

Match Choice and Ghettoization in Evolutionary Games

Nathan Larson*

November 22, 2000

Abstract

Matching is endogenized in an otherwise standard adaptive model of play in 2x2 games by giving agents the option to sometimes continue playing against the same opponent. This renewal option allows agents playing efficiently to isolate themselves from those who are not. Perhaps counterintuitively, efficient play may be less likely to survive in the long run when agents have this additional instrument, even in games with common interests. This is because isolation has two effects: it raises the returns to an efficient but fragile strategy, but it also “ghettoizes” agents playing inefficiently - they rarely learn about the efficient strategy and when they do learn about it, they rarely hear good news. I look at an extension in which agents have long memories about the performance of a strategy. With long memories, good news about an efficient strategy will be more likely to trickle down to ghettoized agents, mitigating the bias in learning. With this bias removed, long run survival of efficient play becomes more likely - even when it is not a static Nash equilibrium. Essentially, agents learn to use the renewal option to punish non-cooperators. However, the speed of learning may still be quite slow, so that ghettoization persists for a long time.

1 Introduction

This paper asks how agents in large populations learn to play games when they have some control over who they play against. Evolutionary models have already shed much light on how simple adaptive dynamics can lead a large population of agents who repeatedly play a game in pairs toward an equilibrium, but for the most part, these models assume that an exogenous process determines which agents play against each other. However, *who* to play is often as important a strategic decision as *how* to play. I build an evolutionary model in which matching is random but agents influence who they play by controlling the length of matches. This has the potential to benefit efficient strategies by allowing agents playing them to partially isolate

*I'd like to thank the people whose criticism and encouragement have kept this project sailing straight. These include but are not limited to Bengt Holmstrom, Sendhil Mullainathan (who inspired the example of the introduction), Thomas Philippon, Pieter Gautier, Daron Acemoglu, and particularly my advisors, Abhijit Banerjee and Glenn Ellison.

themselves from strategic risk. However, isolation is a double-edged sword – agents who are unsuccessful in isolating may be worse off than under random matching, and the behavior of a successful agent may be learned and imitated more slowly if she is isolated. I show that when agents judge their strategies by their recent experience, the latter two effects dominate, in some cases preventing the population from learning an efficient equilibrium that would have been reached under random matching. In contrast, when agents judge strategies on their long term performance, flexible matching always improves long run payoffs and in some cases makes it possible to sustain payoffs higher than those in any Nash equilibrium. These results provide evolutionary foundations for phenomena like the persistence of ghettos and the emergence of cooperative behavior.

My model focuses on large populations in which agents are repeatedly matched to play a 2×2 game. Matching is random, but sometimes agents are given the chance to continue playing against each other rather than rematching. An agent decides whether to renew a match by comparing the payoff she received in that match to other payoffs that she has received in the past. Only when both agents agree does a match have a chance of being renewed. One can think, for example, of firms renewing a contract that has been profitable, or of academics who continue to collaborate after one joint project has gone well.

Meanwhile, at a slower rate, agents can adjust their strategies. This adjustment process has two steps. Agents learn about different strategies by observing and talking to the agents they meet in the matching process. Through this interaction, an agent sometimes learns her opponent’s assessment of his strategy – that is, she learns about some of the payoffs he has received in the past. If the strategy he is playing is different from her own, she compares the performance of the two strategies, using her own past payoffs, and adopts the strategy that appears to be performing better. One may interpret the slow rate of adjustment to reflect implicit friction in changing strategies, perhaps because strategies involve fixed investments or because particular behaviors are habit-forming.

In my *standard model*, an agent assesses her strategies myopically, considering only the payoff she received in the match immediately before her current one. In this context, allowing agents to extend good matches has results that might appear surprising. My first result shows that in coordination games, flexible matching expands the set of initial conditions from which the dynamics converge to the risk dominant equilibrium, even if it is not efficient. In stag hunt games the results are even stronger: the dynamics tilt toward the inefficient equilibrium and always select it as the chance of extending a good match grows large.¹

The intuition behind these results is best presented in an example. Think of molecular biologists repeatedly collaborating on projects in pairs. Each biologist specializes in one of two

¹Unfortunately, there is something of a coordination problem over what the terms “coordination game” and “stag hunt game” mean. My definitions are presented with the model in Section 2.

