market features that guarantee that an equilibrium of the induced game generates a stable matching.

Another related strand of literature is the search and matching literature (e.g., Burdett and Coles, 1997, Eeckhout, 1999, and the survey by Rogerson, Shimer, and Wright, 2005). There, each period, workers and firms randomly encounter each other, observe the resulting match utilities, and decide jointly whether to pursue the match and leave the market or to separate and wait for future periods. In markets with assortative preferences, as time frictions vanish,

the impacts of incomplete information in such team-formation settings.

[6] Bloch and Cantala (2016) also study dynamic allocations. In their setting, a mechanism is a probability distribution over all priority orders consistent with a closed waiting list. Given a priority order, whenever a new object becomes available, agents are proposed the object in sequence and can either accept or reject. They show the benefits of a strict seniority order. Schummer (2016) considers the welfare implications of policies that impact deferral choices by impatient agents waiting in queue. He shows that individuals' level of impatience and risk aversion determines whether allowing deferral of sub-optimal options is desirable.

[7] On the benefits of algorithms that create thicker pools in sparse dynamic allocation environments, see also Ashlagi, Jaillette, and Manshadi (2013).

[8] While most of this literature has been focusing on the design of algorithms to achieve socially desirable matchings Doval (2016) introduces a notion of stability in dynamic environments, and provides conditions under which dynamically stable allocations exist.

generated outcomes are close to a stable matching. A crucial difference with our setting is the stationarity of the market – the perceived distribution of potential partners does not change with time, and each side of the market solves an option value problem.[9]

Last, there is a rather large literature that considers dynamic matching of buyers and sellers and inspects protocols that increase efficiency or allow for Walrasian equilibrium outcomes to emerge as agents become increasingly patient (see, e.g., Satterthwaite and Shneyerov, 2007 and Taylor, 1995).[10]

2 Setup

We study an infinite-horizon dynamic matching market. There are two kinds of agents: squares and rounds. Squares and rounds can stand for potential adoptive parents and children relinquished for adoption, workers and employers, patients and donors, etc.[11]

At each time $t \in \{1, 2, \dots\}$, one square and one round arrive at the market. Each square can be of either type A or B with probability p or $1-p$, respectively, and each round can be of type α or β with probability p or $1-p$, respectively. These types correspond to the attributes of participants – they can stand for the wealth of parents and race of children in the adoption application, level of education of employees and social benefits or promotion likelihoods for employers in labor markets[12], age or tissue types in the organ donation context[13], etc.

[9] In the context of marriage markets, Kocer (2014) considers learning over time and models the choice of temporary interactions with different potential partners as a multi-armed bandit problem. Choo (2015) develops a new model for empirically analyzing dynamic matching in the marriage market and applies that model to recent changes in the patterns of US marriages.

[10] Related, Budish, Cramton, and Shim (2015) consider financial exchanges and argue that high-frequency trading leads to inefficiencies, while frequent batch auctions, uniform-price double auctions that occur at fixed and small time intervals, can provide efficiency improvements. These results are reminiscent of our observations regarding the welfare improvements generated by matchings that occur at the end of each fixed window of time.

[11] The organ donation application shares some features with our model when considering live donations from good samaritan donors, in which case the matching process is two-sided in nature.

[12] In some markets, wages differ across individual employees and can be thought of as transfers, which this paper does not handle. However, Hall and Kruger (2012) suggest that a large fraction of jobs have posted wages. Naturally, these wages may reflect general equilibrium wages tailored to the precise composition of the market. Nonetheless, the fact that these wages are fairly constant and do not fluctuate dramatically suggests they may not respond to particular characteristics of individual employees. Our model speaks to this segment of the market.

[13] This is a simplified representation that aims at capturing heterogeneity in types and the quality of different matches it implies. For organ donation, patients and donors are often classified into coarse categories based on age. However, more than two tissue types are often considered. For example, for kidney transplantation, the medical community currently looks at six tissue types, called major histo-compatibility complex or HLA

In our model, squares seek to match with rounds and vice versa. We denote by $U_x(y)$ the surplus for a type-x participant from matching with a type-y participant. We assume that preferences are assortative: A-squares are more desirable for all rounds and that α-rounds are more desirable for all squares. That is,

$$\begin{aligned} U_A(\alpha) &> U_A(\beta), \qquad U_B(\alpha) > U_B(\beta), \\ U_\alpha(A) &> U_\alpha(B), \qquad U_\beta(A) > U_\beta(B). \end{aligned}$$

It will be convenient to denote:

$$\begin{array}{ll} U_{A\alpha} \equiv U_A(\alpha) + U_\alpha(A), & U_{A\beta} \equiv U_A(\beta) + U_\beta(A), \\ U_{B\alpha} \equiv U_B(\alpha) + U_\alpha(B), & U_{B\beta} \equiv U_B(\beta) + U_\beta(B), \end{array}$$

as well as

$$U \equiv U_{A\alpha} + U_{B\beta} - U_{A\beta} - U_{B\alpha}.$$

We will further assume that $U > 0$ so that the utilitarian efficient matching in a static market creates the maximal number of (A, α) and (B, β) pairs. The value of U captures the efficiency gain from such an assortative matching relative to the anti-assortative matching. Notice that $U > 0$ is tantamount to assuming super-modular assortative preferences (a-la Becker, 1974) and U can be thought of as the degree of super-modularity preferences exhibit.

We assume that each square and round suffer a cost $c > 0$ for each period they spend on the market waiting to be matched. We also assume that agents leave the market only by matching. In Section 7 we provide bounds on the utility of agents from remaining unmatched that assures this assumption is consistent with individual rationality in the processes we analyze.[14]

Several assumptions merit discussion. We assume that preferences are super-modular and that waiting costs are identical for squares and rounds for presentation simplicity. These assumptions are common in the literature and, as we describe in Section 5, lead to a conservative comparison of the optimal and discretionary matching protocols.[15]

antigens. For an extension to richer type sets, see the Online Appendix.

[14] We show that individual rationality is guaranteed when all agents are acceptable and when any β-round receives a utility lower than $U_\beta(B) - \frac{p}{1-p}\left[U_\alpha(A) - U_\alpha(B)\right]$ when leaving the market unmatched (analogously for B-squares).

[15] Our analysis carries through fully if participants are horizontally differentiated; In particular, if A-squares and α-rounds prefer one another and B-squares and β-rounds prefer one another, in which case the utilitarian efficient static matching is assortative without further assumptions. This may be relevant in some child-adoption contexts if both adoptive parents and birth mothers display what is often termed "homophilic preferences," preferring to match with individuals of their own race. We describe our results assuming assortative preferences since they are a leading example in the extant literature and potentially tie to more applications.

The assumption that the distribution of types of rounds mirrors that of squares also simplifies our analysis substantially as will be seen. It essentially implies that if we drew a large population of rounds and squares, the realized distributions of types would be approximately balanced with high probability. This may be a fairly reasonable assumption for certain applications, such as organ donation. Indeed, the distribution of tissue types of donors and patients is arguably similar. Furthermore, the age of a donor is known to have a strong impact on the expected survival of a graft (see, e.g., Gjertson, 2004 and Oien et al., 2007) and younger recipients have been suggested as the natural recipients of higher-quality organs (see Stein, 2011). Our assumptions then fit a world in which both patients and donors are classified as "young" or "mature" and patients' and donors' age distributions are similar. Our assumption also approximately holds for certain attributes in the online dating world (see Hitch, Hortacsu, and Ariely, 2010). Nonetheless, it might be a rather harsh assumption for other applications. As it turns out, the techniques we introduce can be used were we to relax that assumption. We replicate some of our analysis for general asymmetric settings in the Online Appendix.

The other strong assumption we make is that a pair of agents arrives at the market in each period. The analysis would remain virtually identical were we to assume that pairs arrive at random times following, say, a Poisson distribution. Moreover, our results extend directly if we assume that each period a fixed number of square-round pairs (possibly greater than one) enter the market. However, the assumption that participants arrive in pairs is important for the techniques we use. This assumption assures that the market is balanced throughout the matching process. It is a reasonable assumption for some applications. For example, in the adoption process presumably potential adoptive parents and birth mothers make important decisions (whom to match with, whether to leave the process, etc.) at spaced-out intervals. Given the limited variability in the volume of entrants on a monthly basis, the assumption of balanced arrivals provides a decent approximation of reality. Allowing for different arrival rates on both sides of the market introduces new considerations as matching participants in thin markets, while still entailing low waiting costs, imposes now a loss in terms of both the quality of matches and the number of individuals matched. The Online Appendix offers a more thorough discussion of how the analysis might be extended to allow for different arrival rates of squares and rounds.

Another assumption we adopt is the fact that the agents incur a fixed cost c for every period they spend on the market unmatched. An alternative way to model waiting costs would be to consider agents' payoffs as discounted match utilities. As a first step, our criterion allows us for greater tractability. To see why, notice the benefits of matching an agent would then depend on the number of periods that agent already spent on the market. Also, note that the randomness inherent in the environment suggests that the timing of matches is potentially a random variable. Keeping track of expected exponentially discounted values then introduces non-trivial complications.

Last, the assumption that there are only two types of squares and rounds is made for tractability. It corresponds to a coarse description of types in many applications. In the Online Appendix, we illustrate the insights our binary-type analysis allows for environments with richer type sets.

3 Optimal Dynamic Matching

3.1 The Matching Process

At any time $t \in \{1, 2, \dots\}$, after a new square-round pair enters the market, a queue corresponds to a vector $\mathbf{s}^t = (s_A, s_B, s_\alpha, s_\beta)$, where each entry is *the stock of squares or rounds* of a particular type waiting in line. We represent by vector $\mathbf{l}^t = (l_{A\alpha}, l_{A\beta}, l_{B\alpha}, l_{B\beta})$ the profile of matches created at time t. For every $\mathbf{s}^t \in \mathbb{Z}_+^4$, a match profile $\mathbf{l}^t \in \mathbb{Z}_+^4$ has to satisfy a feasibility condition

$$\begin{aligned} l_{x\alpha} + l_{x\beta} &\leq s_x \quad \text{for } x \in \{A, B\}, \\ l_{Ay} + l_{By} &\leq s_y \quad \text{for } y \in \{\alpha, \beta\}. \end{aligned}$$

The surplus generated by the matches is:

$$S(\mathbf{l}) \equiv \sum_{(x,y) \in \{A,B\} \times \{\alpha,\beta\}} l_{xy} U_{xy}.$$

We denote the volume of remaining agents by $\mathbf{k}^t = (k_A, k_B, k_\alpha, k_\beta)$, where

$$\begin{aligned} k_x &= s_x - (l_{x\alpha} + l_{x\beta}) \quad \text{for } x \in \{A, B\}, \\ k_y &= s_y - (l_{Ay} + l_{By}) \quad \text{for } y \in \{\alpha, \beta\}. \end{aligned}$$

The total waiting costs incurred by the remaining agents in period t are then:

$$C(\mathbf{s}, \mathbf{l}) \equiv c \left(\sum_{x \in \{A,B,\alpha,\beta\}} k_x \right).$$

Finally, the welfare generated at time t is

$$w(\mathbf{s}, \mathbf{l}) \equiv S(\mathbf{l}) - C(\mathbf{s}, \mathbf{l}),$$

if the profile of matches $\mathbf{l}$ is feasible given the stock $\mathbf{s}$, and $w(\mathbf{s}, \mathbf{l}) = -\infty$ otherwise. At time $t+1$, the queue $\mathbf{s}^{t+1}$, is determined by the number of remaining agents $\mathbf{k}^t$ and the types of agents arriving at $t+1$. As an initial condition, we have $\mathbf{k}^0 = (0, 0, 0, 0)$. A *mechanism* μ is any rule governing matching profiles. We evaluate a mechanism by considering the *average welfare* it generates:

$$v(\mu) \equiv \liminf_{T\to\infty} \frac{1}{T} E_\mu \left[\sum_{t=1}^{T} w(\mathbf{s}^t, \mathbf{l}^t) \right]. \tag{1}$$

Note that for any mechanism μ, $v(\mu) \in \mathbb{R} \cup \{-\infty\}$, and it is bounded above by $U_{A\alpha}$. This criterion allows us to focus on the long-run performance of mechanisms. We say that μ^* is *optimal* if

$$v(\mu^*) = \sup_{\mu} v(\mu).$$

An important subclass of the set all mechanisms is the set of *stationary and deterministic mechanisms* (SD-mechanisms). The matches created by a SD-mechanism $\mu^{SD} : \mathbb{Z}^4_+ \to \mathbb{Z}^4_+$ at every period depend only on the queue in place at that period.

Consider the class of SD-mechanisms which satisfy the following two conditions:

Condition (1) The mechanism matches (A, α) and (B, β) pairs as soon as they become available;

Condition (2) The mechanism never holds more than $\frac{U}{2c}$ squares (and rounds) in the market.

While for simplicity of exposition from now on we restrict our focus on SD-mechanisms satisfying both conditions (1) and (2), in the Appendix we show that this restriction is without loss of generality since there exists an optimal SD-mechanism that satisfies both Conditions (1) and (2).

To conclude this section, we address another important feature of our set-up, which is the absence of discounting in the planner's objective function: the planner uses the average of the agent's match utilities, net of the waiting costs. An alternative assumption would be considering exponentially-discounted match utilities , evaluated at time $t = 0$ from the planner's perspective. However, substituting our objective function (1) with a discounted sum of match utilities evaluated at $t = 0$ would cause all unmatched agents to be incurring in waiting costs regardless of them being already on the market or not, which would be an unappealing feature of the model given the applications we are interested in.[16]

3.2 Structure of Optimal Dynamic Mechanisms

Conditions (1) and (2) imply that at any point in time an optimal dynamic mechanism entails queues of only A-squares and β-rounds, or only B-squares and α-rounds. That is, the queue can take the form of either $(k, 0, 0, k)$ or $(0, k, k, 0)$, for some $k \geq 0$. The optimal dynamic mechanism is then identified by the maximal stock of A-squares (and β-rounds) and the maximal stock of α-rounds (and B-squares) that are kept waiting in queue. In the following proposition we characterize the structure of the optimal mechanism.[17]

Proposition 1 (Optimal Mechanisms) *An optimal dynamic mechanism is identified by a pair of thresholds* $(\bar{k}_A, \bar{k}_\alpha) \in Z_+$ *such that*

1. *whenever more than* $\bar{k}_A$ *A-squares are present, the excess pairs of type* (A, β) *are matched immediately, and*
2. *whenever more than* $\bar{k}_\alpha$ *α-rounds are present, the excess pairs of type* (B, α) *are matched immediately.*

As will soon be stated formally, the symmetry of our environment assures that, generically, an optimal mechanism corresponds to symmetric thresholds: $\bar{k} = \bar{k}_A = \bar{k}_\alpha$.[18] A dynamic

[16] In fact, note that substituting (1) with a discounted sum of match utilities simplifies the analysis in some respects, as it would generate a unique solution to the optimal SD-mechanism problem. To see this, consider a mechanism that holds both a A-square and an α-round on the market for a finite number of periods rather than matching them immediately as Condition (1) prescribes. Such mechanism does not necessarily decrease the value of (1), while it would decrease a discounted sum of match utilities.

[17] In principle, there could be multiple mechanisms that are identified with the same thresholds. For instance, consider a mechanism in which a pair of type (A, β) is matched whenever there are $\bar{k}_A + 1$ A-squares present, or when there are $\bar{k}_A + 2$ A-squares present. Such a mechanism would be equivalent to a mechanism that matches (A, β) pairs only when there are precisely $\bar{k}_A + 1$ A-squares present. In that sense, we will focus only on the thresholds with the minimal magnitude, which identify outcomes fully, and ignore multiplicity that arises from prescriptions of the social planner over events that are never reached.

[18] As mentioned, in the Online Appendix we work out the extension to asymmetric environments, where the two thresholds may differ.

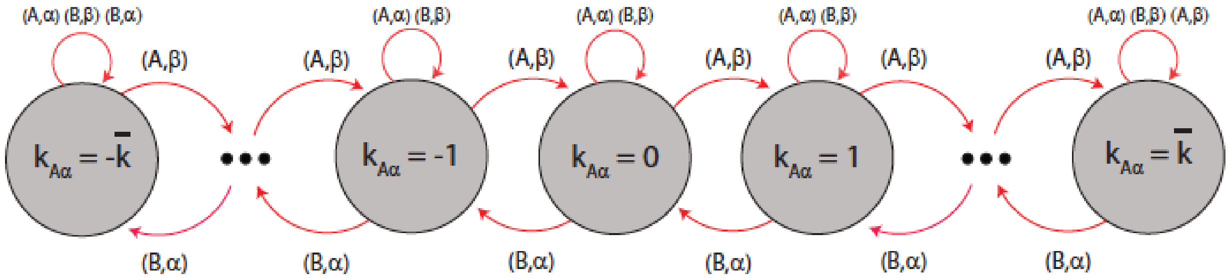

Figure 1: Structure of Optimal and Decentralized Matching Processes

mechanism with symmetric thresholds $(\bar{k}, \bar{k})$ as defined in Proposition 1 is depicted in Figure 1, where $k_{A\alpha} = k_A - k_\alpha$ captures the difference between the length of the queue of A-squares and the length of the queue of α-rounds. We call $k_{A\alpha}$ the *(signed) length of the A-α queue.*

This process induces the following Markov chain. Let $k^t_{A\alpha}$ denote the number of A-squares (or β-rounds) minus the number of α-rounds (or B-squares) at the end of time t, after the arrival of that period's square-round pair and any matches imposed by the mechanism. If an (A,α) or a (B,β) pair arrive in period $t+1$, the mechanism matches an (A,α) or a (B,β) pair immediately, and the state remains the same: $k^t_{A\alpha} = k^{t+1}_{A\alpha}$. Suppose an (A,β) pair arrives in period $t+1$. As long as $0 \le k^t_{A\alpha} < \bar{k}$, the mechanism creates no matches and $k^{t+1}_{A\alpha}$ becomes $k^t_{A\alpha}+1$. If $k^t_{A\alpha} < 0$, the mechanism creates one (A,α) match and one (B,β) match, and $k^{t+1}_{A\alpha}$ becomes $k^t_{A\alpha}+1$. Finally, if $k^t_{A\alpha} = \bar{k}$, the mechanism creates one (A,β) pair, and $k^{t+1}_{A\alpha}$ remains the same, $k^{t+1}_{A\alpha} = k^t_{A\alpha} = \bar{k}$. Analogous transitions occur with the arrival of a (B,α) pair.