research techniques: using newfangled polymerase chain reactions (*PCR*) or using standard test tube methods (*TT*). *PCR* is the more efficient technique when both of the paired researchers use it, but it is fragile: when a *PCR* researcher works with a researcher trained in standard methods, the latter is likely to bungle things that the former will have to fix. Standard test tube methods produce less spectacular results but are more robust to these mismatches. The payoffs are summarized by the following coordination game with (pure strategy) Nash equilibria at (*PCR*, *PCR*) and (*TT*, *TT*), where (*PCR*, *PCR*) is efficient but (*TT*, *TT*) is risk dominant:²

	<i>PCR</i>	<i>TT</i>
<i>PCR</i>	4, 4	0, 2
<i>TT</i>	2, 0	3, 3

When matching is random, using *PCR* researchers earn higher average payoffs precisely when the fraction of *PCR* researchers is greater than 3/5. Each type of researcher learns the payoff to the other technique, and the mechanics of how to use it, by working with researchers who currently use it, so the population drifts in the direction of whichever technique is currently earning higher returns.

When researchers have the option to renew matches, they will try to extend projects with other researchers using the same technique, so the average payoff to each technique will rise relative to random matching – more so for *PCR* which has more to gain from coordination. However, there is another change that is crucial. First, matching becomes serially correlated: because of the possibility of renewal, biologists in *PCR* – *PCR* matches are more likely than average to have been in *PCR* – *PCR* matches recently (and similarly for biologists in *TT* – *TT* matches), while biologists in *PCR* – *TT* matches are more likely than average to have come from “mixed” matches (because other biologists using the same technique are likely to be locked up in long-term matches). Only biologists in these mixed matches are exposed to a different technique, and because of this serial correlation, these biologists tend to have seen the worst side of their own methods lately. As a result, learning favors the standard method because it is more robust to miscoordination. The biologists that could attest to the benefits of using *PCR* reap those benefits because they are isolated from biologists using *TT* – which is precisely why biologists using *TT* rarely learn from them.

If the payoffs to using *TT* are reversed, so that biologists using standard methods do better working with colleagues using the more advanced *PCR* technique, than they do with each other,

²A strategy in a symmetric 2×2 game is risk dominant if it is a best response to a 50 – 50 mixture by one’s opponent. In this context, an equilibrium is risk dominant if it has both agents playing the risk dominant strategy.

then we have a stag hunt game:

	<i>PCR</i>	<i>TT</i>
<i>PCR</i>	4, 4	0, 3
<i>TT</i>	3, 0	2, 2

Now, *TT* researchers would like to renew matches with *PCR* researchers, but are spurned by them. Because renewal decision require joint approval, *PCR* researchers are still able to isolate themselves. However, when a *PCR* researcher does find herself between good matches, the waiting time to find another colleague using *PCR* will be long because she must screen through many *TT* researchers first – more than in the first example, because now none of the standard methods people will be locked up in long-term matches with each other. As a result, when she meets and compares notes with someone using standard methods, she is more likely than before to have dealt with *TT* types recently, so she will be more likely be discouraged with the performance of her technique and pass on negative information it. Thus, proliferation of the more efficient technique will be even less likely in the long run.

The next results are on the speed of learning. Because learning depends on interaction between agents using different strategies, flexible matching has the potential to either accelerate or slow learning, depending on whether agents choose to spend more or less time in cross-strategy interaction. When learning is slow, convergence to a rest point may not happen within any reasonable time frame, and the system may be better characterized by its medium run behavior. I show that in coordination games, agents using the less efficient strategy will be “ghettoized” in the medium run – for endogenous reasons they are almost never exposed to the success of the efficient strategy. For example, those of my biologists who have learned the new, more efficient *PCR* technique are not interested in collaborating with *TT* types to help them learn about a better method – on the contrary, they avoid them lest their experiments be bungled, consigning the *TT* researchers to work mainly with each other and learn about new developments only slowly.

I next consider an extension, the *long memory model*. In this model, an agent uses a long stream of past payoffs to assess her strategy, rather than her most recent payoff, but otherwise, the specification of the model is the same. In this model, flexible matching has dramatically different implications for the long run. Flexibility increases the chance that the population will converge to a more efficient rest point in all 2×2 games. Furthermore, when matching is flexible enough, a dominated but socially efficient strategy, such as cooperation in the prisoner’s dilemma, can survive in the long run. As matching becomes completely flexible, I obtain a selection result: in the long run agents always earn the greatest lower payoff in the game (that is, the maximum over the lesser of the two payoffs earned at each profile of actions). In particular, the efficient equilibrium is selected in coordination games, and cooperation is

selected in the prisoner's dilemma.

The different results arise because when agents have long memories, their assessments average out the cyclicity of payoffs. In the language of the earlier example, this means that *PCR* researchers currently mismatched with *TT* researchers are no less sanguine about their technique than well-matched *PCR* researchers, because both groups have similar long run experiences. This tends to distribute good news about the efficient technique, which in the standard model was locked up in *PCR – PCR* matches, more evenly across the population. When this bias in learning is removed, the main effect of the option to renew matches is to give agents an instrument with which to coordinate on efficient strategies and avoid inefficient and predatory ones, thus moving the dynamics in the direction of higher long run payoffs.

However, these results come with a caveat attached: the speed of learning effects in the standard model carry over wholesale to the long memory model. As a result, the salutary long run effects of flexibility when agents have long memories may come at the cost of a long period of ghettoization.

My example of biologists ghettoizing colleagues who have not yet learned a new technique lends itself well to the pairwise interaction structure of my model, but the effects described have wide applicability. I will just cite a few examples:

- Firms may not be exposed to an innovation that could streamline their supply chains because suppliers and customers who have adopted it find it inefficient to work with firms that have not.
- Adolescents in poor neighborhoods may form skewed beliefs about the returns to education if successful students have incentives to leave the neighborhood but those who fail come back to relate their experience.
- A poor tennis player may find it difficult to improve his game because more talented players would rather play with people on their own level.

In each of these cases, from the momentous to the mundane, the widespread long run adoption of an efficient strategy may be hindered or prevented rather than helped by giving agents more control over who they interact with. My model should shed insight on a broad class of these sorts of situations.

This paper lies at a crossroads. A substantial literature follows the seminal contributions of Kandori, Mailath, and Rob (1993) (hereafter KMR) and Young (1993), among others, in attempting to explain equilibrium selection in large populations by modeling agents' strategies as an adaptive dynamical system, while usually assuming that matching is exogenous. In contrast, the literature on search and matching explores the implications of giving agents varying amounts of control over who they interact with but typically takes their types to be fixed. In

giving my agents both of these strategic choices – they can change their strategies and change their partners – I will rely on insight garnered from both of these literatures.

The closest models to mine in the evolutionary game theory literature are Ely (1996) and Sethi (1998). Ely makes the point that when agents have both a high degree of control over who they interact with and good information about the state of play, they can identify and selectively match with agents playing efficiently, thus making efficient play robust. My model also endogenizes interaction, but in a quite different way: my agents will only gather and act on information locally (through the agents they interact with). Not surprisingly, the spectrum of behavior predicted by my model will be quite different; in particular, the prospects for efficiency are generally much bleaker. Sethi studies an evolutionary model in which the rate at which a strategy is adopted depends not only on its payoff but also on how hard it is to learn. He demonstrates with a series of counterexamples that several weathered results of the evolutionary literature can fail in this case. I wholeheartedly concur with his emphasis on thinking more concretely about how strategies are learned; my model advances this agenda by showing how the difficulty of learning different strategies can emerge endogenously from the decisions of individual agent about who to interact with. My specification of learning is similar in spirit to recent models of social learning like Ellison and Fudenberg (1993 and 1995) and Bala and Goyal (1998), and one could loosely interpret my model as describing social learning of an activity with spillovers. Finally, this work is motivated in part by the empirical literature on neighborhood effects; in certain cases, my model will generate the sort of stratification that this literature discusses.

The rest of the paper proceeds as follows. Section 2 presents the model. Section 3 derives results under flexible matching for several classes of 2×2 games. Section 4 presents an extension in which agents have long time horizons and shows that the role of flexibility in facilitating coordination is recovered. In some cases, cooperative outcomes outside of the set of static Nash equilibria are sustained in the long run. Section 5 discusses the results in the context of related work and Section 6 concludes.

2 The Model

The underlying game has two players with identical strategy sets $S = \{H, L\}$ and symmetric payoffs π_{ij} , $i, j \in S$, given by the game form below, where π_{LL} has been normalized to 0.

	<i>H</i>	<i>L</i>
<i>H</i>	<i>a, a</i>	<i>c, d</i>
<i>L</i>	<i>d, c</i>	<i>0, 0</i>

Throughout the paper I will require these payoffs to be generic:

Condition 1 (GP) No two payoffs are equal: $a \neq c$, $c \neq d$, and $a \neq d$.

Additionally, I will always take $a > 0$ without loss of generality, so that (H, H) Pareto dominates (L, L) .