Therefore, we can describe the probabilistic transition as follows. Denote by

$$\mathbf{x}^t \equiv (x^t_{-\bar{k}}, x^t_{-\bar{k}+1}, ..., x^t_{\bar{k}-1}, x^t_{\bar{k}})^{tr} \in \{0,1\}^{2\bar{k}+1}$$

the timed vector capturing the state, $x^t_i = 1(k^t_{A\alpha} = i)$ – that is, x^t_i is an indicator that takes the value of 1 if the state is i and 0 otherwise. Then,

$$\mathbf{x}^{t+1} = \mathbf{T}_{\bar{k}} \mathbf{x}^t,$$

where

$$\mathbf{T}_{\bar{k}} = \begin{pmatrix} 1-p(1-p) & p(1-p) & \dots & 0 & 0 \\ p(1-p) & 1-2p(1-p) & \dots & 0 & 0 \\ 0 & p(1-p) & \dots & & \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ & & \dots & p(1-p) & 0 \\ 0 & 0 & \dots & 1-2p(1-p) & p(1-p) \\ 0 & 0 & \dots & p(1-p) & 1-p(1-p) \end{pmatrix}. \quad (2)$$

The above Markov chain is ergodic (i.e., irreducible, aperiodic, and positively recurrent). Therefore, an optimal mechanism corresponds to a matching process that reaches a steady state with a unique stationary distribution. For $T_{\bar{k}}$, the steady-state distribution is uniform so that each state $k_{A\alpha} = -\bar{k}, \ldots, \bar{k}$ occurs with an equal probability of $\frac{1}{2\bar{k}+1}$.

3.3 Optimal Thresholds

In order to characterize the optimal threshold, we first evaluate the welfare corresponding to any arbitrary symmetric threshold. First, we compute the average total waiting costs incurred by agents waiting in line for one period of time. During the transition from time $t-1$ to time t, $2|k_{A\alpha}^{t-1}|$ agents wait in line. So the total costs of waiting incurred during this one time period is $2|k_{A\alpha}^{t-1}|c$. Thus, a mechanism with threshold $\bar{k}$ results in expected total costs of waiting equal to

$$\frac{1}{2\bar{k}+1}\left(\sum_{k_{A\alpha}=-\bar{k}}^{\bar{k}} 2|k_{A\alpha}|\right)c = \frac{2\bar{k}(\bar{k}+1)c}{2\bar{k}+1}.$$

Next we compute the average total surplus generated during one time period, tracking the Markov process described above. A new arrived pair is of type (A, α) with probability p^2, in which case the optimal mechanism generates a surplus equal to $U_{A\alpha}$. Similarly, when a new pair of type (B, β) arrives, which occurs with probability $(1-p)^2$, the optimal mechanism generates a surplus equal to $U_{B\beta}$.

Suppose an (A, β) pair arrives at time t. If $k_{A\alpha}^{t-1} < 0$, the mechanism creates one (A, α) and one (B, β) pair, generating a surplus equal to $U_{A\alpha} + U_{B\beta}$. If $0 \leq k_{A\alpha}^{t-1} < \bar{k}$, the mechanism creates no matches (and no additional surplus), and if $k_{A\alpha}^{t-1} = \bar{k}$, the mechanism creates one (A, β) pair and generates surplus equal to $U_{A\beta}$. Analogous conclusions pertain to the case in which a (B, α) pair arrives. Thus, a mechanism with threshold $\bar{k}$ generates an expected total surplus equal to

$$\begin{aligned} & p^2 U_{A\alpha} + (1-p)^2 U_{B\beta} + \frac{2p(1-p)}{2\bar{k}+1}\left[\bar{k}\left(U_{A\alpha} + U_{B\beta}\right) + \frac{U_{A\beta} + U_{B\alpha}}{2}\right] \\ = \; & pU_{A\alpha} + (1-p)U_{B\beta} - \frac{p(1-p)U}{2\bar{k}+1}. \end{aligned}$$

Therefore, the net expected total welfare per period, accounting for waiting costs, is:

$$pU_{A\alpha} + (1-p)U_{B\beta} - \frac{p(1-p)U}{2\bar{k}+1} - \frac{2\bar{k}(\bar{k}+1)c}{2\bar{k}+1}. \tag{3}$$

When the cost of waiting c is high, waiting for potentially higher-quality matches does not justify any waiting and the mechanism matches agents instantaneously as they arrive. For sufficiently low costs, however, the optimal mechanism exhibits non-trivial waiting. The optimal threshold $\bar{k}^{opt}$ maximizes the welfare as given in (3).

The following proposition summarizes our discussion and provides the full characterization of the optimal mechanism.

Proposition 2 (Optimal Thresholds) *The threshold*

$$\bar{k}^{opt} = \left\lfloor \sqrt{\frac{p(1-p)U}{2c}} \right\rfloor$$

identifies an optimal dynamic mechanism. In this optimal mechanism, all available (A, α) and (B, β) pairs, and any number of (A, β) or (B, α) pairs exceeding $\bar{k}^{opt}$, are matched immediately. Furthermore, the optimal mechanism is generically unique.[19]

The optimal threshold increases with the probability of any incongruent pair, $p(1-p)$, and with the degree of super-modularity U, which reflects the value of generating assortative matches. It decreases with waiting costs. In fact, when waiting costs are prohibitively high, namely when $c > \frac{p(1-p)U}{2}$, the maximal queue length is $\bar{k}^{opt} = 0$ and all matches are instantaneous.

3.4 Welfare

We now turn to the expected per-period welfare in the steady state under the optimal mechanism. Were we to consider no costs of waiting, the optimal mechanism would naturally entail long waits to get the maximal possible match surplus asymptotically by matching A-squares with α-rounds and B-squares with β-rounds. We denote the resulting welfare by:

$$S_\infty \equiv pU_{A\alpha} + (1-p)U_{B\beta}.$$

The optimal threshold identified in Proposition 2 allows us to characterize the welfare achieved by the optimal mechanism through equation (3) and to get the following corollary.[20]

[19] Multiplicity arises only when $\sqrt{\frac{p(1-p)U}{2c}}$ is an integer.

[20] In the Appendix, we provide the analytical formula for $\Theta(c)$ in terms of the fundamental parameters of our setting.

Corollary 1 (Optimal Welfare) *The welfare under the optimal mechanism is given by $W^{opt}(c) = S_\infty - \Theta(c)$, where $\Theta(c)$ is continuous, increasing, and concave in c, $\lim_{c \to 0} \Theta(c) = 0$, and $\Theta(c) = p(1-p)U$ for all $c \geq \frac{p(1-p)U}{2}$.*

As waiting costs approach 0, the welfare induced by the optimal mechanism approaches S_∞. For costs large enough, the optimal mechanism matches all square-round pairs instantaneously as they arrive and the resulting welfare is $S_\infty - p(1-p)U$. For intermediate costs, the optimal mechanism generates welfare that is naturally in between these two values.[21,22] The observation that welfare under the optimal mechanism decreases as c increases is rather intuitive. Indeed, suppose $c_1 > c_2$. Were we to implement the optimal mechanism with waiting cost c_1 when the waiting cost is c_2, the distribution of matches would remain identical, while waiting costs would go down, thereby leading to greater welfare overall. This implies that the welfare under the optimal mechanism with waiting cost c_2, which would be at least weakly higher, is greater than that corresponding to waiting cost c_1. The amount by which welfare decreases when waiting costs increase depends on the number of agents expected to wait in line in the steady state. The higher the waiting costs, the lower the number of agents waiting in line on average. Therefore, the impact of an increase in costs by a fixed increment is greater at smaller costs, which leads to the concavity of $\Theta(c)$.[23]

4 Discretionary Matching

4.1 Matching Process

Many dynamic matching processes are in essence discretionary, in the sense that participants have the choice of declining a match they do not wish to form: child adoption in the US and

[21] Notice that the value of S_∞ is effectively the analogue of the value generated by an "omnsicient" planner in our setting, which is used as one benchmark in Akbarpour, Li, and Oveis Gharan (2017). Corollary 1 suggests that the omniscient planner's value is a valid feasible benchmark when waiting costs vanish.

[22] In fact, simple algebraic manipulations imply that:

$$S_\infty - \sqrt{2p(1-p)Uc} - c \leq W^{opt}(c) \leq S_\infty - \sqrt{2p(1-p)Uc} + c.$$

[23] Continuity follows directly from concavity. Alternatively, fix any mechanism that is optimal for some waiting costs. An increase in waiting costs reduces the resulting welfare continuously, in fact linearly. Thus, we cannot get discontinuous reductions in welfare when choosing an optimal mechanism as waiting costs increase.

abroad, job searches in many industries, etc. It is therefore important to understand the implications of discretionary dynamics, particularly when considering centralized interventions.

In our discretionary matching process, we assume individuals join the market in sequence and decide when to match with a potential partner immediately and when to stay in the market and wait for a potentially superior match. While the discretionary setting we study still requires some centralized governance, as matches occur according to some order, it provides a convenient benchmark for studying dynamic matching markets that are lightly regulated.

We assume that at each period t, there are three stages. First, a square and round enter the market with random attributes as before: with probability p the square is an A-square and with probability p the round is an α-round. Second, individuals of each type are ordered by some priority rule that we describe later. In the third stage, each square and round declare their demand – whether a square will match only with an α-round, or is willing to match with either an α-round or a β-round, and whether a round will match only with an A-square, or with either an A-square or a B-square. Given the order and the participants' demands, the market clears sequentially as we describe below. Any remaining participants proceed to period $t+1$ at the additional cost of c.

4.2 A Formal Model

In each period t, one square w^t and one round r^t arrive at the market, and their types are realized. Upon their arrival, a period-t stage-game begins:

$$G^t \equiv \{I^t, (D_i)_{i \in I^t}, \phi, (u_i(\cdot;\phi) : \Pi_{i \in I^t} D_i \to \mathbb{R})_{\forall i \in I^t}\}.$$

The components of G^t are defined as follows. The set of players is $I^t \equiv A^t \cup B^t \cup \alpha^t \cup \beta^t$, where $A^t \subseteq \{w^{t'} : 1 \leq t' \leq t\}$ is the set of A-squares present in the market in period t, and the other sets - B^t, α^t, and β^t - are defined similarly. Each A-square, say player i, in I^t chooses an action in $D_i = \{\alpha, \beta\}$, with α denoting a demand for only α-rounds and β denoting a demand for *both* types of rounds.[24] The action sets for the other agents' types are defined similarly. A *priority rule* ϕ assigns a linear order over each set A^t, B^t, α^t, and β^t. First, we consider a *first-in-first-out (FIFO) protocol*, which assigns a linear order $\succ$ over, say, A^t such that

$$\forall w^{t'}, w^{t''} \in A^t, \quad w^{t'} \succ w^{t''} \iff t' < t'' \leq t.$$

[24] This restriction on the actions' space is made for simplicity of exposition. An equilibrium similar to the one we describe below arises if we allow players to demand only inferior matches on the other side of the market, or to submit no demand at all.

Indeed, there is anecdotal evidence that order of arrivals affects the order of matches in many markets, and FIFO is a commonly used protocol among those. For instance, in the child adoption context, many countries follow a FIFO protocol to match relinquished children to adoptive parents.[25] In Section 6.2 and in the Online Appendix, we consider alternative priority protocols.

The stage-game payoffs are determined by a sequential market clearing. First, we take A-squares and α-rounds in the order induced by ϕ and form as many (A, α) pairs as possible (regardless of their demands).[26] If there are remaining A-squares demanding β-rounds, they are matched with β-rounds sequentially according to ϕ and independently of the demands made by the β-rounds. Analogously, if there are remaining α-rounds demanding B-squares, they are matched with B-squares sequentially according to ϕ and independently of the demands made by the B-squares. Last, the remaining B-squares and β-rounds who demand each others' types form matches sequentially in the order induced by ϕ. The stage-game payoff for a type-x agent matched with a type-y agent is $U_x(y)$. If a player remains unmatched, her stage-game payoff is $-c$.

We complete the definition of a dynamic discretionary matching game by characterizing the evolution of the stage-games G^t, and each player's dynamic game payoff. The initial set of players is $I^0 \equiv \emptyset$. All players in I^t who remain unmatched in period t form I^{t+1} together with new arrivals w^{t+1} and r^{t+1}. Consider a player i, who arrives in period t and is matched at t''. Such player receives stage-game payoffs $(u_i^t, u_i^{t+1}, \dots)$, and a dynamic game payoff $\sum_{t'=t}^{\infty} u_i^{t'} (\in \mathbb{R} \cup \{-\infty\})$, where $u_i^{t'} = 0$ for $t' > t''$.

The dynamic game has complete information and arbitrary (dynamic) strategies. Each player i, say an A-square, chooses a demand in every period from when she arrives until she matches. A (dynamic) strategy σ_i indicates the probability of demanding α in each of these periods and can depend on the complete history from $t = 0$ on. As before, let $\mathbf{s}^t = (s_A^t, s_B^t, s_\alpha^t, s_\beta^t)$ be the *state* (or stock) at period t and let $q_i^t \in \mathbb{Z}_+$ be player i's *rank* according to ϕ in period t.[27] Let $\theta_i^t \equiv (\mathbf{s}^t, q_i)$ denote an *augmented state* for player i.

[25] For example, see the protocol adopted by the China Center of Children's Welfare and Adoption (CCCWA) here: http://www.aacadoption.com/programs/china-program.html.

[26] This market-clearing assumption allows us to simplify some steps of the proofs, and avoid inefficient equilibrium scenarios in which (A, α) pairs remain on the market unmatched.

[27] That is, under FIFO, $q_i^t = 1$ if player i arrived before all other A-squares in A^t, $q_i^t = 2$ if player i arrived second among all other A-squares in A^t, and so on.

Definition 1. *A strategy σ_i is a* ***stationary and deterministic strategy (SD-strategy)*** *for an A-square i, if there exists $\psi_i^A : \{(\mathbf{s}, q) \in \mathbb{Z}_+^5\} \to \{\alpha, \beta\}$ such that, for any t such that $i \in A^t$ and $\theta_i^t = (\mathbf{s}^t, q_i^t)$ player i demands $\psi_i^A(\mathbf{s}^t, q_i^t)$.*

We similarly define SD-strategies for B-squares, α-rounds, and β-rounds. A *symmetric, stationary, and deterministic strategy profile* (which we name *stationary* strategy profile*) is a set of SD-strategies, such that all players of the same type use the same strategy, i.e. $\psi_i^x = \psi^x$ for all t, $i \in x^t$, and $x = A, B, \alpha, \beta$. We denote a *stationary* strategy profile* by $\Psi = (\psi^A, \psi^B, \psi^\alpha, \psi^\beta)$.

Definition 2. *A stationary* strategy-profile Ψ is a* ***stationary* equilibrium*** *if it is a subgame perfect equilibrium of the dynamic matching game.*[28]

For simplicity, we assume from now on a symmetric setting (some results pertaining to asymmetric settings appear in the Online Appendix)–that is, we assume:

$$U_A(\alpha) - U_A(\beta) = U_\alpha(A) - U_\alpha(B) \text{ and } U_B(\alpha) - U_B(\beta) = U_\beta(A) - U_\beta(B).$$

Last, we assume that the environment is regular in that

$$p(U_A(\alpha) - U_A(\beta)) \neq kc$$

for all natural numbers $k \in \mathbb{N}$. Regularity assures that neither squares nor rounds are ever indifferent between waiting in queue and matching immediately with an available partner.[29]

4.3 Equilibrium Characterization with the FIFO Protocol

In this section, we present necessary conditions for a stationary* equilibrium that are sufficient to compute the equilibrium welfare. We guarantee the existence of stationary* equilibria and provide their characterization in the Appendix.

By construction, at the beginning of a period, the queue cannot entail both A-squares and α-rounds. As before, we denote the (signed) length of the A-α queue after an arrival of a new pair by $s_{A\alpha} \equiv s_A - s_\alpha$, and after agents form matches by $k_{A\alpha} \equiv k_A - k_\alpha$. We first

[28] Note that in a stationary* equilibrium we allow a player's deviation to be any dynamic strategy, including history-dependent and random ones.

[29] The assumption of regularity is not crucial and similar analysis follows without regularity for any arbitrary tie-breaking rule. However, the presentation is far simpler for regular environments.

consider the decisions of an A-square (analogous analysis holds for an α-round). Let us first consider the case in which *an A-square arrives at the market, and an α-round is available* (one that had either been waiting in the queue or that has just arrived at the market as well). In this case, an A-square is matched immediately to an α-round, the identities of whom are prescribed by the order of arrival. In particular, if the arriving A-square is the first in line, that square is matched to an α-round. If there are A-squares already in queue, this implies that the available α-round arrived at the same time of our A-square, and that α-round will be matched with the first A-square in the queue. In that case, the newly arrived A-square has a choice of whether to wait in line or match with a β-round. However, notice that this square's decision is equivalent to the last A-square who had arrived and decided to wait. Therefore, in a stationary* equilibrium, the new A-square decides to wait and the queues remain as they were.

Suppose now that *an A-square enters the market and an α-round is not available.* This implies that there is at least one β-round available (that the square can match with). Therefore, the A-square has to decide whether to match immediately with the β-round or to wait in line based on the number of A-squares already waiting. An immediate match with a β-round delivers $U_A(\beta)$, whereas waiting in line till matching with an α-round eventually delivers $U_A(\alpha)$ at an uncertain cost of waiting.

Note that once an A-square decides to wait in the queue (rather than match immediately with a β-round), she will wait until matching with an α-round (rather than leave the queue by matching with a β-round at a later point). Indeed, as matches form on a FIFO basis, her position in the queue moves up over time, and the expected time until matching with an α-round becomes shorter. That is, if it is optimal for her to wait in the queue upon entry, it is optimal for her to wait at any later period. The expected waiting time till a match with an α-round is therefore solely determined by the number of other A-squares who precede the square in the queue. The following result identifies bounds on the size of the A-α queue:

Lemma 1 (FIFO Thresholds) *In all stationary* equilibria under FIFO, in all periods,*

$-\bar{k}^{fifo} \leq k_{A\alpha} \leq \bar{k}^{fifo}$ *where*[30]

$$\bar{k}^{fifo} \equiv \left\lfloor \frac{p(U_A(\alpha) - U_A(\beta))}{c} \right\rfloor = \left\lfloor \frac{p(U_\alpha(A) - U_\alpha(B))}{c} \right\rfloor . \tag{4}$$

Intuitively, the time till an α-round enters the market is distributed geometrically (with parameter p), so that the expected time till an α-round arrives at the market is $\frac{1}{p}$. An A-square which is k-th in line in the queue will be matched when the k-th α-rounds arrive, which is expected to occur in $\frac{k}{p}$ periods. The expected waiting costs are therefore $\frac{kc}{p}$, which generate an increase in match utility of $U_A(\alpha) - U_A(\beta)$ (relative to matching with a β-round immediately). An A-square will wait as long as the expected benefit of waiting exceeds its costs, i.e., whenever

$$\frac{kc}{p} < U_A(\alpha) - U_A(\beta),$$

which is the comparison underlying the maximal size of the queue described in Lemma 1. Our regularity assumption further guarantees that an A-square or an α-round is never indifferent between waiting in line and matching immediately. Whenever there are fewer than $\bar{k}^{fifo}$ A-squares in the queue, a new A-square will wait in the market. Whenever there are $\bar{k}^{fifo}$ or more A-squares in the queue, the new A-square prefers to match with a β-round immediately. An analogous description holds for α-rounds and our symmetry assumptions assure that the maximal queue length is identical for A-squares and α-rounds.