There is a continuum of agents (of measure 1) who are repeatedly matched in pairs to play the game above in continuous time. (So the payoffs listed above are flow payoffs.) At time t , the measure of agents playing strategy i is x_i^t , with $x^t = (x_H^t, x_L^t)$. Friction in changing strategies means that x^t adjusts slowly, and for brevity I will sometimes refer to agents of type i when I mean agents currently playing i ; it should be clear that these “types” are by no means permanent.

Each agent forms an assessment of the returns she expects from using her current strategy; this is her current *perceived payoff*. I assume that an agent’s perceived payoff is equal to the flow payoff she received in her last match before the current one. This assumption captures the idea that she bases her expectations on her recent experiences with the strategy. There are several reasons that agents may behave this way. For example, they may be myopic or have limited memory. Or even if they care about the future, they may doubt their ability to predict it accurately and so settle for making fairly accurate short run forecasts based on the recent past. The set of possible perceived payoffs is $\{a, c\}$ for agents playing H and $\{0, d\}$ for agents playing L , and I denote by π_i^h the perceived payoff of an agent playing strategy i with history h (that is, whose last opponent played $h \in \{H, L\}$).

The full model consists of two additional elements: a matching process (MP) that describes how agents cycle into and out of matches, and adaptive dynamics (AD) that describe how agents learn about different strategies and sometimes switch to them. I describe these in turn.

2.1 The Matching Process

I define the state of matching $m^t = (m_{HH}^t, m_{HL}^t, m_{LH}^t, m_{LL}^t)$, where m_{ij}^t is the measure of agents of type i currently matched with opponents of type j . We have $m_{iH}^t + m_{iL}^t = x_i^t$ and $\sum_{i,j} m_{ij}^t = 1$, so agents are never unmatched. Furthermore, we have $m_{HL}^t = m_{LH}^t$ because matching is one-to-one. Together, these imply that for given population shares x^t , the state of matching can be completely characterized by m_{HL}^t , the measure of agents of each type in mixed matches. Agents will control matching entirely through the option to renew matches that both agents are pleased with. The matching process works in the following way. All matches dissolve independently at Poisson rate σ . When a match dissolves, two things happen instantaneously before the agents are rematched. First, each of the two agents chooses either to *renew* the match or to *exit*. These decisions are simultaneous and independent. Then each agent’s perceived payoff is updated to her flow payoff from the match that just ended. Next the agents are rematched. If both opted to renew, then they are rematched together with

probability $1 - 1/s$, where $s \geq 1$. If either agent opted to exit, or if both opted to renew but failed (with probability $1/s$), then the agents flow into a rematching pool consisting of all of the agents whose matches just ended and were not successfully renewed. There they are immediately and randomly rematched.³ For now, assume that all agents playing i make the same decision about whether to renew ij matches, for all i and j , and that this decision is constant over time. (Later I will show this to be true when agents choose optimally.) Then we can write σ_{ij} for the exit rate at which members of ij matches flow out of unrenewed matches and into the rematching pool, where σ_{ij} takes one of the two values σ or σ/s .

I assume that strategy adjustment is much slower than the matching process, so that x^t can be taken as fixed while m^t adjusts to its steady state.⁴ Then the state of matching evolves according to

$$\dot{m}_{ij}^t = -\sigma_{ij}m_{ij}^t + y_i^t \frac{y_j^t}{Y^t} \quad (1)$$

The first term represents the rate at which ij matches end. In the second term, y_i^t is the total flow of i 's into rematching; of these, a fraction y_j^t/Y^t rematch with a j , where Y^t is the total flow of all agents into rematching:

$$\begin{aligned} y_H^t &= \sigma_{HH}m_{HH}^t + \sigma_{HL}m_{HL}^t \\ y_L^t &= \sigma_{LL}m_{LL}^t + \sigma_{LH}m_{LH}^t \\ Y^t &= y_H^t + y_L^t \end{aligned}$$

Four equations are described by (1), but because matching is complete and balanced, only one is independent. Expressed in terms of \dot{m}_{HL}^t , matching is characterized by

$$\dot{m}_{HL} = \frac{(1 - S)m_{HL}^2 - m_{HL} + x_H x_L}{D} \quad (MP)$$

³Because the measure of agents in the rematching pool is 0, the law of large numbers does not immediately ensure either that the proportions of agents of each type in the pool are exactly given by their flow rates, or that random matching within the pool produces exactly the expected number of matches of each type. I make the convention that both of these equivalences hold. This assumption could be rigorously derived by starting with a finite number of agents and discrete time and taking limits in the appropriate order: first letting the number of agents go to infinity, and then letting period length go to zero.