We now turn to the decisions of B-squares and β-rounds. A β-round (similarly, a B-square) may decide to wait voluntarily hoping to match with an A-square who will become available when the line for A-squares exceeds $\bar{k}^{fifo}$. In principle, there are two effects at work. The first is similar to that experienced by the A-squares waiting in line: the longer the queue of β-rounds waiting ahead in line, the longer the new β-round has to wait. The second effect, however, is due to A-squares' behavior in equilibrium: the longer is the queue, the closer A-squares are to reach the threshold $\bar{k}^{fifo}$ and to start accepting matches with β-rounds.

However, at least the last β-round in the queue has an incentive to match immediately with any square. To gain intuition, consider the first β-round, say player i, arriving at the market. There cannot be other A-squares waiting in the market since any such squares would have

[30] In the Appendix, we show that in all stationary* equilibria the full support of the A-α queue is $\{-\bar{k}^{fifo}, .., \bar{k}^{fifo}\}$. Therefore the bounds described in this lemma are achieved in equilibrium. Moreover, we show that in equilibrium an A-square (α-round) never matches with a β-round (B-square) if the A-α queue is strictly below $\bar{k}^{fifo}$ (above $-\bar{k}^{fifo}$).

arrived with β-rounds, in contradiction to our β-round being the first in line. Suppose that player i arrives with a B-square. We are going to show that the expected payoff for player i is at most $U_\beta(B)$, which implies that i is (weakly) better off by matching immediately. Observe that by Lemma 1 the first $\bar{k}^{fifo}$ A-squares wait in line until they are matched with an α-round. Thus, player i has to wait for the arrival of at least $\bar{k}^{fifo}+1$ A-squares to match with an A-square. Given this observation, consider now the following hypothetical optimal stopping problem for player i: in each period, player i can choose between matching with a B-square, or wait for the $(\bar{k}^{fifo}+1)$-th arriving A-square, which we assume to be available for player i. Because of this assumption, the expected payoff for player i from the hypothetical optimal stopping problem is (weakly) higher than the player's expected payoff from the dynamic matching game. It is easy to see that the expected payoff from the stopping problem is $U_\beta(B)$. Indeed, the expected costs of waiting until the first available A-square is $\frac{(\bar{k}^{fifo}+1)c}{p}$, which is strictly greater than the benefit from waiting as we have

$$U_\beta(A) - U_\beta(B) < U_\alpha(A) - U_\alpha(B) = U_A(\alpha) - U_A(\beta) < \frac{\left(\bar{k}^{fifo}+1\right)c}{p}.$$

In the Appendix, we show that this intuition yields the following result.

Lemma 2 (Equilibrium under FIFO) *There exists a stationary* equilibrium such that there are never both a B-square and a β-round waiting in the market.*

Lemma 2 implies that there exists a stationary* equilibrium that follows a protocol similar to that implemented by the optimal mechanism and depicted in Figure 1, though the thresholds governing them may be different. Whenever an A-square and an α-round, or a B-square and a β-round, enter the market together, a match of either (A,α) or (B,β) is generated immediately, and the lengths of the queues do not change in that period. The length of the queues can increase or decrease only upon the arrival of an (A,β) pair or a (B,α) pair. As long as the queue (of either A-squares or α-rounds) is strictly shorter than its maximum of $\bar{k}^{fifo}$, the arrival of a new incongruent pair can either increase or decrease the length of the queue by precisely one.

Note that Lemma 1 and its discussion in the Appendix guarantee that since the behavior of A-squares and α-rounds is the same in all stationary* equilibria, so is the welfare generated by matches involving A-squares and α-rounds. Therefore, the stationary* equilibrium described by Lemma 2, in which B-squares and a β-rounds do not delay their matches to each other, is the one that maximizes welfare, as stated in the following corollary.

Corollary 2 *A stationary* * *equilibrium such that there are never both a B-square and a β-round waiting in the market is welfare-maximizing among all stationary* * *equilibria under FIFO.*

4.4 Steady State of Discretionary Matching

As for the optimal mechanism, the length of the A-α queue $k_{A\alpha}$ in the equilibrium described in Lemma 2 and Corollary 2 is characterized by a Markov chain. The transition matrix follows the description in Section 3.1. Formally, denote by

$$x^t \equiv (x^t_{-\bar{k}^{fifo}}, x^t_{-\bar{k}^{fifo}+1}, ..., x^t_{\bar{k}^{fifo}-1}, x^t_{\bar{k}^{fifo}})^{tr} \in \{0,1\}^{2\bar{k}^{fifo}+1}$$

the state of the market in period t – that is, x^t_i is an indicator that takes the value of 1 if the (signed) queue length is precisely i and 0 otherwise. Then,

$$\mathbf{x}^{t+1} = \mathbf{T}_{\bar{k}^{fifo}}\mathbf{x}^t,$$

where $T_{\bar{k}^{fifo}}$ is that described in (2). Since this Markov chain is ergodic, it is characterized by a unique stationary distribution, which is uniform across the $2\bar{k}^{fifo}+1$ states.

The following proposition provides the characterization of the equilibrium steady state of the discretionary process under the FIFO protocol.

Proposition 3 (Discretionary Steady State) *The welfare-maximizing stationary* * *equilibrium under FIFO is associated with a unique steady state distribution over queue lengths, such that the length of the A-α queue $k_{A\alpha} = k_A - k_\alpha$ is uniformly distributed over $\{-\bar{k}^{fifo}, \ldots, \bar{k}^{fifo}\}$, and in any period, the queues contain either*

1. $k_{A\alpha}$ A-squares and β-rounds, and no α-rounds and B-squares, or
2. $|k_{A\alpha}|$ α-rounds and B-squares, and no A-squares and β-rounds.

The threshold $\bar{k}^{fifo}$ is determined by the decisions of A-squares and α-rounds to wait, as specified in Lemma 1. The crucial difference between the discretionary and optimal mechanism is the threshold placed on the maximal stock of A-squares or α-rounds in waiting. Notice that a decision to wait in the market by, say, a square imposes a negative externality on succeeding squares, as it potentially affects their waiting time, and possibly the quality of their matches.[31] On the other hand, a decision to wait may impose a positive externality on future desirable agents on the other side of the market, who would face a ready desirable agent upon arrival. As it turns out, on net, externalities are negative and agents wait "too much." Recalling that we denoted by $\bar{k}^{opt}$ the threshold characterizing the optimal mechanism (specified in Proposition 2), we have the following corollary:

Corollary 3 (Thresholds' Comparison) *Maximal waiting queues are longer under FIFO than they are under the optimal mechanism. That is, $\bar{k}^{opt} \leq \bar{k}^{fifo}$, with strict inequality for sufficiently small waiting costs c.*

To gather some intuition on Corollary 3, consider the case in which in equilibrium all incongruent pairs match immediately ($\bar{k}^{fifo} = 0$), which happens when $\frac{c}{p} > U_A(\alpha) - U_A(\beta)$. Consider now an incongruent pair (A, β) arriving at a discretionary market when there are no α-rounds available, and let us evaluate the externalities generated by a decision of the A-square to wait rather than to match immediately with a β-round, when everybody else is matching instantaneously. This decision affects negatively that β-round, who will have to wait until the next α-round arrives to the market, at a total waiting cost of $\frac{c}{p}$. The event of the next α-round arriving with a B-square (which happens with probability $1-p$) generates additional externalities. Specifically, the β-round also loses the potential match surplus $U_\beta(A) - U_\beta(B)$ and, on the positive side, the next α-round arriving at the market realizes a match surplus of $U_\alpha(A) - U_\alpha(B)$. Under symmetry, such positive externality is lower than the cost of waiting incurred by the β-round since $\frac{c}{p} > U_A(\alpha) - U_A(\beta) > (1-p)(U_A(\alpha) - U_A(\beta))$.[32] Therefore, the social planner would not make the A-square wait and would instead match him immediately. Corollary 3 guarantees that this intuition generalizes to any optimal threshold $\bar{k}^{opt}$.

[31] From a welfare perspective, the externality on the quality of the match is of less importance. As long as a social planner views identical agents as interchangeable, an immediate mismatch or a later mismatch have similar welfare consequences.

[32] The A-square decision to wait generates further negative externalities for both a potential A-square and a potential B-square who arrive at the market with the first α-round.

4.5 Welfare

Since the protocols are similar except for the queues' thresholds, the expected per-period welfare in the steady state characterized in Proposition 3 can be found using an analogous derivation to that carried out for the optimal mechanism. This derivation leads to an expression mirroring equation (3), accounting for the discretionary process' threshold $\bar{k}^{fifo}$. That is, the expected per-period net welfare is given by:

$$W^{fifo}(c) = S_\infty - \frac{p(1-p)U}{2\bar{k}^{fifo}+1} - \frac{2\bar{k}^{fifo}(\bar{k}^{fifo}+1)c}{2\bar{k}^{fifo}+1},$$

where $\bar{k}^{fifo}$ is defined in (4). To summarize, we have the following corollary:

Corollary 4 (Decentralized Welfare) *The maximum equilibrium welfare under FIFO is given by* $W^{fifo}(c) = S_\infty - \Psi(c)$, *where* $\lim_{c\to 0}\Psi(c) = p\left(U_A(\alpha) - U_A(\beta)\right)$, *and* $\Psi(c) = p(1-p)U$ *for all* $c \geq p\left(U_A(\alpha) - U_A(\beta)\right)$.

Recall Corollary 1, which characterized the welfare under the optimal mechanism. By definition, the welfare generated under the optimal mechanism is higher than that generated by the discretionary process, so that $\Theta(c) \leq \Psi(c)$ for all c. While the optimal mechanism generates welfare that is decreasing in waiting costs, this is not necessarily the case under the discretionary process. Furthermore, while the welfare under the optimal mechanism approaches S_∞ as waiting costs diminish, this is not the case under the discretionary process. As waiting costs become very small, there is a race between two forces. For any given threshold, the overall waiting costs decline. However, in equilibrium, discretionary thresholds increase, leading to greater expected wait times. As it turns out, the balance between these two forces generates significant welfare losses, given by $p\left(U_A(\alpha) - U_A(\beta)\right)$, even for vanishingly small costs. The next section provides a detailed comparison of the two procedures in terms of welfare.

5 Welfare Comparisons

By construction, the optimal mechanism generates welfare that is at least as high as that generated by the discretionary process.[33] In this section we inspect how the welfare wedge

[33] It is possible to show that the optimal mechanism represent a Pareto improvement with respect to the decentralized setting. In fact, it is easy to see that B-squares and β-rounds are better off under the optimal

responds to the underlying parameters of the environment, suggesting the settings in which centralized intervention might be particularly useful.

The following proposition captures the effect on the welfare wedge $W^{opt}(c) - W^{fifo}(c)$ of both waiting costs c, the frequency p of A-squares or α-rounds, and the utility benefit for an α-round from matching with an A-square rather than a B-square (equivalently, the utility benefit for an A-square from matching with an α-round rather than a β-round).

Proposition 6 (Welfare Wedge – Comparative Statics)

1. *For any interval* $[\underline{c}, \bar{c})$, *where* $\underline{c} > 0$, *there is a partition* $\{[c_i, c_{i+1})\}_{i=1}^{M-1}$, *where* $\underline{c} = c_1 < c_2 < ... < c_M = \bar{c}$, *such that* $W^{opt}(c) - W^{fifo}(c)$ *is continuous and increasing over* (c_i, c_{i+1}) *and*
$$W^{opt}(c_i) - W^{fifo}(c_i) > W^{opt}(c_{i+1}) - W^{fifo}(c_{i+1})$$
for all $i = 1, ..., M-1$.

2. *As* c *becomes vanishingly small, the welfare gap* $W^{opt}(c) - W^{fifo}(c)$ *converges to a value that is increasing in* $p \in [0, 1)$ *and in* $U_A(\alpha) - U_A(\beta)$.

To see the intuition for the comparative statics corresponding to waiting costs, notice that an increase in costs has two effects on the welfare gap. Since the equilibrium threshold under the discretionary process is greater than the optimal threshold (Corollary 3), an increase in waiting costs has a direct effect of magnifying the welfare gap. Nonetheless, there is also an indirect effect of an increase in waiting costs that arises from the potential changes in the induced thresholds. Consider a slight increase in waiting costs such that the optimal threshold does not change, but the discretionary threshold decreases. The discretionary process is then "closer" to the optimal process – both the matching surplus and the waiting costs are closer and the welfare gap decreases. In fact, as costs become prohibitively high, both processes lead to instantaneous matches and identical welfare levels. As we show in the proof of Proposition 4, the indirect effect overwhelms the direct effect at precisely such transition points and acts to shrink the welfare gap. The construction of the partition is done as follows. Each atom $[c_i, c_{i+1})$ of the partition corresponds to constant thresholds under the discretionary

mechanism (as it implies better matches and shorter waiting times). Moreover, one can show that, as long as $\bar{k}^{opt} \geq 2$, the expected payoff of A-squares and α-rounds as they enter the market is higher under the optimal mechanism than under decentralization as well.

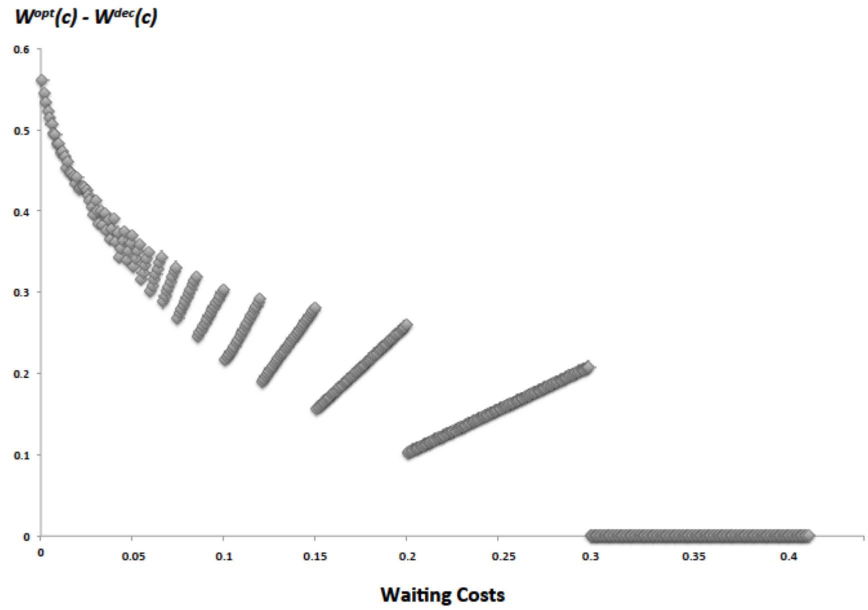

Figure 2: Welfare Gap between Optimal and Discretionary Matching as a Function of Costs

process. Over these intervals, only the direct effect operates and the welfare gap is increasing. Each of the endpoints $\{c_i\}_i$ corresponds to a decrease of the discretionary threshold by one. That is, the discretionary threshold is not constant over $[c_i, c_{i+2})$ for all i. Therefore, when comparing two such endpoints, the indirect effect kicks in and the decreasing trend of the welfare gap emerges. Figure 2 depicts the resulting pattern the welfare gap exhibits for $U_A(\alpha) = U_\alpha(A) = 3$, $U_A(\beta) = U_\alpha(B) = U_B(\alpha) = U_\beta(A) = 1$, $U_B(\beta) = U_\beta(B) = 0$, and $p = 0.3$. As suggested by the proposition, $W^{opt}(c) - W^{fifo}(c)$ is piece-wise increasing in c. Nevertheless, overall, the gap has a decreasing trend.

To glean some intuition on the comparative statics the welfare gap displays with respect to p, consider two type distributions governed by p_1 and p_2 such that $p_1 < p_2 = mp_1$, $m > 1$. Suppose further that costs are so low that the match surplus in the discretionary process is very close to the optimum, S_∞. The individual incentives to wait for A-squares and α-rounds are higher under p_2 than under p_1. In fact, in the discretionary setting, the distribution of steady state queue length is the uniform distribution where, from (4), under p_2, roughly $1 - 1/m$ of the probability mass is allocated to queue lengths larger than those realized under p_1. For each of these large steady state queue lengths, we have more pairs of agents waiting, i.e., increased per-period waiting costs. The optimal mechanism internalizes the negative externalities, so the effect of the increased waiting costs is weaker. On the other hand, the benefit of this increase in queue length is a lower chance of producing mismatches. However, for sufficiently low c, the match surplus under p_1 is already close to its optimum of S_∞ and this effect is weak; in particular, the difference in terms of match surplus that the optimal and discretionary processes generate is similar under p_1 and p_2. Therefore, for sufficiently low c, the dominant effect is the one produced by the differential in expected waiting costs, which generates our comparative statics.[34] Notice that if $p = 0$ or $p = 1$, both the optimal mechanism and the discretionary processes generate the same welfare.

Going back to our assumption of super-modular preferences, note that the construction of the optimal mechanism would remain essentially identical were preferences sub-modular (with an appropriate relabeling of market participants). However, in the discretionary setting, sub-modular preferences would lead to a negative welfare effect compounding the negative externalities present in our setup. Namely, individual incentives would be misaligned with market-wide ones. In that respect, our comparison of optimal and discretionary processes assuming super-modular preferences is a conservative one.

Similarly, considering waiting costs that differ across the two sides of the market would lead to a greater welfare wedge as well. Intuitively, suppose that squares experience a waiting cost of c_S and rounds experience a waiting cost of c_R, where $c_S > c_R$, with an average cost of $c = (c_S + c_R)/2$. The optimal mechanism with asymmetric costs would coincide with that corresponding to identical costs of c since per-pair costs are the same in both cases. In the discretionary process, A-squares would be willing to wait when the queue of A-squares is no longer than $\bar{k}^{fifo}$ and α-rounds would be willing to wait when the queue of α-rounds is no longer than $\bar{k}^{fifo}$, where

$$\bar{k}^{fifo} = \left\lfloor \frac{p\left(U_A(\alpha) - U_A(\beta)\right)}{c_S} \right\rfloor \quad \text{and} \quad \bar{k}^{fifo} = \left\lfloor \frac{p\left(U_\alpha(A) - U_\alpha(B)\right)}{c_R} \right\rfloor .$$

Suppose $\frac{p(U_A(\alpha)-U_A(\beta))}{c_x} \in \mathbb{N}$ for $x = S, R$ to avoid rounding issues. From convexity, it follows that the threshold $\bar{k}^{fifo}$ corresponding to identical costs of c satisfies $\bar{k}^{fifo} \leq (\bar{k}^{fifo} + \bar{k}^{fifo})/2$. Therefore, the excessive waiting discretionary processes exhibit would be even more pronounced when costs are asymmetric across market sides.[35]

[34] In fact, we can show that for any $\Delta p > 0$, there exists $\delta > 0$ such that for every $c < \delta$ and $p \in [0, 1 - \Delta p]$, we have that the welfare wedge under $p + \Delta p$ and c is greater than under p and c. Furthermore, $\delta \to 0$ as $\Delta p \to 0$.

[35] In fact, in such a setting, β-rounds may wait in the market even when B-squares are present, leading to

6 Transfers and Alternative Protocols

So far, we have shown that intervention in dynamic matching markets can have a substantial impact on welfare, at least when centralization is carried out using the optimal dynamic mechanism. However, the full-fledged optimal mechanism may be hard to implement. It requires that the formation of matches, even those of individuals who would prefer to wait in the market, be within the purview of the centralized planner. It also requires the central planner to monitor the market continuously to determine when matches should be formed, which may be administratively costly. In this section (and in the Online Appendix) we show that improvements to discretionary settings under FIFO can be achieved by mechanisms that relax one of these two requirements. We start by showing that a tax scheme imposed on those who wait in the market can yield the optimal waiting patterns, if tailored appropriately. We then analyze an alternative setting that does not require the clearinghouse to monitor the market continuously and can provide a substantial welfare improvements over the discretionary matching process under FIFO: a discretionary environment in which matches are formed following a last-in-first-out (LIFO) protocol; in addition, in the Online Appendix, we analyze two alternative protocols, one in which the centralized clearinghouse matches all available agents every fixed number of periods, and the second in which, in a discretionary market, agents' priorities at every period are determined randomly.

6.1 Optimal Taxation

Certainly, a fixed per-period tax on waiting $\tau \geq 0$ can be set so that the resulting de-facto cost of waiting, $c+\tau$, is such that the discretionary process generates the optimal threshold (with costs c).[36] Namely, τ can be set so that

$$\frac{p\left(U_A(\alpha)-U_A(\beta)\right)}{c+\tau}=\sqrt{\frac{p(1-p)U}{2c}}. \tag{5}$$

In fact, if collected taxes in period t are given to new entrants in period $t+1$, they would have no effects on either overall welfare or individual incentives to wait once in the market. Nonetheless, there is a risk that such a policy would introduce a strong incentive for agents to enter the market to begin with, only to gather the previous generation's taxes. As it turns

another channel of inefficient waiting. We show formally how our main results carry over to this more general environment in the Online Appendix.

[36] This is always possible from Corollary 3.

out, there is a tax scheme that is budget balanced such that the expected tax (or subsidy) for an entering agent who is not privy to the pattern of queues in place is nil. Under such a scheme, no agent is tempted to enter the market only for the sake of reaping the benefits of taxes. Indeed, consider a linear tax scheme – agents who are k-th in line pay $\tau^* k$ to the matching institution, regardless of whether they are a square or a round and regardless of their type. To achieve budget balance, the collected taxes are equally redistributed back to the existing agents in the market in each period. When the length of queue is $\hat{k}$, the resulting net tax (that is added to the fixed cost c) for an agent who is k-th in line is

$$\tau^* k - \frac{2 \cdot \sum_{k=1}^{\hat{k}} \tau^* k}{2\hat{k}} = \tau^* \left(k - \frac{\hat{k}+1}{2} \right).$$

In particular, the net added tax on waiting for the last agent in the queue is $\frac{\tau^*(\hat{k}-1)}{2}$, which is increasing in the queue's length $\hat{k}$. Thus, using the definition of τ from equation (5), we want the tax levied on the last agents in the queue of length $\bar{k}^{opt}$ to satisfy

$$\frac{\tau^*(\bar{k}^{opt} - 1)}{2} = \tau \quad \Longleftrightarrow \quad \tau^* = \frac{2\tau}{\bar{k}^{opt} - 1}.$$

Such a taxation policy might still be difficult to administer in terms of the financial activity it would entail – time-dependent taxes and redistribution of resources at each period. Without taxation, the optimal mechanism may be viewed as too complex administratively as well. Indeed, optimality dictates matchings that occur at points in time that are not perfectly predictable at the outset of the process and the clearinghouse needs to monitor the queues of individuals continuously. In the next subsection we offer the analysis of a simple mechanism that, while not optimal, can generate substantial welfare improvements relative to the discretionary process.

6.2 Last-In-First-Out

6.2.1 Equilibrium Characterization under LIFO

We now consider a discretionary setting that has the same structure described in Section 4.2 but, once every agent on the market has specified their demands, matches form according to the *last-in-first-out* (LIFO) protocol. This protocol assigns a linear order $\succ$ over, say, A^t such that

$$\forall s^{t'}, s^{t''} \in A^t, \quad s^{t''} \succ s^{t'} \iff t' < t'' \leq t.$$

We first consider the decisions of A-squares (and we omit an analogous discussion for α-rounds). If an A-square finds an α-round upon arrival, the A-square matches with the last arrived α-round. If no α-round is available, the A-square needs to decide either to match with the last β-round, who must have just arrived together with the A-square, or to wait in the queue. Under LIFO, this decision is independent of other A-squares who have been waiting in the queue. Rather, the decision depends on the anticipated behavior of A-squares who will arrive at the market in the future periods.

Let us consider a SD-strategy ψ_A with a threshold $\bar{k}_A$ for A-squares. At any period, if no α-round is available, an A-square, say player i, waits by demanding α as long as the rank q_i according to LIFO is at most $\bar{k}_A$: i.e., when there are less than $\bar{k}_A$ other A-squares who arrived *after* player i. To gain intuition, suppose that all A-squares (including player i) use the threshold $\bar{k}_A = 1$. If player i finds no available α-round upon arrival, then she waits by demanding α. In the next period, player i continues to wait if either a pair (A, α) or a pair (B, β) arrive, because in these scenarios her rank according to LIFO remains the same: a new pair (A, α) will be matched to each other, and the queue of A-square is unaffected by an arrival of (B, β). However, if an (A, β) pair arrives, the new A-square, who also plays uses $\bar{k}_A = 1$, demands α. According to ψ_A, player i demands β and leaves the market. Finally, if a (B, α) pair arrives, player i matches to the α-round. To summarize, player i exits the market matched with either an α-round or a β-round (with probability 1/2 each) as soon as the first incongruent pair arrives. Since the expected number of periods until the first arrival of an incongruent pair is $\frac{1}{2p(1-p)}$, the expected payoff for player i is

$$\frac{U_A(\alpha) + U_A(\beta)}{2} - \frac{c}{2p(1-p)}.$$

Consider a possible deviation for player i, which requires i to demand β-rounds when she finds no available α-round upon her arrival. This deviation is not strictly profitable if and only if

$$U_A(\beta) \leq \frac{U_A(\alpha) + U_A(\beta)}{2} - \frac{c}{2p(1-p)},$$

which we can rewrite as

$$\frac{p(1-p)\left(U_A(\alpha) - U_A(\beta)\right)}{c} \geq 1 = \frac{\bar{k}_A(\bar{k}_A + 1)}{2}. \tag{6}$$

Let us consider another potential deviation by player i: if one more A-square arrives after player i and no α-round is available, player i, instead of demanding β, may consider to increase her threshold to $\bar{k}'_A = 2$ and remain in the market.[37] If player i uses the threshold $\bar{k}'_A = 2$, while all other A-squares use $\bar{k}_A = 1$, player i will match to an α-round for sure. We use an *Absorbing Markov Chain* to compute the expected continuation payoff for player i. The *event time* τ increases upon arrival of an incongruent pair. The state space is $\{1, 2, \alpha\}$: the two *transient states* (1 and 2) denote the A-square's rank, and the *absorbing state* (α) denotes player i's matching to an α-round. The matrix of transition probabilities p_{ij} from state i to state j is

$$P = \begin{bmatrix} Q & R \\ \mathbf{0} & 1 \end{bmatrix}, \quad \text{where} \quad Q = \begin{bmatrix} 0 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \quad \text{and} \quad R = \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}.$$

The matrix Q represents transition probabilities between transient states.[38] Let

$$T \equiv (I_2 - Q)^{-1} \cdot 1 = 4 \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}.$$

If the initial state of the absorbing Markov chain is 2, it is well-known in the absorbing Markov chains literature that the chain will be absorbed by state α within $T_2 = 6$ expected periods of event time. Therefore, if player i deviates by increasing her threshold perpetually to $\bar{k}'_A = 2$, the expected continuation payoffs is $U_A(\alpha) - \frac{6c}{2p(1-p)}$. Such deviation is not strictly profitable if

$$U_A(\beta) \geq U_A(\alpha) - \frac{6c}{2p(1-p)},$$

which is equivalent to

$$\frac{p(1-p)\,(U_A(\alpha) - U_A(\beta))}{c} \leq 3 = \frac{(\bar{k}+1)(\bar{k}+2)}{2}. \tag{7}$$

In the Online Appendix, we show that (6) and (7) generalize to an arbitrary threshold $\bar{k}_A \in \mathbb{Z}_+$ as follows

$$\frac{\bar{k}_A(\bar{k}_A+1)}{2} \leq \frac{p(1-p)(U_A(\alpha) - U_A(\beta))}{c} \leq \frac{(\bar{k}_A+1)(\bar{k}_A+2)}{2}.$$

In particular, these inequalities are necessary conditions for any SD-strategy with threshold $\bar{k}_A$ to be a part of an equilibrium: the first inequality guarantees that an A-square with rank

[37] In particular, we are considering a perpetual deviation from ψ_A to a SD-strategy with threshold $\bar{k}' = 2$, rather than a one-shot deviation.

[38] Take any event time τ, and suppose that the state at time τ is 2: i.e., there is an another A-square waiting, who arrived after player i. The event time τ progresses to $\tau+1$ by an arrival of an incongruent pair. If the incongruent pair is (B, α), the rank of player i moves up to 1. This transition incurs with probability $Q_{21} = 1/2$. Otherwise, the new incongruent pair is (A, β). According to ψ_A, the A-square who has been waiting with player i demands β and leaves the market, leaving player i's rank at 2. This transition incurs with probability $Q_{22} = 1/2$. If the state in period τ is 1, and (B, α) arrives, then player i matches to the α-round. This last transition incurs with probability $R_{1\alpha} = 1/2$.

$\bar{k}_A$ would not deviate by demanding β when no α-rounds are available in the market, and the second inequality guarantees that an A-square, whose ranking would become $\bar{k}_A + 1$ by waiting, does not prefer to do so rather than matching with a β-round.

Similarly to the FIFO environment, we assume that the environment is regular in that

$$p(1-p)(U_A(\alpha) - U_A(\beta)) \neq kc$$

for all natural numbers $k \in \mathbb{N}$.

The following results characterize stationary* equilibria under LIFO and allow us to compare the welfare achieved in this setting to both the centralized mechanism and discretionary settings governed by the FIFO protocol.

Lemma 3 (Thresholds under LIFO) *In all stationary* equilibria under LIFO in which A-squares and α-rounds use threshold strategies, in all periods, $-\bar{k}^{lifo} \leq k_{A\alpha} \leq \bar{k}^{lifo}$ where*[39]

$$\bar{k}^{lifo} \equiv \left\lfloor \sqrt{\frac{2p(1-p)(U_A(\alpha) - U_A(\beta))}{c} + \frac{1}{4}} - \frac{1}{2} \right\rfloor .$$

Let us turn to the decisions of β-rounds (and omit a similar analysis for B-squares). Suppose that A-squares play a SD-strategy with threshold $\bar{k}_A$. Only when an (A, β) pair arrives, there may be an A-square who may demand β, and she matches the latest arriving β-round. Thus, if a β-round remains unmatched upon arrival, he will never be matched with an A-square later. It follows that every β-round has an incentive to leave the market as soon as possible. The following result guarantees the existence of a stationary* equilibrium such that A-squares and α-rounds use a threshold strategy identified by $\bar{k}^{lifo}$, and B-squares and β-rounds always match immediately whenever it is possible. Similarly to the FIFO case, such equilibrium is welfare-maximizing among all stationary* equilibria.

Lemma 4 (Equilibrium under LIFO) *There exists a stationary* equilibrium in which A-squares and α-rounds use a threshold $\bar{k}^{lifo}$ and such that there can never be both B-squares and β-rounds waiting in the market. Such equilibrium is welfare-maximizing among all stationary* equilibria.*

[39] In the Appendix, we show that in all stationary* equilibria the full support of the A-α queue is $\{-\bar{k}^{lifo}, .., \bar{k}^{lifo}\}$. Therefore the bounds described in this lemma are achieved in equilibrium.

6.2.2 Steady State and Welfare under LIFO

The following proposition provides the characterization of the equilibrium steady state of the discretionary process under the LIFO protocol.

Proposition 5 (Discretionary Steady State under LIFO) *The welfare-maximizing stationary* equilibrium under LIFO is associated with a unique steady state distribution over queue lengths, such that the length of the A-α queue $k_{A\alpha} = k_A - k_\alpha$ is uniformly distributed over $\{-\bar{k}^{lifo}, \ldots, \bar{k}^{lifo}\}$, and in any period, the queues contain either*

1. *$k_{A\alpha}$ A-squares and β-rounds, and no α-rounds and B-squares, or*
2. *$|k_{A\alpha}|$ α-rounds and B-squares, and no A-squares and β-rounds.*

The threshold $\bar{k}^{lifo}$ is determined by the decisions of A-squares and α-rounds to wait, as specified in Propositions 7 and 8. Recalling that we denoted by $\bar{k}^{opt}$ the threshold characterizing the optimal mechanism (described in Proposition 2), we have the following corollary:

Corollary 5 (Thresholds and Welfare Comparisons under LIFO)

1. *For sufficiently small waiting costs c, the maximal waiting queues under LIFO are longer than under the optimal mechanism, but shorter than under FIFO–that is, $\bar{k}^{opt} < \bar{k}^{lifo} < \bar{k}^{fifo}$.*
2. *The LIFO protocol is asymptotically efficient–that is, the maximum equilibrium welfare under LIFO is given by $W^{lifo}(c) = S_\infty - \Gamma(c)$, where $\lim_{c \to 0} \Gamma(c) = 0$.*

Corollary 5 suggests that the LIFO protocol could represent a substantial improvement with respect to the FIFO protocol in discretionary settings. However, the LIFO protocol has well-known implementation problems, as agents may have an incentive to leave the market and re-enter to improve their odds for a shorter wait and a better match. Figure 3 depicts the welfare losses generated by both the FIFO and the LIFO protocols with respect to the optimal mechanism for the parameter values $U_{A\alpha} = 3$, $U_{A\beta} = U_{B\alpha} = 1$, $U_{B\beta} = 0$, and $p_A = p_\alpha = \frac{1}{3}$. The figure suggests that the welfare gap decreases significantly under LIFO.

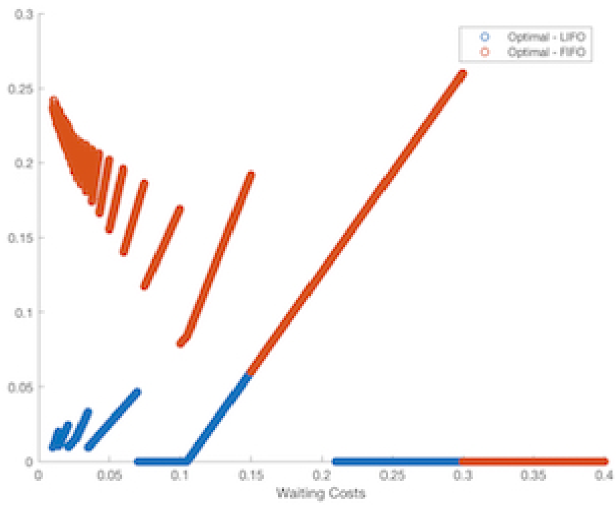

Figure 3: Welfare Gaps between Optimal and Discretionary Matching (FIFO and LIFO) as a Function of Costs

7 Individual Rationality

Throughout the paper we assumed that agents leave the market only after being matched. Agents do not leave the market unmatched, regardless of their expected utility. Certainly, if remaining unmatched generates zero utility, our assumption violates individual rationality as some agents (in particular, B-squares or β-rounds) sometimes stay in the market simply for lack of available agents who will match with them. These agents could consequently earn a negative expected utility even when $U_x(y) > 0$ for all x, y. Assume that any agent prefers to match with any other agent immediately over remaining unmatched. We now provide a bound on the value of remaining unmatched, or the value of an outside option agents have, that assures all the matching protocols we discuss are individually rational. Notice that since, from Corollary 3, the discretionary threshold is higher than the optimal threshold, it suffices to find such a bound that guarantees that the discretionary process is individually rational.

In the discretionary process, A-squares or α-rounds always have the possibility of matching with β-rounds or B-squares instantaneously when they decide to wait in the market. Therefore, when deciding to wait they expect an even greater utility and individual rationality holds for them. Consider now β-rounds (analogously, B-squares). A β-round who is k-th in line can always declare all squares as acceptable in each period. Notice that, by construction, $k \leq \bar{k}^{fifo}$. The time between arrivals of B-squares is distributed geometrically with probability $1-p$. Therefore, with such a strategy, the expected time for the β-round to match with a B-square is at most $k/(1-p)$, yielding a match utility of $U_\beta(B)$. The wait time till matching with a B-square could be even shorter if β-rounds who precede the β-round in question are willing to match only with A-squares. Furthermore, the β-round could end up matching with an A-square before k B-squares arrive at the market. It follows that such a strategy guarantees an expected utility of at least

$$U_\beta(B) - \frac{kc}{1-p} \geq U_\beta(B) - \frac{\bar{k}^{fifo}c}{1-p} \geq U_\beta(B) - \frac{p}{1-p}\left(U_\alpha(A) - U_\alpha(B)\right) \equiv U^{\min}.$$

If β-rounds follow a different strategy in equilibrium, it must be that their utility is at least as high. Therefore, as long as the value of remaining unmatched is lower than $U^{\min}$, individual rationality holds under both the optimal and the discretionary processes (analogous calculations follow for B-squares and, under full symmetry of utilities, the bound corresponding to them is also $U^{\min}$). In addition, since the optimal fixed-window size is smaller than the discretionary equilibrium threshold, the construction above assures that this bound on remaining unmatched assures individual rationality for the fixed-window protocol as well.

8 Conclusions

In this paper we considered a dynamic matching setting and identified the optimal matching mechanism in an environment as such. The optimal mechanism always matches congruent pairs immediately and holds on to a stock of incongruent pairs up to a certain threshold. When matching follows a discretionary process, a similar matching protocol ensues in equilibrium, but the induced thresholds for waiting in the market are larger as individuals do not internalize the net negative externalities they impose on those who follow. This difference generates a potentially significant welfare wedge between discretionary processes and centralized clearinghouses, even when waiting costs are vanishingly small. Our results provide guidance as to the features of the economy that could make centralized intervention more appealing.

We also offer some simple interventions to discretionary markets – transfer schemes that can induce optimal outcomes, and an alternative last-in-first-out priority protocol, which are arguably far less complex than the full-fledged optimal mechanism and can provide substantial welfare improvements relative to the first-in-first-out protocol.

There are several natural other interventions and extensions that we analyze and discuss in the Online Appendix: a simple centralized mechanism that matches individuals at fixed-time intervals, a discretionary setting in which agents are ranked at every period according to a random priority rule, general asymmetric environments, a richer set of participant types, and different arrival processes.

9 Appendix

9.1 Proofs Regarding the Optimal Mechanism

First, we present an auxiliary result (Lemma A1), and then we move on to show, in Theorem A1, that the restriction on SD-mechanisms satisfying Conditions (1) and (2) we imposed in the main text is without loss of generality.