⁴This approach allows me to work with a one dimensional system subject to a constraint, permitting complete analytical results to be derived. I expect that a two dimensional system in x^t and m^t would have the same qualitative features.

where

$$S = \frac{\sigma_{HL}^2}{\sigma_{HH}\sigma_{LL}}$$

$$D = \frac{\sigma_{HH}(x_H - m_{HL}) + \sigma_{LL}(x_L - m_{HL}) + 2\sigma_{HL}m_{HL}}{\sigma_{HH}\sigma_{LL}} > 0$$

and time superscripts have been dropped for clarity. The next proposition demonstrates that holding x^t fixed, matching converges to a steady state.

Proposition 1 (*MP*) has a globally asymptotically stable steady state $m^*(x, S)$ given by

$$m_{HL}^*(x, S) = \frac{\sqrt{1 + 4(S-1)x_H x_L} - 1}{2(S-1)}$$

$$m_{HH}^*(x, S) = x_H - m_{HL}^*$$

$$m_{LL}^*(x, S) = x_L - m_{HL}^*$$

Proof. Appendix. ■

Where the dependence of the steady state on x and S does not need to be emphasized, I will sometimes write just m^* .

2.2 The choice to exit or renew

Now I specify how agents decide whether to renew or not and show that the characterization above will be consistent with their choices. An agent opts to renew matches in which she earns the larger of the two payoffs to her strategy and exits from matches in which she earns the lesser of the two payoffs. The motivation is that an agent earning her worst payoff always expects to do at least weakly better outside her current match, while an agent earning her best payoff expects to do weakly worse. With these decision rules, the exit rate from an ij match depends only on the flow payoffs a , c , and d , and so is constant across agents and over time as assumed earlier. The exit rates are given by

$$\begin{aligned} \sigma_{HH} &= \sigma (= \sigma/s) \text{ if } a < c \text{ (if } a > c) \\ \sigma_{LL} &= \sigma (= \sigma/s) \text{ if } d > 0 \text{ (if } d < 0) \\ \sigma_{HL} &= \sigma (= \sigma/s) \text{ if } a > c \text{ or } d < 0 \text{ (if } c > a \text{ and } d > 0) \end{aligned} \tag{2}$$

The following game forms illustrate how this decision rule is exercised. In each game, pairings that are not renewed are flagged with an R , a C , or both, depending on who opts to exit – the row player, the column player, or both.

	H	L	
H	1, 1	-3, -1 ^{RC}	
L	-1, -3 ^{RC}	0, 0	
	(A)		

	H	L	
H	2, 2	-3, 1 ^R	
L	1, -3 ^C	0, 0 ^{RC}	
	(B)		

	H	L	
H	1, 1 ^{RC}	3, 2	
L	2, 3	0, 0 ^{RC}	
	(C)		

2.3 Adaptive Dynamics

I complete the description of the model by specifying how agents learn about different strategies and how they change strategies. My agents will learn about strategies they are not currently playing through exposure to them. Specifically, at Poisson rate μ , the partners in a match learn each other's perceived payoff. If both agents are playing the same strategy, then neither takes any action based on this information. However, if the agents are playing different strategies, the agent with the lower perceived payoff switches to the strategy of the agent with the higher perceived payoff after the current match ends with probability $|\pi_i^k - \pi_j^l|/L$ (where π_i^k and π_j^l are the perceived payoffs of the two agents, and L is a constant greater than the maximum difference between payoffs.) Thus, we have the following dynamical system:

$$\dot{x}_H^t = \lambda \sum_{k \in \{H,L\}} \sum_{l \in \{H,L\}} m^t(H_k, L_l) (\pi_H^k - \pi_L^l) \quad (3)$$

Here, $\lambda = \mu/L$, and $m^t(H_k, L_l)$ is the measure of matches at time t between agents playing H with perceived payoff π_H^k and agents playing L with perceived payoff π_L^l . Note also that π_H^k , the perceived payoff of an agent whose last opponent played k , is equal to π_{Hk} , the flow payoff to playing H against k . Define $m^*(H_k, L_l|x^t, S)$ to be the value of $m^t(H_k, L_l)$ in the steady state of (MP) . The adaptive dynamics are defined by (3), subject to the constraint that (MP) always be at its steady state:

$$\dot{x}_H^t = \lambda \sum_{k \in \{H,L\}} \sum_{l \in \{H,L\}} m^*(H_k, L_l|x^t, S) (\pi_{Hk} - \pi_{Ll}) \quad (AD)$$

This formulation embodies several assumptions. First, strategy change is slow. The model relies on this in two ways: agents cannot change strategies within a match, and the state of matching can be approximated by its steady state. The first is crucial; if matched agents can quickly renegotiate to efficient play, then the coordination problem is moot. The second is purely for tractability; if we were to instead model the coevolution of m^t and x^t as a two dimensional system, the qualitative conclusions would be very similar. There are many contexts in which short run commitment to strategies seems plausible, for example if behavior is habit-forming or if strategies involve investment.