Lemma A1 *(1) For any mechanism μ, there exists a mechanism μ', with $v(\mu') \geq v(\mu)$, that never holds both A-squares and α-rounds if they are available in the market; (2) For any mechanism μ, there exists a mechanism μ', with $v(\mu') \geq v(\mu)$, that never holds more than $\frac{U}{2c}$ squares (and rounds) in the market.*

Proof of Lemma A1 (1) Take any mechanism μ that may hold some (A, α) pairs after some histories. Let us consider another mechanism μ' which creates the same set of matches as μ at every history, except that (i) when μ holds a pair of (A, α), say agents (i, j), μ' matches the pair as soon as they are available, (ii) μ' does not create any match that μ creates involving either i or j, and (iii) if μ matches i to a round $r(\neq j)$ in period $t' > t$, and matches j to a square $s(\neq i)$ in period $t'' > t$, then μ' forms the match (s, r) in period $\max\{t', t''\}$. It is clear from the construction that $v(\mu') \geq v(\mu)$ because μ' involves strictly lower waiting costs and a weakly higher match surplus than μ does for every finite time horizon. We omit an analogous proof that we can further improve μ' by matching (B, β) pairs immediately.

(2) Take a mechanism μ, and assume, without loss of generality from the previous step, that μ never holds an (A, α) or a (B, β) pair in the market. Assume additionally that μ matches squares and rounds on the first-in-first-out basis, in the sense that when μ matches, for example, an A-square to a round, it selects the A-square who arrived first among all available A-squares. The last restriction does not affect $v(\mu)$ because agents of the same type are interchangeable from a welfare perspective. Suppose that μ may hold more than $\frac{U}{2c}$ (A, β) pairs at some histories. We construct another mechanism μ' which creates the same set of

matches as μ at every history, except that (i) μ holds $\{A_i, \beta_i\}_{i=1}^n$ with $n > \frac{U}{2c}$ (where the ranking i is obtained by the time of arrival at the market), μ' holds only $\{A_i, \beta_i\}_{1 \le i \le U/2c}$ and matches all remaining (A, β) pairs immediately, (ii) μ' does not create any match that μ creates with agents in $\{A_i, \beta_i\}_{U/2c < i \le n}$, and (iii) for any A_i and β_i with $\frac{U}{2c} < i \le n$ (i.e., agents held by μ but matched immediately by μ'), if μ matches A_i to a round $r(\neq \beta_i)$ in period t' and β_i to a square $s(\neq A_i)$ in period t'', then μ' matches (s, r) in period $\max\{t', t''\}$.

We claim that $v(\mu') \ge v(\mu)$. To see this, take any pair (A_r, β_r) held by the mechanism μ but matched by the mechanism μ' at some period t (i.e., $r > \frac{U}{2c}$). Recall that μ never holds (A, α) or (B, β) in the market, and it matches agents on the first-in-first-out basis. As only one pair arrives in each period, μ will either (a) match (A_r, β_r) to each other before or at period $t + (\frac{U}{2c})$, or (b) it will match (A_r, β_r) either to each other or to other agents, in some periods after $t + (\frac{U}{2c})$. In both instances (a) and (b), μ generates a lower average welfare than μ' for every finite time horizon. To see this, note that in case (a), μ generates strictly higher waiting costs but the same match surplus, and in case (b) the additional waiting costs generated by μ are strictly higher than $(\frac{U}{2c})(2c)$ and therefore they exceed the highest match surplus gain $U = U_{A\alpha} + U_{B\beta} - U_{A\beta} - U_{B\alpha}$ that can be generated by holding (A_r, β_r) and matching them with other agents. We omit an analogous proof showing that we can further improve μ' by not holding more than $\frac{U}{2c}$ (B, α) pairs on the market. ■

Lemma A1 allows us to simplify our problem as the following Markov decision problem with agents arriving in incongruent pairs, a finite set of states, and a finite set of actions:

$$(MDP, s^0) \equiv \{T, S, s^0, (A_s)_{s \in S}, (r(s,k), p(\cdot|k))_{s \in S, k \in A_s}\},$$

where s^0 denotes a particular initial state. Each component is defined as follows:

1. $T \equiv \{0, 1, 2, \dots\}$ is *the set of decision epochs.* Epochs correspond to times at which an incongruent pair (A, β) or (B, α) arrives. Since the probability of an incongruent pair arriving at any period is $2p(1-p)$, the expected time between epochs is $\frac{1}{2p(1-p)}$.

2. $S \equiv \{z \in \mathbb{Z} : -(U/2c) - 1 \le z \le (U/2c) + 1\}$ is *the set of possible states (or stocks).* Each state $s_{A\alpha} \equiv s_A - s_\alpha \in S$ represents the (signed) number of mismatched pairs of type (A, β) or (B, α) in the market. Since we restrict our attention to mechanisms that do not hold more than $U/2c$ squares (and rounds), a state, which takes a new arriving pair into account, has to belong to the set $\{-(U/2c) - 1, ..., (U/2c) + 1\}$.

3. $s^0 = 0$ is the *the initial state*. Initially, no agent waits.

4. $K \equiv \{z \in \mathbb{Z} : -U/2c \leq z \leq U/2c\}$ is *the set of available actions*. Each $k \in K$ represents the (signed) number of incongruent pairs held in the market from one period to the next.

5. $r(s, k)$ is *the reward function*: for every $s \in S$, $k \in K$,

$$r(s,k) = \begin{cases} (s-k)U_{A\beta} - \frac{kc}{2p(1-p)} & \text{if } s \geq k \geq 0 \\ (|s|-|k|)U_{B\alpha} - \frac{|k|c}{2p(1-p)} & \text{if } s \leq k \leq 0 \\ -\infty & \text{otherwise.} \end{cases}$$

The expected waiting cost incurred to any agent who waits for one epoch is $\frac{c}{2p(1-p)}$. The reward function returns $-\infty$ if an action is infeasible. For all feasible actions, the values of the reward function are in the interval $\left[-\frac{U}{4p(1-p)}, \frac{U}{2c}U_{A\alpha}\right]$.

6. $p(s, k)$ is the *transition probability* that the system is in state $s \in S$ at time $\tau + 1$, after the action k has been chosen at time τ.

$$p(s,k) = \begin{cases} 1/2 & \text{for } s = k-1, k+1, \\ 0 & \text{otherwise.} \end{cases}$$

(MDP, s^0) is stationary in the sense that the reward function $r(s, k)$ and the transition probability function $p(s, k)$ do not depend on time, or epochs, explicitly. A *policy* of (MDP, s^0) is any rule, deterministic or randomized, governing the choice of actions. Such rule may depend on the entire history of the process up to that point. The value of a policy μ is then,

$$v(\mu) \equiv \liminf_{T\to\infty} \frac{1}{T} E_\mu \left[\sum_{\tau=1}^{T} r(s^\tau, k^\tau)\right].$$

An *SD-policy* of (MDP, s^0) applies the same decision rule $\mu^{SD} : S \to K$ at every state in S. The value of μ^{SD} is then

$$v(\mu^{SD}) = \lim_{N\to\infty} \frac{1}{N} E \left[\sum_{\tau=1}^{N} r(s^\tau, \mu^{SD}(s^\tau))\right].$$

Such limit exists, as guaranteed, for example, by Proposition 8.1.1-b in Puterman (2005).

We are now ready to show the next result, which guarantees that restricting the attention to mechanisms satisfying Conditions (1) and (2) in the main text is without loss of generality.

Theorem A1 *There exists an optimal SD-mechanism that satisfies Conditions (1) and (2).*

Proof of Theorem A1 In what follows, we prove that there exists an optimal SD-policy of (MDP, s^0). By Lemma A1 and the discussion above, an optimal SD-policy of (MDP, s^0) defines an optimal SD-mechanism of our matching problem. Let us recall the following result from Ross (2014).

Theorem Ross (1) (V.2.1 in Ross (2014) *If there exists a bounded function $h(s)$, $s \in S$, and a constant g such that*

$$g + h(s) = \max_{k \in K} \left[r(s,k) + \sum_{s' \in S} p(s', k) h(s') \mid s \in S \right] \text{ for all } s \in S, \tag{8}$$

1. there exists an SD-policy μ^* such that

$$g = \nu(\mu^*) = \sup_{\mu} \nu(\mu);$$

2. and μ^* is any SD-policy that, for each $s \in S$, prescribes an action k that maximizes the RHS of (8).

We extend the notion of (MDP, s^0) by allowing the initial state to be an arbitrary $s \in S$, and we denote the Markov decision problem with an arbitrary initial state by (MDP). It is straightforward to extend the definition of a policy and a SD-policy to (MDP). For any $0 < \delta < 1$, initial state $s \in S$, and a policy μ of (MDP), define

$$V_\delta(\mu; s) \equiv E_\mu \left[\sum_{\tau=0}^{\infty} r(s^\tau, k^\tau) \delta^\tau | s \right],$$

where E_μ represents an expectation conditional to policy μ being implemented. For each $s \in S$, $V_\delta(\mu; s)$ represents the expected total discounted return earned when the policy μ of (MDP) is employed. Since the reward function is bounded by the interval $\left[-\frac{U}{4p(1-p)}, \frac{U}{2c} U_{A\alpha} \right]$, and $0 < \delta < 1$, the expectation is well-defined for all policies that implement feasible actions. For any $\delta \in (0,1)$ and initial state $s \in S$, let

$$V_\delta(s) \equiv \sup_{\mu} V_\delta(\mu; s).$$

The following second theorem from Ross (2014) describes sufficient conditions to be able to apply Theorem Ross (1), and therefore to guarantee that there exists an optimal SD-policy of (MDP), which is also an optimal SD-policy of (MDP, s^0) and defines an optimal SD-mechanism for our dynamic matching problem.

Theorem Ross (2) (V.2.2 in Ross (2014)) *If there exists $M < \infty$ such that*

$$|V_\delta(s) - V_\delta(0)| < M \text{ for all } \delta \in (0,1) \text{ and } s \in S,$$

then there exists a bounded function $h(s)$ and a constant g satisfying (8).

By Theorems Ross (1) and Ross (2), proving the following claim is sufficient to complete the proof of Theorem A1.

Claim (1): *There exists $M < \infty$, such that, for any $s \in S$, $\delta \in (0,1)$, and policy μ of (MDP), there exists another policy μ' of (MDP) with*

$$|V_\delta(\mu; s) - V_\delta(\mu'; 0)| < M$$

Proof of Claim (1): Take any initial state $s \in S$, $\delta \in (0,1)$, and policy μ of (MDP). Let $s > 0$ (a similar proof, which we omit, applies for the case of $s < 0$). We assume that μ matches agents on a FIFO basis (note that the assumption does not change the value $V_\delta(\mu; s)$ because agents of the same type are interchangeable from a welfare perspective). Let $m \equiv \lfloor (U/2c) + 1 \rfloor$, denoting the maximum number of incongruent pairs that μ would hold on the market. Observe that, during the first $3m$ periods, $3m$ incongruent pairs arrive at the market. Let n be the number of (A, β) pairs arriving during the first $3m$ periods, so $3m - n$ is the number of (B, α) pairs arriving. . Suppose that $n < 2m$, so at least m (B, α) pairs arrived. As the policy μ matches agents on a FIFO basis and it never holds (A, α) or (B, β) pairs, *all agents (of types A or β) who were initially in the market would be matched by μ within the first $3m$ periods.* Next, suppose $n \geq 2m$. As μ holds at most m incongruent pairs at any time, it would hold at most m (A, β) pairs at the end of period $3m$. Because of the FIFO protocol, *all (A, β) pairs held by μ at the end of period $3m$ must have arrived after the initial period.*

We construct another policy μ' of (MDP) which differs from μ only when the initial state is 0. If the initial state is 0, in each of the first $3m$ periods, μ' holds all agents, who arrived after the initial period and would have been held by μ if the initial state were s and the same types of agents have arrived. In the same $3m$ periods, the policy μ' with the initial state 0 matches all other agents arbitrarily and immediately.

For the observation at the beginning of the proof, after period $3m$, the policy μ' with the initial state 0 creates the same matches as μ would if the initial state were s (note that the last part of the construction is feasible, because all agents held by μ at the end of period $3m$ after initial state s arrived after the initial period, so they are also available for μ' to hold with initial state 0). Therefore, the reward function values generated by μ' with initial state 0 differ from those generated by μ' with initial state s only for the first $3m$ periods. Thus,

$$|V_\delta(\mu; s) - V_\delta(\mu'; 0)| \leq 3m \left(\frac{U}{2c} U_{A\alpha} + \frac{U}{4p(1-p)} \right).$$

where the inequality is guaranteed by the fact that the reward function is bounded in $\left[-\frac{U}{4p(1-p)}, \frac{U}{2c} U_{A\alpha} \right]$. Note that the right hand side is independent of s and δ. Therefore, Claim (1) holds for $M \equiv 3 \left(\frac{U}{2c} + 1 \right) \left(\frac{U}{2c} U_{A\alpha} + \frac{U}{4p(1-p)} \right)$. ■

Proof of Proposition 1:

Step 1 (Existence of Thresholds $(\bar{k}_A, \bar{k}_\alpha)$)

Any stationary and deterministic policy d of (MDP) is associated with two thresholds, each of which represents the largest number of incongruent pairs of either type (A, β) or (B, α) to hold.[40] Define

$$\begin{aligned} \bar{k}_A &\equiv \min\{s \mid s > 0, d(s) < s\} - 1, \quad \text{and} \\ \bar{k}_\alpha &\equiv \min\{|s| \mid s < 0, d(s) > s\} + 1. \end{aligned}$$

The thresholds $(\bar{k}_A, \bar{k}_\alpha)$ are well-defined. Indeed, policies maintain only a bounded number of unmatched pairs in the market. In particular, $d(\tilde{s}) < \tilde{s}$ and $d(-\tilde{s}) > -\tilde{s}$.We claim that the value of a policy is uniquely determined by the thresholds and decisions at the thresholds. Given policy d, define

$$d'(s) \equiv \begin{cases} d(s) & \text{if } -\bar{k}_\alpha \leq s \leq \bar{k}_A \\ d(\bar{k}_A) & \text{if } s > \bar{k}_A \\ d(-\bar{k}_\alpha) & \text{if } s < -\bar{k}_\alpha. \end{cases}$$

The Markov processes induced by d and d', namely $\{(s^\tau, r(s^\tau, d(s^\tau))\}_{\tau=0}^\infty$ and $\{(s^\tau, r(s^\tau, d'^\tau))\}_{\tau=0}^\infty$, are identical. Thus, $\tilde{v}(d') = \tilde{v}(d)$. We can therefore characterize any policy d by its corresponding thresholds $(\bar{k}_A, \bar{k}_\alpha)$ and decisions at the thresholds $(d(\bar{k}_A), d(\bar{k}_\alpha))$.[41]

Step 2 (Stationary Distribution of $k_{A\alpha}$)

[40] Indeed, suppose a policy dictates matches to be formed when the number of, say, (A, β) pairs exceeds k_A^1 or $k_A^2 > k_A^1$. The number of (A, β) pairs would then never surpass k_A^2, so that the relevant threshold for outcomes would be the minimal threshold k_A^1.

[41] As mentioned in the body of the text, there is multiplicity regarding prescriptions for states that are never reached. With thresholds $\bar{k}_A$ and $\bar{k}_\alpha$ the market never has more than $\bar{k}_A + 1$ A-squares or more than $\bar{k}_\alpha + 1$ α-rounds. The specification of what happens outside of these regions therefore has no impact on outcomes.

Given a policy d, in the next result we study an ergodic Markov process and characterize the unique stationary distribution of $k_{A\alpha}$.

Claim (1) *Take* $\bar{k}_A, \bar{k}_\alpha \in [1, \tilde{s}-1] \cap Z_+$, *and* $z_A, z_\alpha \in Z_+$ *with* $z_A \leq \bar{k}_A$ *and* $z_\alpha \leq \bar{k}_\alpha$. *A policy* d *of (MDP) defined by*

$$d(s) \equiv \begin{cases} s & \text{if } -\bar{k}_\alpha \leq s \leq \bar{k}_A \\ \bar{k}_A - z_A & \text{if } s > \bar{k}_A \\ -\bar{k}_\alpha + z_\alpha & \text{if } s < -\bar{k}_\alpha. \end{cases}$$

induces a Markov chain of $k_{A\alpha}$. *The unique steady state distribution* $\pi = (\pi_{-\bar{k}_\alpha}, \ldots, \pi_{\bar{k}_A})$ *is such that:*

1. *(Middle Range) for* $-\bar{k}_\alpha + z_\alpha \leq k \leq \bar{k}_A - z_A$, $\pi_k = \pi_0 = \frac{1}{\bar{k}_A + \bar{k}_\alpha - z_A/2 - z_\alpha/2 + 1}$,
2. *(Upper Range) for* $z = 1, \ldots, z_A$, $\pi_{\bar{k}_A - z_A + z} = \pi_0 \left(1 - \frac{z}{z_A + 1}\right)$,
3. *(Lower Range) for* $z = 1, \ldots, z_\alpha$, $\pi_{-\bar{k}_\alpha + z_\alpha - z} = \pi_0 \left(1 - \frac{z}{z_\alpha + 1}\right)$.

That is, the stationary distribution is uniform in the middle range. The stationary probability mass decreases as $k_{A\alpha}$ *approaches* $\bar{k}_A$ *or* $\bar{k}_\alpha$.

Proof of Claim (1): Denote by

$$\mathbf{x}^\tau \equiv (x^\tau_{-\bar{k}_\alpha}, x^\tau_{-\bar{k}_\alpha+1}, \ldots, x^\tau_{\bar{k}_A-1}, x^\tau_{\bar{k}_A})^{tr} \in \{0,1\}^{\bar{k}_A + \bar{k}_\alpha + 1}$$

the timed vector such that $x_i^\tau = \mathbf{1}(k_{A\alpha} = i)$. Then, $\mathbf{x}^{\tau+1} = T_d \mathbf{x}^\tau$, where

$$\mathbf{T}_d = \begin{pmatrix} 0 & 1/2 & \ldots & 0 & 0 \\ 1/2 & 0 & \ldots & 0 & 0 \\ 0 & 1/2 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1/2 & 0 \\ 0 & 0 & \ldots & 0 & 1/2 \\ 0 & 0 & \ldots & 1/2 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 1/2 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 1/2 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 \end{pmatrix}.$$

The second matrix on the right hand side has two non-zero elements valued at 1/2. Each represents two scenarios, a transition from $\bar{k}_A$ upon the arrival of an (A, β) pair to $\bar{k}_A - z_A$, and a transition from $-\bar{k}_\alpha$ upon the arrival of a (B, α) pair to $-\bar{k}_\alpha + z_\alpha$. The first matrix includes all other transitions. The Markov chain is ergodic, and the unique stationary distribution of $k_{A\alpha}$ exists. Then, $\boldsymbol{\pi}$ in Claim (1) is the unique stationary distribution.

Step 3 (Welfare)

We compute the average welfare (i.e., total welfare per period) for any stationary and deterministic mechanism μ. Let d be the associated policy of (MDP) with thresholds $(\bar{k}_A, \bar{k}_\alpha)$ and the decisions at the thresholds identified by (z_A, z_α).[42] Denote the corresponding unique stationary distribution over states by $\boldsymbol{\pi} = (\pi_{-\bar{k}_\alpha}, \pi_{-\bar{k}_\alpha+1}, \ldots, \pi_{\bar{k}_A-1}, \pi_{\bar{k}_A})$.

First, we compute the average total surplus generated in one time period. A newly arrived pair is of type (A, α) with probability p^2, in which case the optimal mechanism generates a surplus equal to $U_{A\alpha}$. Similarly, a newly arrived pair is of type (B, β) with probability $(1-p)^2$, in which case the optimal mechanism generates a surplus equal to $U_{B\beta}$.

Suppose an (A, β) pair arrives at time t when the stock is $k_{A\alpha}^{t-1}$. If $k_{A\alpha}^{t-1} < 0$, the mechanism creates one (A, α) and one (B, β) pair, generating a surplus equal to $U_{A\alpha} + U_{B\beta}$. If $0 \leq k_{A\alpha}^{t-1} < \bar{k}_A$, the mechanism creates no matches (and no additional surplus), and if $k_{A\alpha}^{t-1} = \bar{k}_A$, the mechanism creates $(z_A + 1)$ matches of (A, β) pairs. Analogous conclusions pertain to the case in which a (B, α) pair arrives. The expected match surplus per period is therefore:

$$\begin{aligned} & p^2 U_{A\alpha} + (1-p)^2 U_{B\beta} + (1-\pi_0)p(1-p)(U_{A\alpha} + U_{B\beta}) \\ & + \pi_{\bar{k}_A} p(1-p)(z_A+1)U_{A\beta} + \pi_{\bar{k}_\alpha} p(1-p)(z_\alpha+1)U_{B\alpha} \\ = \; & pU_{A\alpha} + (1-p)U_{B\beta} - \pi_0 p(1-p)U. \end{aligned}$$

Next we compute the average total waiting costs incurred by agents waiting in line for one period. During the transition from time t culminating at stock $k_{A\alpha}$ to time $t+1$, $2|k_{A\alpha}|$ agents wait in line. So the total costs of waiting incurred during this one time period is $2|k_{A\alpha}|c$. Thus, a mechanism with thresholds $(\bar{k}_A, \bar{k}_\alpha)$ results in expected total costs of waiting equal to

$$\sum_{k=1}^{\bar{k}_A} 2c\pi_k|k| + \sum_{k=-1}^{-\bar{k}_\alpha} 2c\pi_k|k|$$

The first term equals to

$$\begin{aligned} & (2c\pi_0) \sum_{k=1}^{\bar{k}_A - z_A} k + (2c\pi_0) \sum_{z=1}^{z_A} \left(1 - \frac{1}{z_A+1}\right)(\bar{k}_A - z_A + z) \\ = \; & (2c\pi_0) \left(\frac{(\bar{k}_A - z_A)(\bar{k}_A - z_A + 1)}{2} + \frac{(\bar{k}_A - z_A)z_A}{2} + \frac{z_A(z_A+1)}{2} - \frac{z_A(2z_A+1)}{6} \right) \\ = \; & (c\pi_0) \left((\bar{k}_A - z_A)(\bar{k}_A + 1) + \frac{z_A(z_A+2)}{3} \right). \end{aligned}$$

[42] That is, for $s > \bar{k}_A, d(s) = \bar{k}_A - z_A$ and for $s < -\bar{k}_\alpha, d(s) = -\bar{k}_\alpha + z_\alpha$.

The second term is computed similarly. The average welfare of the mechanism μ is then

$$W(\bar{k}_A, \bar{k}_\alpha, z_A, z_\alpha) = pU_{A\alpha} + (1-p)U_{B\beta} - \pi_0 p(1-p)U$$
$$-(c\pi_0)\left((\bar{k}_A - z_A)(\bar{k}_A+1) + (\bar{k}_\alpha - z_\alpha)(\bar{k}_\alpha+1) + \frac{z_A(z_A+2)+z_\alpha(z_\alpha+2)}{3}\right), \quad (9)$$

where $\pi_0 = \frac{2}{2\bar{k}_A+2\bar{k}_\alpha-z_A-z_\alpha+2}$.

Step 4 (Matching at Most One Pair at a Time)

We show that, in finding an optimal mechanism, we can focus on mechanisms satisfying $z_A = z_\alpha = 0$. In fact, generically this restriction is necessary for a mechanism to be optimal. The proof is immediate from the following claim, which completes the proof of Proposition 1.

Claim (2) *Fix any $\bar{k}_\alpha$ and $z_\alpha (\leq \bar{k}_\alpha)$. For any $\bar{k}_A \geq 1$ and $0 \leq z_A \leq \bar{k}_A - 1$,*

$$W(\bar{k}_A, \bar{k}_\alpha, z_A+1, z_\alpha) \geq W(\bar{k}_A, \bar{k}_\alpha, z_A, z_\alpha)$$

implies

$$W(\bar{k}_A - 1, \bar{k}_\alpha, z_A, z_\alpha) \geq W(\bar{k}_A, \bar{k}_\alpha, z_A+1, z_\alpha).$$

That is, whenever a mechanism with a larger z_A leads to a higher average welfare, we can find a mechanism with an even higher average welfare by decreasing the threshold $\bar{k}_A$, while adhering to a smaller z_A.

Proof of Claim (2): Let

$$\begin{aligned} \phi &= 2\bar{k}_A + 2\bar{k}_\alpha - z_A - z_\alpha + 1 \quad \text{and} \\ \psi &= (\bar{k}_A - z_A)(\bar{k}_A+1) + (\bar{k}_\alpha - z_\alpha)(\bar{k}_\alpha+1) + \frac{z_A(z_A+2)+z_\alpha(z_\alpha+2)}{3}. \end{aligned}$$

The first inequality in Lemma A2 holds if and only if

$$\frac{p(1-p)U}{\phi} + \frac{c}{\phi}\left(\psi - \bar{k}_A + \frac{2z_A}{3}\right) \leq \frac{p(1-p)U}{\phi+1} + \frac{c}{\phi+1}\psi,$$

or equivalently

$$p(1-p)U + c\psi - (\phi+1)c\left(\bar{k}_A - \frac{2z_A}{3}\right) \leq 0. \quad (10)$$

The second inequality in Lemma A2 holds if and only if

$$\frac{p(1-p)U}{\phi-1} + \frac{c}{\phi-1}\left(\psi - 2\bar{k}_A + z_A\right) \leq \frac{p(1-p)U}{\phi} + \frac{c}{\phi}\left(\psi - \bar{k}_A + \frac{2z_A}{3}\right),$$

or equivalently

$$p(1-p)U + c\psi - c\left(\bar{k}_A - \frac{2z_A}{3}\right) - \phi c\left(\bar{k}_A - \frac{z_A}{3}\right) \leq 0. \tag{11}$$

Clearly, ((10)) implies ((11)).

Claim (2) complete the proof of Proposition 1. Furthermore, Claim (2) illustrates that there is always an optimal mechanism identified by $z_A = z_\alpha = 0$. From the proof, notice that if $z_A > 0$, inequality (10) implies that inequality (11) holds with a strict inequality. Therefore, in any optimal mechanism, $z_A, z_\alpha < 2$. In fact, multiplicity can emerge only when there is multiplicity in the thresholds $\bar{k}_A, \bar{k}_\alpha$ fixing $z_A = z_\alpha = 0$. Indeed, suppose there is an optimal mechanism with $\bar{k}_A$ and $z_A = 1$ and some $\bar{k}_\alpha, z_\alpha$. From the proof of Claim (2), it follows that

$$W(\bar{k}_A, \bar{k}_\alpha, 1, z_\alpha) - W(\bar{k}_A, \bar{k}_\alpha, 0, z_\alpha) = W(\bar{k}_A - 1, \bar{k}_\alpha, 0, z_\alpha) - W(\bar{k}_A, \bar{k}_\alpha, 1, z_\alpha).$$

The optimality of $\bar{k}_A$ and $z_A = 1$ implies that, in the above equality, both sides equal to 0 (otherwise, the mechanism identified by $\bar{k}_A - 1$ and $z_A = 0$, with $\bar{k}_\alpha, z_\alpha$, would generate greater welfare). In particular, there are optimal mechanisms identified by both $\bar{k}_A - 1$ and $z_A = 0$ as well as $\bar{k}_A$ and $z_A = 0$. ■

9.2 Proofs Regarding Discretionary Matching

9.2.1 Players' Markov Decision Problem (MDP)

In this section we study the stationary* equilibria under several priority protocols. A key first step is to formalize each player's dynamic decision problem, defined by other players' equilibrium strategies, as a Markov decision problem (MDP). We now formalize an A-square's problem (and omit a similar formalization for the other types of players). Note that this formalization applies to the case of the FIFO priority protocol, as well as to the alternative protocols studied in Section 5 and in the Online Appendix.

Fix any priority rule, and take an A-square, say player i, who arrived in period $t_0 \geq 1$. Assume that all other players follow a stationary* strategy-profile Ψ_{-i}. Player i solves an infinite-horizon dynamic decision problem, defined by Ψ_{-i}. For each period $t \geq t_0$, let $\theta_i^t = (\mathbf{s}^t, q_i^t)$ denote the player's augmented state, where $\mathbf{s}^t = (s_A^t, s_B^t, s_\alpha^t, s_\beta^t)$ denotes the state of the market, and q_i^t denotes player i's rank among A-squares, or $q_i^t = 0$ if player i matched before period t. We denote by Θ_i the set of augmented states that player i may experience. In each period $t \geq t_0$, player i's chooses a demand $d_i \in \{\alpha, \beta\}$, where α represents demand for an α-round, and β represents demand for *any* round. The stage-game payoff $u_i(d_i, \theta_i, \Psi_{-i})$ is either match surplus ($U_A(\alpha)$ or $U_A(\beta)$), waiting costs $-c$, or 0 (if $q_i = 0$). The initial augmented state is $\theta_i^{t_0} = (\mathbf{s}^{t_0}, q_i^{t_0})$ such that either $q_i^{t_0} = s_A^{t_0}$ under FIFO or $q_i^{t_0} = 1$ under LIFO. The transition of augmented state is straightforward from our description of the model, so we omit.

A strategy σ_i is any rule for choosing demands. A choice may depend on history of the process up to that point, including the history from periods even before player i's arrival, and it may be random. The payoff for player i from strategy σ_i is

$$U_i(\sigma_i; \theta_i, \Psi_{-i}) \equiv E_{\sigma_i}\left[\sum_{t=t_0}^{\infty} u_i(d_i, \theta_i^t, \Psi_{-i}) : \theta_i^{t_0} = \theta_i\right].$$ [43]

A stationary* strategy applies the same decision rule $\psi_i : \Theta_i \to \{\alpha, \beta\}$ in every period. The payoff from a stationary* strategy ψ_i is

$$U_i(\psi_i; \theta_i, \Psi_{-i}) = E\left[\sum_{t=t_0}^{\infty} u_i(\psi_i(\theta_i^t), \theta_i^t, \Psi_{-i}) : \theta_i^{t_0} = \theta_i\right].$$

We focus on player i's Markov random strategies in the sense that a choice in each period is independent of past history. This restriction is without loss of generality because

$$\sup_{\sigma_i \in \Sigma_i} U_i(\sigma_i; \theta_i, \Psi_{-i}) = \sup_{\sigma_i \in \Sigma_i^{MR}} U_i(\sigma_i; \theta_i, \Psi_{-i}), \tag{12}$$

where Σ_i and Σ_i^{MR} denote the set of all strategies and all Markov random strategies, respectively (Proposition 7.1.1 of Puterman (2014)). The restriction to Markov random strategies allows us to normalize player i's arrival time as $t_0 = 0$ and write

$$U_i(\sigma_i; \theta_i, \Psi_{-i}) = E_{\sigma_i}\left[\sum_{t=0}^{\infty} u_i(d_i, \theta_i^t, \Psi_{-i}) : \theta_i^0 = \theta_i\right] \quad \text{for each } \sigma_i \in \Sigma_i^{MR}.$$

We extend player i's decision problem as a MDP, with an arbitrary initial state. That is, an initial state θ_i^0 can be any in $\Theta_i \subseteq \{(\mathbf{s}, q_i) \in \mathbb{Z}_+^5 : s_A + s_B = s_\alpha + s_\beta, 1 \leq q_i \leq s_A\}$.[44] A *policy* μ_i of MDP is any Markovian random rule for choosing demands in the MDP. A *stationary and deterministic policy* (SD-policy, for short) applies the same decision rule in every period.

[44] Recall that, the initial condition $\theta_i^0 = (\mathbf{s}^0, q_i^0)$ should satisfy $q_i^0 = s_A^0$ under FIFO and $q_i^0 = 1$ under LIFO. We remove such restrictions in (MDP).

We denote a SD-policy by $\psi_i : \Theta_i \to \{\alpha, \beta\}$. The *value of a policy* μ_i, for each initial state θ_i, is defined by

$$v_i(\mu_i; \theta_i, \Psi_{-i}) \equiv U_i(\mu_i; \theta_i, \Psi_{-i}).$$

Last, the *value of the MDP*, for each initial state θ_i, is

$$v_i^*(\theta_i; \Psi_{-i}) \equiv \sup_{\mu_i} v_i(\mu_i; \theta_i, \Psi_{-i}).$$

We characterize the value of the MDP, $v_i^*(\cdot; \Psi_{-i}) : \Theta_i \to \mathbb{R} \cup \{-\infty\}$ and find an optimal SD-policy, whose value is equal to $v_i^*(\theta_i; \Psi_{-i})$ for every initial state θ_i. An optimal SD-policy of MDP defines a best-response stationary strategy for player i given any initial state. A stationary* strategy profile $\Psi = (\psi_A, \psi_B, \psi_\alpha, \psi_\beta)$ is a stationary* equilibrium if, for every A-square (and similarly for other types), ψ_A is an optimal SD-policy of the MDP defined by all other player's equilibrium strategies Ψ_{-i}.

We recall the following definition and theorems from Puterman (2014) which are associated to player i's problem, but hold for general Markov decision problems.

Definition 3. *(Optimality Equation, Equation 6.2.2 with $\lambda = 1$, or Equation 7.1.8 of Puterman (2014)) We refer to the following system of equations as the* optimality equation

$$v(\theta_i) = \max_{d \in \{\alpha, \beta\}} \left[u_i(d_i, \theta_i, \Psi_{-i}) + \sum_{\theta_i' \in \Theta_i} p(\theta_i' | \theta_i, d_i, \Psi_{-i}) v(\theta_i') \right], \tag{13}$$

for all $\theta_i \in \Theta_i$.

Theorem Puterman (1) (7.1.3 of Puterman (2014)) *The value of MDP, $v_i^*(\cdot; \Psi_{-i})$, is a solution of the optimality equation (13).*

Theorem Puterman (2) (7.2.5 (a) of Puterman (2014)) *A policy μ_i^* is optimal if and only if the value of the policy $v_i^*(\cdot; \Psi_{-i}) : \Theta_i \to R \cup \{-\infty\}$ is a solution of the optimality equation (13).*

Note that the value of MDP is not a unique solution of the optimality equation. For example, we can add a constant to the value and find another solution. This non-uniqueness of a solution of the optimality equation makes the analysis with the total expected utility different from the standard approach to the model with time-discount. Finally, observe that the state space Θ_i for player i can be finite under some stationary* strategy-profile Ψ_{-i} chosen by other players. We have

Theorem Puterman (3) (7.1.9 of Puterman (2014)) *If Θ_i is finite, then there exists an optimal SD-policy.*

9.2.2 Proofs Regarding Stationary* Equilibria under FIFO

Lemma A3 *Under FIFO, if $\Psi^* = (\psi_A^*, \psi_B^*, \psi_\alpha^*, \psi_\beta^*)$ is a stationary* equilibrium, then*

$$\psi_A^*(\mathbf{s}, q) = \begin{cases} \alpha \text{ or } \beta & \text{if } q \leq s_\alpha, \\ \alpha & \text{if } 1 \leq q - s_\alpha \leq \bar{k}^{fifo}, \\ \beta & \text{otherwise}, \end{cases} \tag{14}$$

where

$$\bar{k}^{fifo} \equiv \left\lfloor \frac{p(U_A(\alpha) - U_A(\beta))}{c} \right\rfloor = \left\lfloor \frac{p(U_\alpha(A) - U_\alpha(B))}{c} \right\rfloor.$$

We omit an analogous lemma for α-rounds.

Proof of Lemma A3: We show that, if Ψ^* is a stationary* equilibrium, for any augmented state $\theta_i = (\mathbf{s}, q_i)$ for an A-square i we have

$$\psi_A^*(\theta_i) = \begin{cases} \alpha & \text{if } 1 \leq q_i - s_\alpha \leq \bar{k}^{fifo} \\ \beta & \text{if } q_i - s_\alpha > \bar{k}^{fifo}. \end{cases} \tag{15}$$

The proof is by induction. First, we characterize the equilibrium behavior of an A-square who finds no available α-round in a period, and, if she remains unmatched, *she would become the first in the queue in the next period.* Take any stationary* strategy-profile $\Psi = (\psi_A, \psi_B, \psi_\alpha, \psi_\beta)$. Take any A-square, say player i, whose augmented state in a period $t_0 = 0$ satisfies $q_i^{t_0} = s_\alpha^{t_0} + 1$. That is, in period t_0, player i finds no available α-round (i.e., $q_i^{t_0} > s_\alpha^{t_0}$) and, if she is not matched, she would become the first A-square in the queue (i.e., $q_i^{t_0} - s_\alpha^{t_0} = 1$). Player i solves the following problem:

$$v_i^*(\theta_i; \Psi_{-i}) \equiv \sup_{\sigma_i \in \Sigma_i^{MR}} E_{\sigma_i}\left[\sum_{t=0}^{\infty} u_i(d_i, \theta_i^t, \Psi_{-i}) : \theta_i^0 = \theta_i\right].$$[45]

Since $s_\beta^0 \geq q_i^0 + s_B^0 - s_\alpha^0 > 0$,we know that at least one β-round is available in period $t_0 = 0$. Also, the first α-round to arrive at the market will be available to match for player i, but it takes in expectation $1/p$ periods until such α-round arrives. Therefore, we have

$$v_i^*(\theta_i; \Psi_{-i}) = \max\left\{U_A(\beta), U_A(\alpha) - \frac{c}{p}\right\}.$$

Hence, if ψ_A^* is part of a stationary* equilibrium, it must be that,

$$\psi_A^*(\theta_i) = \begin{cases} \alpha & \text{if } q_i - s_\alpha = 1 \leq \bar{k}^{fifo} \\ \beta & \text{if } q_i - s_\alpha = 1 > \bar{k}^{fifo}. \end{cases} \tag{16}$$

That is, ((15)) holds for every $\theta_i = (\mathbf{s}, q_i)$ with $q_i - s_\alpha = 1$.

Next, we complete the induction. Take any $k \in \mathbb{Z}_{++}$ and a stationary* strategy-profile Ψ such that ψ_A satisfies ((15)) for every augmented state $\theta = (\mathbf{s}, q)$ with $q - s_\alpha \leq k$. Consider any A-square, say player i, whose augmented state in period 0 (normalized) satisfies $q_i^0 = s_\alpha^0 + (k+1)$. Assume that every other A-square, say player j with $q_j^0 \leq s_\alpha^0 + k < q_i^0$, plays ψ_A. Given that each player's rank in the queue only improves over time, the first $\min\{k, \bar{k}^{fifo}\}$ arriving α-rounds in the future are not available for player i, but the next arriving α-round will be. In expectation, it takes $\frac{\min\{k,\bar{k}^{fifo}\}+1}{p}$ periods until a α-round becomes available for player i. As such,

$$\begin{aligned} v_i^*(\theta_i; \Psi_{-i}) &= \max\left\{U_A(\beta), U_A(\alpha) - \frac{(\min\{k, \bar{k}^{fifo}\}+1)c}{p}\right\} \\ &= \max\left\{U_A(\beta), U_A(\alpha) - \frac{(k+1)c}{p}, U_A(\alpha) - \frac{(\bar{k}^{fifo}+1)c}{p}\right\} \\ &= \max\left\{U_A(\beta), U_A(\alpha) - \frac{(k+1)c}{p}\right\}, \end{aligned}$$

where the last equality follows from the definition of $\bar{k}^{fifo}$.

Therefore, if ψ_A is part of a stationary* equilibrium, $\psi_A(\theta_i)$ must satisfy ((15)) for any augmented state $\theta_i = (\mathbf{s}, q_i)$ with $q_i - s_\alpha = k+1$. ■

Lemma A4 There exists a stationary* equilibrium $\Psi^* = (\psi_A^*, \psi_B^*, \psi_\alpha^*, \psi_\beta^*)$ such that

(a) ψ_A^* (and ψ_α^*) satisfies ((14)) (respectively, a similar condition for α-rounds), and

(b) $\psi_\beta^*(\mathbf{s}, s_\beta) = B$ and $\psi_B^*(\mathbf{s}, s_B) = \beta$, whenever $s_B > 0$ and $s_\beta > 0$.

Proof of Lemma A4: First, let us address the decisions of the A-squares. Take any stationary* strategy-profile $\Psi = (\psi_A, \psi_B, \psi_\alpha, \psi_\beta)$ that satisfies the conditions (a) and (b) in the claim. We prove that ψ_A is a best-response for an A-square, say player i, regardless of her initial augmented state. Let Θ_A be the set of all possible augmented states for player i, conditional on Ψ_{-i}. That is,

$$\Theta_A \equiv \{(\mathbf{s}, q) \in \mathbb{Z}_+^5 : -\bar{k}^{fifo} - 1 \leq s_{A\alpha} \leq \bar{k}^{fifo} + 2,\ 0 \leq q_i \leq s_A\}.$$[46]

We extend the player i's decision problem as a MDP with an arbitrary initial state (ignoring that her initial state in the discretionary matching satisfies $q_i = s_A$). That is, player i's MDP is

$$v_i^*(\theta_i; \Psi_{-i}) \equiv \sup_{\mu_i \in \Sigma_i^{MR}} v_i(\mu_i; \theta_i, \Psi_{-i}), \quad \text{for all } \theta_i \in \Theta_A,$$

where

$$v_i(\theta_i; \mu_i, \Psi_{-i}) \equiv U_i(\mu_i; \theta_i, \Psi_{-i}) \equiv E_{\mu_i}\left[\sum_{t=0}^{\infty} u_i(d_i, \theta_i^t, \Psi_{-i}) : \theta_i^0 = \theta_i\right].$$

If player i follows the SD-policy ψ_A that satisfies ((14)), then, for $\theta_i = (\mathbf{s}, q_i)$

$$\begin{aligned} v_i(\theta_i; \psi_A, \Psi_{-i}) &= E\left[\sum_{t=0}^{\infty} u_i(\psi_A(\theta_i^t), \theta_i^t, \Psi_{-i}) : \theta_i^0 = \theta_i\right] \\ &= \begin{cases} U_A(\alpha) & \text{if } q_i \le s_\alpha \\ U_A(\alpha) - \frac{(q_i - s_\alpha)c}{p} & \text{if } 1 \le q_i - s_\alpha \le \bar{k}^{fifo} \\ U_A(\beta) & \text{otherwise.} \end{cases} \end{aligned}$$

From the construction of ψ_A in the proof of Lemma A3, it is easy to verify that $v_i(\cdot; \psi_A, \Psi_{-i}) : \Theta_A \to \mathbb{R} \cup \{-\infty\}$ solves the optimality equation (13):

$$v(\theta_i) = \max_{d \in \{\alpha, \beta\}} \left[u_i(d_i, \theta_i, \Psi_{-i}) + \sum_{\theta_i' \in \Theta_A} p(\theta_i' | \theta_i, d_i, \Psi_{-i}) v(\theta_i')\right] \quad \text{for all } \theta_i \in \Theta_A.$$

Thus, by Theorem Puterman (2), ψ_A is a SD optimal policy of the MDP for player i. In particular, each A-square is best-responding by playing ψ_A, regardless her initial augmented state.

Next, let us address the decisions by β-rounds. Let Θ_β denote the set of all possible augmented states that a β-round may experience:

$$\Theta_\beta \equiv \{(\mathbf{s}, q) \in \mathbb{Z}_+^5 : s_A + s_B = s_\alpha + s_\beta,\, q \le s_\beta\},$$

where $q = 0$ represents the augmented state after the player is matched.

We take ψ_A and ψ_α satisfying condition (a) in Lemma A4. Then, we want to construct a SD-strategy $\psi_\beta : \Theta_\beta \to \{A, B\}$ (and $\psi_B : \Theta_B \to \{\alpha, \beta\}$, whose similar construction we omit) such that $\Psi = (\psi_A, \psi_B, \psi_\alpha, \psi_\beta)$ constitutes a stationary* equilibrium. We start the construction of ψ_β (and ψ_B) with the following assumption:

Assumption A1 For any $(\mathbf{s}, s_\beta) \in \Theta_\beta$, $(\mathbf{s}, s_B) \in \Theta_B$ with $s_\beta > 0$ and $s_B > 0$,

(1) $\psi_\beta(\mathbf{s}, s_\beta) = \psi_\beta(\mathbf{s}, s_\beta - 1) = ... = \psi_\beta(\mathbf{s}, 1 + s^+_{A\alpha}) = B$;

(2) $\psi_B(\mathbf{s}, s_B) = \psi_B(\mathbf{s}, s_B - 1) = ... = \psi_B(\mathbf{s}, 1 + s^-_{A\alpha}) = \beta$.

Note that Assumption A1 is consistent with condition (b) in Lemma A4. The construction of ψ_β that follows will guarantee that ψ_β is a best-response for a β-round in any period and with any initial augmented state, under Assumption A1 applied to the other players. Finally, we will justify Assumption A1 as describing best-response strategies.

Take any β-round, say player i, and any stationary* strategy-profile Ψ_{-i} such that ψ_A and ψ_α satisfy ((14)) and Assumption A1 holds. Then, there is no period in which both A-squares and α-rounds wait at the market. Therefore, for any t, the stock $\mathbf{k}^t \equiv (k^t_A, k^t_B, k^t_\alpha, k^t_\beta)$ satisfies $k^t_A k^t_\alpha = 0$ and $-\bar{k}^{fifo} \leq k^t_{A\alpha} \leq \bar{k}^{fifo}$. In addition, by Assumption A1, there is no period in which at least two B-squares and two β-rounds wait by demanding α and A, respectively.

We characterize the set of augmented states that player i can experience, which we denote by $\Theta'_\beta \subseteq \Theta_\beta$. Let $K \subseteq \mathbb{Z}^4_+$ denote the set of possible states at the end of each period, taking into account ψ_A and ψ_α, Assumption A1, and every possible unilateral deviation by an individual β-round or B-square. That is,

$$\mathbf{k} \equiv (k_A, k_B, k_\alpha, k_\beta) \in K \iff \begin{array}{ll} (i) & k_A + k_B = k_\alpha + k_\beta, \\ (ii) & k_A k_\alpha = 0, \\ (iii) & -\bar{k}^{fifo} \leq k_{A\alpha} \leq \bar{k}^{fifo}, \\ (iv) & k_{A\alpha} \geq 0 \implies k_B \leq 1, \text{ and } k_{A\alpha} \leq 0 \implies k_\beta \leq 1. \end{array}$$

Then, Θ'_β is a subset of Θ_β such that

$$(\mathbf{s}, q) \in \Theta'_\beta \iff (\exists \mathbf{k} \in K) \quad s.t.\ \mathbf{s} - \mathbf{k} \in \{(1,0,1,0), (0,1,0,1), (1,0,0,1), (0,1,1,0)\}.$$

It is clear that Θ'_β is finite, and for any $(\mathbf{s}, q) \in \Theta'_\beta$, we have $0 \leq q \leq \bar{k}^{fifo} + 2$. It is sufficient to define ψ_β over augmented states in Θ'_β only, as an augmented state $(\mathbf{s}, q) \notin \Theta'_\beta$ never occurs. Note that, under the FIFO protocol, the ranking of a β-round, say player i, improves as he waits in the market. Thus, player i's continuation payoff from the MDP after his rank becomes 1 is independent of his actions in a state with a rank lower than 1. For each possible ranking of a β-round, $q \in \{1, 2, \ldots, \bar{k}^{fifo}+2\}$, let $\Theta'_{\beta,q}$ be the set of augmented state with rank q (i.e., $\Theta'_{\beta,q} \equiv \{(\mathbf{s}, q) \in \Theta'_\beta\}$). We construct $\psi_{\beta,q} : \Theta'_{\beta,q} \to \{A, B\}$, sequentially from $q = 1$ to $q = \bar{k}^{fifo} + 2$, and define $\psi_\beta : \Theta'_\beta \to \{A, B\}$ as $\psi_\beta(\mathbf{s}, q) \equiv \psi_{\beta,q}(\mathbf{s})$. In the construction, we will guarantee that ψ_β constitutes a best-response for a β-square, taking as given ψ_A, ψ_α, and Assumption A1 applied to the other players.

Step 1: Construction of $\psi_{\beta,1}$

Consider a β-round, say player i, who is the first in the queue at some period. Player i solves a dynamic decision problem, defined by ψ_A, ψ_α, and Assumption A1 (applied to other players' strategies). We extend the player i's dynamic decision problem as a MDP with an arbitrary initial state $\theta_i \in \Theta'_{\beta,1}$. Let $v^*(\theta_i)$ denote the maximal expected total payoffs for player i with an initial augmented state θ_i. Theorem Puterman (3) guarantees that there exists a SD optimal policy. Moreover, any policy whose values solve the optimality equation, is optimal by Theorem Puterman (2),, which allows us to choose a particular SD optimal policy $\psi_{\beta,1}$ that is consistent with Assumption A1. To proceed with the construction, we show the following Claims (1) and (2).

Claim (1) *For any $\theta_i \in (s, 1) \in \Theta_{\beta,1}$,*

$$v^*(\theta_i) \leq \max\left\{U_\beta(B), U_\beta(A) - \frac{(\bar{k}^{fifo} - s_{A\alpha} + 1)c}{p}\right\}.$$

Proof of Claim (1): Take a β-round, say player i, who is at the first in the queue for β-rounds in some period which we normalize to be $t_0 = 0$ and augmented state $(\mathbf{s}^0, 1) \in \Theta'_{\beta,1}$. Given ψ_A, to match with an A-square, player i must wait for at least $\bar{k}^{fifo} - s^0_{A\alpha} + 1$ additional arrivals of A-squares.

Consider now the following optimal stopping problem:

> **[P]** *A boy (β) stands under an apple (A) tree and holds a banana (B). In each period, one apple falls from the tree with probability p. The boy can consume exactly one piece of fruit, either an apple or the banana. He prefers an apple, with payoff $U_\beta(A)$, to the banana, with payoff $U_\beta(B)$. Thus, while he can consume the banana and walk away with $U_\beta(B)$ in any period, he may want to wait for falling apples. The first $\bar{k}^*(\equiv \bar{k}^{fifo} - s^0_{A\alpha})$ apples should be handed over to the owner of the tree. He incurs a cost c for each period of waiting, without consuming any fruit.*

Let $\Theta_{(P)} \equiv \{0, 1, \dots, \bar{k}^* + 1\} \cup \{\Delta\}$ denote the state space of [P], where Δ denotes the state after the boy consumes a piece of fruit. In each period t and state $\theta^t_{(P)} \in \Theta_{(P)} \backslash \{\Delta\}$, the boy chooses a demand $d \in \{A, B\}$. The stage payoff from demand A is either $U_\beta(A)$ in state $\theta^t_{(P)} = \bar{k}^* + 1$, or $-c$ in any other state in $\Theta_{(P)} \backslash \{\Delta\}$. The stage payoff from demand B is $U_\beta(B)$ in any state in $\Theta_{(P)} \backslash \{\Delta\}$. In state Δ (i.e., after consuming a piece of fruit), the boy gets zero stage payoff forever. The value of [P] with an arbitrary initial state $\theta \in \Theta_{(P)}$ is

$$v^*_{(P)}(\theta) \equiv \sup_\mu E_\mu \left[\sum_{t=0}^{\infty} u(d, \theta^t) : \theta^0 = \theta \right].$$ [47]

It is clear from the description of [P] that $v^*_{(P)}(0)$ constitutes an upper bound for the maximal expected total payoff of player i (i.e., $v^*(\theta_i)$). In fact, unlike player i, the boy in [P] can always consume a banana and walk away. Also, while player i must wait *at least* $\bar{k}^* + 1$ arrivals of A-squares to match with an A-square, the boy in [P] is guaranteed to get the $(\bar{k}^* + 1)$-th falling apple. As such, to prove the claim, it is sufficient to show that

$$v^*_{(P)}(0) \le \max \left\{ U_\beta(B), U_\beta(A) - \frac{(\bar{k}^{fifo} - s^0_{A\alpha} + 1)c}{p} \right\}. \tag{17}$$

Let

$$\bar{k}^{**} \equiv \left\lfloor \frac{c(U_\beta(A) - U_\beta(B))}{p} \right\rfloor \le \bar{k}^{fifo}.$$

(i) Suppose that $\bar{k}^* < \bar{k}^{**}$. Then, compared to consuming a banana immediately, it is weakly more profitable to wait until $\bar{k}^* + 1 = \bar{k}^{fifo} - s^0_{A\alpha} + 1 (\le \bar{k}^{**})$ apples fall. Once the boy decides to wait, he will continue to wait until he obtains an apple. Thus,

$$v^*_{(P)}(0) = U_\beta(A) - \frac{(\bar{k}^* + 1)c}{p} = U_\beta(A) - \frac{(\bar{k}^{fifo} - s^0_{A\alpha} + 1)c}{p}.$$

(ii) Suppose that $\bar{k}^* = \bar{k}^{**}$. As $v^*_{(P)}(\cdot)$ solves the optimality equation (*13*), we have

$$v^*_{(P)}(0) = \max \left\{ U_\beta(B), -c + p \left(v^*_{(P)}(1) \right) + (1-p) \left(v^*_{(P)}(0) \right) \right\}.$$

Suppose, toward a contradiction, that

$$v^*_{(P)}(0) = -c + p \left(v^*_{(P)}(1) \right) + (1-p) v^*_{(P)}(0) > U_\beta(B).$$

Then,

$$\begin{aligned} v^*_{(P)}(0) &= v^*_{(P)}(1) - \frac{c}{p} = \left(U_\beta(A) - \frac{\bar{k}^* c}{p} \right) - \frac{c}{p} \\ &= U_\beta(A) - \frac{(\bar{k}^{**} + 1)c}{p} > U_\beta(B). \end{aligned}$$

where the second equality follows from case (i) above (after the first apple falls, the boy needs to hand over only $\bar{k}^* - 1 (< \bar{k}^{**})$ additional apples to the owner). Notice that the last inequality contradicts to the definition of $\bar{k}^{**}$. Therefore, $v^*_{(P)}(0) \leq U_\beta(B)$.

(iii) Suppose that $\bar{k}^* > \bar{k}^{**}$. More apples should be handed over to the owner than the previous case, so $v^*_{(P)}(0) \leq U_\beta(B)$. This concludes the proof of Claim (1).

Claim (2) *There exists a SD optimal policy* $\psi_{\beta,1} : \Theta'_{\beta,1} \to \{A, B\}$ *of the MDP for player* i *such that*

$$\psi_{\beta,1}(\theta_i) = B, \quad \text{for all } \theta_i = (\mathbf{s}, 1) \in \Theta'_{\beta,1} \text{ with } s_{A\alpha} < 1.$$

Proof of Claim (2): Let $\psi_{\beta,1} : \Theta'_{\beta,1} \to \{A, B\}$ such that

$$\psi_{\beta,1}(\theta_i) = \begin{cases} A & \text{if } v^*(\theta_i) > U_\beta(B), \\ B & \text{if } v^*(\theta_i) \leq U_\beta(B). \end{cases}$$

Then, $v_i(\cdot; \psi_{\beta,1}) : \Theta'_{\beta,1} \to \mathbb{R} \cup \{-\infty\}$ is a solution of the optimality equation of the MDP for player i. It follows from Theorem Puterman (2) that the SD-policy $\psi_{\beta,1}$ is optimal.

By Claim (1), for any $\theta_i = (\mathbf{s}, 1) \in \Theta'_{\beta,1}$ with $s_{A\alpha} < 1$,

$$v^*_i(\theta_i) \leq \max\left\{U_\beta(B), U_\beta(A) - \frac{(\bar{k}^{fifo} + 1)c}{p}\right\} = U_\beta(B),$$

so that $\psi_{\beta,1}(\theta_i) = B$. This concludes the proof of Claim (2).

Step 2: Construction of $\psi_{\beta,q+1}$ *given* $(\psi_{\beta,1}, \psi_{\beta,2}, \ldots, \psi_{\beta,q})$

Fix $q \in \{1, 2, \ldots, \bar{k}^{fifo} + 1\}$. For a β-round, say player i, who enters as q-th in line, we extend the player's dynamic decision problem as a MDP with an arbitrary initial augmented state set $\Theta'_{\beta,\leq q} \equiv \bigcup_{q' \leq q} \Theta'_{\beta,q'}$. Note that the MDP for player i is defined by ψ_A, ψ_α, $\psi_{\beta,<q} \equiv (\psi_{\beta,q-1}, \psi_{\beta,q-2}, \ldots, \psi_{\beta,1})$, $\psi_{B,<q} \equiv (\psi_{B,q-1}, \psi_{B,q-2}, \ldots, \psi_{B,1})$, and by Assumption A1 applied to other players' strategies.

Inductive Hypothesis: There exists $\psi_{\beta,q'} : \Theta'_{\beta,q'} \to \{A, B\}$ for $q' \in \{1, 2, \ldots, q\}$ such that,

1. $\psi_{\beta,\leq q} = (\psi_{\beta,q}, \psi_{\beta,q-1}, \ldots, \psi_{\beta,1})$ is a SD optimal policy for player i, and

2. the maximal total expected payoff for player i, given any $\theta_i = (\mathbf{s}, q_i) \in \Theta'_{\beta,q}$ is

$$v_i^*(\theta_i) \leq \max\left\{U_\beta(B), U_\beta(A) - \frac{(\bar{k}^{fifo} - s_{A\alpha} + q_i)c}{p}\right\}.$$

Now, consider a β-round, say player j, who is $(q+1)$-th in line at some period $t_0 = 0$ (normalized). Player j solves a dynamic decision problem. As before, we extend the player j's problem as a MDP with an arbitrary initial augmented state in the set $\Theta'_{\beta,\leq q+1} \equiv \bigcup_{q' \leq q+1} \Theta'_{\beta,q'}$. Note that player j's MDP is defined by ψ_A, ψ_α, $\psi_{B,\leq q}$, $\psi_{\beta,\leq q}$, and by Assumption A1 applied to other players' strategies. As the set of augmented state for player j is still finite, there exists a SD optimal policy (see Theorem Puterman (3)). Moreover, any policy whose values solve the optimality equation, is optimal (see Theorem Puterman (2)). In particular, it is optimal for player j to follow any SD optimal policy of his MDP until his rank becomes q, after which he switches to any optimal policy of an MDP for a β-round who enters as q-th in line. In all, to find an optimal policy for player j, it is sufficient to find a function $\psi_{\beta,q+1} : \Theta'_{\beta,q+1} \to \{A, B\}$ that is consistent with Assumption A1.

Claim (3) For any $\theta_j = (\mathbf{s}, q+1) \in \Theta'_{\beta,q+1}$,

$$v_j^*(\theta_j) \leq \max\left\{U_\beta(B), U_\beta(A) - \frac{(\bar{k}^{fifo} - s_{A\alpha} + q + 1)c}{p}\right\}.$$

Proof of Claim (3) Take a β-round, say player j, who is $(q+1)$-th in the queue for β-rounds in period $t_0 = 0$ (normalized) and augmented state $(\mathbf{s}^0, q+1) \in \Theta'_{\beta,q+1}$. We make two observations:

1. (i) in any augmented state $\theta_j = (\mathbf{s}, q+1) \in \Theta'_{\beta,q+1}$, if there exists any $q' < q+1$ such that $\psi_{(\mathbf{s},q')} = B$, the maximum expected continuation payoff for player j is at most $U_\beta(B)$.

   (ii) in any augmented state $\theta_j = (\mathbf{s}, q+1)$ with $s_{A\alpha} = \bar{k}^{fifo} + 1$, the first β-round in the queue matches to an A-square. Thus, the maximum expected continuation payoff for player j (i.e., $v_j^*(\theta_j)$) equals to $v_j^*((\mathbf{s}', q))$ where $\mathbf{s}'$ denotes the augmented state after matching the first β-round with an A-square.[48] As $s'_{A\alpha} = \bar{k}^{fifo}$, by Induction Hypothesis,

$$v_j^*(\theta_j) = v_j^*(\mathbf{s}', q) \leq \max\left\{U_\beta(B), U_\beta(A) - \frac{qc}{p}\right\}.$$

Player j either matches with a B-square and receive $U_\beta(B)$ while his rank is $q+1$ or has an augmented state at some period before matching. Moreover, starting from an arbitrary initial augmented state $\theta_j^0 = (\mathbf{s}^0, q+1)$, the second case incurs only after $\bar{k}^{fifo} - s^0_{A\alpha} + 1$ arrivals of A-squares, at the minimum.

Consider the following optimal stopping problem:

[P'] *A boy (β) stands under an apple (A) tree and holds a banana (B). In each period, one apple falls from the tree with probability p. The boy can consume exactly one piece of fruit, either an apple or the banana. He (weakly) prefers an apple, with payoff $U'_\beta(A) \equiv \max\left\{U_\beta(B), U_\beta(A) - \frac{qc}{p}\right\}$, to the banana, with payoff $U_\beta(B)$. Thus, while he can consume the banana and walk away in any period, we may want to wait for falling apples. The first $\bar{k}^*(\equiv \bar{k}^{fifo} - s^0_{A\alpha})$ falling apples should be handed over to the owner of the apple tree. He incurs a cost c for each period of waiting, without consuming any fruit.*

Similar to the proof of Claim (1), let $\Theta_{(P')} \equiv \{0, 1, \ldots, \bar{k}^* + 1\} \cup \{\Delta\}$ denote the state space of [P'], where Δ denotes the state after the boy consumes a fruit. The value of [P'] with an arbitrary initial state $\theta \in \Theta_{(P')}$ exists (similar to the existence of the value of [P]).

The value of [P'] with the initial condition 0, denoted by $v^*_{(P')}(0)$, is an upper bound of the maximal expected payoff for player j. Unlike player j, the boy in [P'] can always consume a banana and walk away. While player j must wait $\bar{k}^* + 1$ arrivals of A-squares at the minimum to get a further expected continuation payoff $U'_\beta(A)$, the boy in [P'] is guaranteed to get $U'_\beta(A)$ after $\bar{k}^* + 1$ falling apples. As such, it is sufficient to prove that

$$v^*_{(P')}(0) \leq \max\left\{U_\beta(B), U_\beta(A) - \frac{\bar{k}^{fifo} - s^0_{A\alpha} + q + 1}{p}\right\}.$$

Let

$$\bar{k}^{**} \equiv \left\lfloor \frac{c(U'_\beta(A) - U_\beta(B))}{p} \right\rfloor \leq \bar{k}^{fifo}.$$

[48] That is, $(s'_A, s'_B, s'_\alpha, s'_\beta) = (s_A, s_B, s_\alpha, s_\beta) - (1, 0, 0, 1)$.

Similarly to the proof of Claim (1), we consider three cases:

(i) Suppose that $\bar{k}^* < \bar{k}^{**}$. Compared to consuming a banana immediately, it is weakly more profitable to wait until $\bar{k}^* + 1 = \bar{k}^{fifo} - s^0_{A\alpha} + 1 (\leq \bar{k}^{**})$ apples fall. Once the boy waits, he will continue to wait until he obtains an apple. Thus,

$$\begin{aligned} v^*_{(P')}(0) &= U'_\beta(A) - \frac{(\bar{k}^* + 1)c}{p} = U'_\beta(A) - \frac{(\bar{k}^{fifo} - s^0_{A\alpha} + 1)c}{p} \\ &\leq \max\left\{U_\beta(B), U_\beta(A) - \frac{(\bar{k}^{fifo} - s^0_{A\alpha} + q + 1)c}{p}\right\}. \end{aligned}$$

(ii) Suppose that $\bar{k}^* = \bar{k}^{**}$. As $v^*_{(P')}(\cdot)$ solves the optimality equation (see Theorem Puterman (1)),

$$v^*_{(P')}(0) = \max\left\{U_\beta(B), -c + p(v^*_{(P')}(1)) + (1-p)(v^*_{(P')}(0))\right\}.$$

Suppose, toward a contradiction, that

$$v^*_{(P')}(0) = -c + p(v^*_{(P')}(1)) + (1-p)(v^*_{(P')}(0)) > U_\beta(B).$$

Then,

$$\begin{aligned} v^*_{(P')}(0) &= v^*_{(P')}(1) - \frac{c}{p} \\ &= \left(U'_\beta(A) - \frac{\bar{k}^* c}{p}\right) - \frac{c}{p} = U'_\beta(A) - \frac{(\bar{k}^{**} + 1)c}{p} \\ &> U_\beta(B). \end{aligned}$$

where the second equality is from case (i). After the first falling apple, the boy needs to hand over only $\bar{k}^* - 1 (< \bar{k}^{**})$ additional apples to the owner. Notice that the last inequality contradicts to the definition of $\bar{k}^{**}$. Therefore, $v^*_{(P')}(0) \leq U_\beta(B)$.

(iii) Suppose that $\bar{k}^* > \bar{k}^{**}$. More apples should be handed over to the owner than the previous case, so $v^*_{(P')}(0) \leq U_\beta(B)$. This concludes the proof of Claim (3).

Claim (4) There exists $\psi_{\beta,q+1} : \Theta'_{\beta,q+1} \to \{A, B\}$ with

$$\psi_{\beta,q+1}(\theta_j) = B, \quad \text{for all } \theta_j = (\mathbf{s}, q+1) \in \Theta'_{\beta,q+1} \text{ with } s_{A\alpha} < q+1,$$

such that $\psi_{\beta,\leq q+1} = (\psi_{\beta,q+1}, \psi_{\beta,q}, \ldots, \psi_{\beta,1})$ (sequentially following as player j's rank moves up) is a SD optimal policy of the MDP for player j.

Proof of Claim (4) Let $\psi_{\beta,\leq q+1} : \Theta'_{\beta,q+1} \to \{A, B\}$ such that

$$\psi_{\beta,q+1}(\theta_j) = \begin{cases} A & \text{if } v_j^*(\theta_j) > U_\beta(B) \\ B & \text{if } v_j^*(\theta_j) \leq U_\beta(B). \end{cases}$$

Then, $v(\cdot; \psi_{\beta,\leq q+1}) : \Theta'_{\beta,\leq q+1} \to \mathbb{R} \cup \{-\infty\})$ solves the optimality equation of the MDP for player j. It follows from Theorem Puterman (2) that $\psi_{\beta,\leq q+1}$ is optimal.

By Claim (1), for any $\theta_j = (\mathbf{s}, q+1) \in \Theta'_{\beta,q+1}$ with $s_{A\alpha} < q+1$,

$$v_j^*(\theta_j) \leq \max\left\{U_\beta(B), U_\beta(A) - \frac{(\bar{k}^{fifo} + 1)c}{p}\right\} = U_\beta(B),$$

so $\psi_{\beta,q+1}(\theta_j) = B$. This concludes the proof of Claim (4).

To conclude the proof of Lemma A4, let us turn to Assumption A1. Thus far, we have constructed ψ_β, which subscribes a best-response for a β-square for any initial augmented state, given ψ_A, ψ_α, and Assumption A1 applied to strategies for each player's opponents. To conclude the proof we need to guarantee that the ψ_β we constructed satisfies Assumption A1. Take any stationary* strategy-profile $\Psi = (\psi_A, \psi_B, \psi_\alpha, \psi_\beta)$ such that ψ_A and ψ_α satisfy ((14)), and ψ_β and ψ_B are constructed as described above. Suppose that both B-squares and β-rounds exist in the market in a period t after a new pair arrives. We consider the case of $s^t_{A\alpha} \geq 0$ (and omit an analogous proof for the case of $s^t_{A\alpha} < 0$). For any β-round, say player i, with rank $q_i > s^t_{A\alpha}$, Claims (2) and (4) imply that player i would demand B. The counterpart results of Claims (2) and (4) for B-squares imply that every B-square, say player j, demands β as $q_j \geq 1 = 1 + s_{A\alpha,-}$. Therefore, Assumption A1 describes best response behavior. This concludes the proof of Lemma A4. ■

Proof of Lemma 1: This proof follows directly from Lemmas A3 and A4 above. ■

Proof of Lemma 2. We show that the second condition in Lemma A4 guarantees Lemma 2. Take a stationary* equilibrium Ψ^* satisfying conditions (a) and (b) in Lemma A4. Initially, there is no agent waiting in the market. Suppose that both a B-square and a β-round exist in some period t, for the first time ever. Given ((14)) (and a similar condition for ψ^*_α), it must be either (i) $s^t_A \geq 0$, $s^t_\alpha = 0$, $s^t_B = 1$, and $s^t_\beta = s^t_A + s^t_B$, or (ii) $s^t_A = 0$, $s^t_\alpha \geq 0$, $s^t_B = s^t_\alpha + s^t_\beta$, and $s^t_\beta = 1$. In both instances, there exist a B-square who finds no available α-round and demands β, and a β-round who finds no available A-square and demands B. As such, one pair of (B, β) will be matched, and only incongruent pair of agents (i.e., either A-squares and

β-rounds, or B-squares and α-rounds) wait until period $t+1$. A similar argument shows that in any period that a B-square and a β-round coexists, for the second time, third time, etc., one (B, β) pair will be formed. ■

Proof of Proposition 3: First, the (signed) length of the A-α queue, denoted by $k_{A\alpha}$ constitutes an ergodic Markov chain. Following arguments in the body of the paper, the unique steady state distribution of $k_{A\alpha}$ is the uniform distribution over $\{-\bar{k}^{dec}, -\bar{k}^{dec}+1, \ldots, \bar{k}^{dec}\}$. At any time t, suppose that $k^t_{A\alpha} > 0$. Clearly, the queue has no α-rounds. As equal numbers of squares and rounds enter and exit the market, it must be that $k_{A\alpha} + k_B = k_\beta$. Lemma 2 guarantees that $k_B = 0$, therefore $k_\beta = k_{A\alpha}$. Similarly for $k_{A\alpha} \leq 0$. ■

Proof of Corollary 3: Whenever $c > \frac{p(1-p)U}{2}$, the optimal mechanism matches arriving agents immediately, $\bar{k}^{opt} = 0$, and $\bar{k}^{opt} \leq \bar{k}^{dec}$. Suppose, then, that $c < \frac{p(1-p)U}{2}$. We then have

$$\sqrt{\frac{p(1-p)U}{2c}} < \frac{p(1-p)U}{2c} \leq \frac{pU}{2c} \leq \frac{p(U_A(\alpha) - U_A(\beta))}{c}.$$

and the result follows from the definitions of $\bar{k}^{opt}$ and $\bar{k}^{dec}$. ■

10 References

Akbarpour, Mohammad, Shengwu Li, and Shayan Oveis Gharan. 2017. "Thickness and Information in Dynamic Matching Markets," mimeo.

Anderson, Ross, Itai Ashlagi, David Gamarnik, and Yash Kanoria. 2015. "Efficient Dynamic Barter Exchange," mimeo.

Ashlagi, Itai, Patrick Jaillet, and Vahideh H. Manshadi. 2013. "Kidney Exchange in Dynamic Sparse Heterogenous Pools," mimeo.

Baccara, Mariagiovanna, Allan Collard-Wexler, Leonardo Felli, and Leeat Yariv. 2014. "Child-Adoption Matching: Preferences for Gender and Race," *American Economic Journal: Applied Economics*, **6(3)**, 133-158.

Becker, Gary S. 1974. "A Theory of Marriage: Part II," *The Journal of Political Economy*, **82(2)**, S11-S26.

Bloch, Francis and David Cantala. 2016. "Dynamic Allocation of Objects to Queuing Agents with Private Values," mimeo.

Budish, Eric, Peter Cramton, and John Shim. 2015. "The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response," *The Quarterly Journal of Economics*, forthcoming.

Burdett, Ken and Melvyn G. Coles. 1997. "Marriage and Class," *The Quarterly Journal of Economics*, **112(1)**, 141-168.

Choo, Eugene. 2015. "Dynamic Marriage Matching: An Empirical Framework," *Econometrica.*, **83(4)**, 1373-1423.

Doval, Laura. 2016. "A Theory of Stability in Dynamic Matching Markets," mimeo.

Eeckhout, Jan. 1999. "Bilateral Search and Vertical Heterogeneity," *International Economic Review*, **40(4)**, 869-887.

Feller, William. 1972. *An Introduction to Probability Theory and Its Applications, Volume II* (2nd ed.). New York: John Wiley & Sons.

Fershtman, Daniel and Alessandro Pavan. 2016. "Matching Auctions: Experimentation and Cross-Subsidization," mimeo.

Gjertson, David W. 2004. "Explainable Variation in Renal Transplant Outcomes: A Comparison of Standard and Expanded Criteria Donors," *Clinical Transplants*, **2004**, 303-314.

Gurvich, Itai and Amy Ward. 2014. "On the Dynamic Control of Matching Queues," *Stochastic Systems*, **4(2)**, 479-523.

Haeringer, Guillaume and Myrna Wooders. 2011. "discretionary Job Matching," *International Journal of Game Theory*, **40**, 1-28.

Hall, Robert E. and Alan B. Krueger. 2012. "Evidence on the Incidence of Wage Posting, Wage Bargaining, and On-the-Job Search." *American Economic Journal: Macroeconomics*, **4(4)**, 56-67.

Herbst, Holger and Benjamin Schickner. 2016. "Dynamic Formation of Teams: When Does Waiting for Good Matches Pay Off," mimeo.

Hitch, Gunter J., Ali Hortacsu, and Dan Ariely. 2010. "Matching and Sorting in Online Dating," *The American Economic Review*, **100(1)**, 130-163.

Hu, Ming and Yun Zhou. 2016. "Dynamic Type Matching," mimeo.

Kocer, Yilmaz. 2014. "Dynamic Matching and Learning," working slides.

Leshno, Jacob. 2015. "Dynamic Matching in Overloaded Waiting Lists," mimeo.

Loertscher, Simon, Ellen V. Muir, and Peter G. Taylor. 2016. "Optimal Market Thickness and Clearing," mimeo.

Niederle, Muriel and Leeat Yariv. 2009. "discretionary Matching with Aligned Preferences," mimeo.

Oien, Cecilia M., Anna V. Reisaeter, Torbjørn Leivestad, Friedo W. Dekker, and Pål-Dag Line. 2007. "Living Donor Kidney Transplantation: The Effects of Donor Age and Gender on Short- and Long-term Outcomes," *Transplantation*, **83(5)**, 600-606.

Pais, Joana V. 2008. "Incentives in discretionary random matching markets," *Games and Economic Behavior*, **64**, 632-649.

Puterman, Martin L. 2005. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, Wiley.

Rogerson, Richard, Robert Shimer, and Randall Wright. 2005. "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, **XLIII**, 959-988.

Ross, Sheldon M. 2014 *Introduction to Stochastic Dynamic Programming*, Academic Press.

Satterthwaite, Mark and Artyom Shneyerov. 2007. "Dynamic Matching, Two-sided Incomplete Information, and Participation Costs: Existence and Convergence to Perfect Competition," *Econometrica*, **75(1)**, 155-200.

Schummer, James. 2016. "Influencing Waiting Lists," mimeo.

Stein, Rob. 2011. "Under Kidney Transplant Proposal, Younger Patients would Get the Best Organs," *The Washington Post*, February 24.

Taylor, Curtis. 1995. "The Long Side of the Market and the Short End of the Stick: Bargaining Power and Price Formation in Buyers', Sellers', and Balanced Markets," *The Quarterly Journal of Economics*, 110(3), 837-855.

Tyurin, Ilya S. 2010. "An improvement of upper estimates of the constants in the Lyapunov theorem," *Russian Mathematical Surveys*, **65(3(393))**, 201-202.

Ünver, Utku. 2010. "Dynamic Kidney Exchange," *Review of Economic Studies*, **77**, 372-414.

Zenios, Stefanos A. 1999. "Modeling the transplant waiting list: A queueing model with reneging," *Queueing Systems*, **31**, 239-251.