Second, learning is very primitive. An agent considers only the most rudimentary local

information when making decisions – her own payoff history and her partner’s payoff history – and uses that information in a very simple way. At first glance, this assumption appears both attractive and dangerous. It is attractive because agents are not required to know anything about the game form or about the global distribution of play, so none of the results will depend on agents being (perhaps unreasonably) well-informed. The apparent danger is that the results may be driven by taking agents’ decisions to be *too* simple and uninformed; I will address this concern in Section 3.3.

3 Characterization and Results

This section will work through the implications of the model. First, I examine the benchmark case of random matching and show that the dynamics collapse to a familiar workhorse of evolutionary game theory, the replicator dynamics. Then, after developing some intermediate results about steady state matching, I demonstrate that the system with flexible matching can be analyzed by looking at a parallel system in which matching is random but the population shares of H and L are biased. This technique is used to characterize the effect of flexible matching for several classes of games. The results extend the intuition developed in the introduction; for example, in coordination games, flexibility will favor the risk dominant equilibrium, while in stag hunt games, flexibility will favor the inefficient equilibrium.

The interesting action in (AD) arises through the possibility that the agents who learn from each other have systematically different histories than the population at large. In this case, the dynamics will not reflect the average payoff difference between strategies because disproportionate weight will be placed on particular payoff differences. The next section shows that the converse is true: when presence in a HL match is uncorrelated with one’s past matches, as is the case under random matching, the dynamics can be summarized by the difference between average payoffs.

The Random Matching Benchmark In steady state under random matching, an agent’s current match is uninformative about her history – the chance that her last opponent played H is always just x_H^t . Then the probability that two matched agents have histories k and l is just $x_k^t x_l^t$, regardless of their current strategies, so in particular, the fraction of HL matches with histories kl is just $x_k^t x_l^t$, and $m^*(H_k, L_l | x^t, 1) = x_k^t x_l^t m_{HL}^*(x^t, 1)$. Furthermore, the steady state measure of mixed matches is just $m_{HL}^*(x^t, 1) = x_H^t x_L^t$. Using these facts (and dropping

the time index), we can rewrite (AD):

$$\dot{x}_H = \lambda m_{HL}^* \sum_{k \in \{H,L\}} \sum_{l \in \{H,L\}} x_k x_l (\pi_{Hk} - \pi_{Ll}) \quad (4)$$

$$\begin{aligned} &= \lambda m_{HL}^* \left(\sum_{k \in \{H,L\}} x_k \pi_{Hk} (x_H + x_L) - \sum_{l \in \{H,L\}} x_l \pi_{Ll} (x_H + x_L) \right) \\ &= \lambda m_{HL}^* (\pi_H - \pi_L) \\ &= \lambda x_H x_L (\pi_H - \pi_L) \end{aligned} \quad (5)$$

where $\pi_H = x_H \pi_{HH} + x_L \pi_{HL}$ and $\pi_L = x_L \pi_{LL} + x_H \pi_{LH}$ are the average payoffs earned by agents using strategies H and L . It is important to recognize that no agents actually use these average payoffs when making a decision. Instead, agents compare two short run assessments of payoffs, each of which is either too optimistic or too pessimistic when viewed from the perspective of average payoffs. Then, from this perspective, some agents will be switching strategies too eagerly, some not eagerly enough, and some may be switching to the “wrong” strategy. However, when matching is random, these effects tend to cancel out, and the aggregate dynamics depend only on average payoffs.⁵ Writing the average payoff over the entire population as $\bar{\pi} = x_H \pi_H + x_L \pi_L$, and observing that $\pi_H - \bar{\pi} = x_L (\pi_H - \pi_L)$, (5) becomes

$$\frac{\dot{x}_H}{x_H} = \lambda (\pi_H - \bar{\pi}) \quad (6)$$

which is just the familiar replicator dynamics. When the underlying game has pure strategy equilibria at (H, H) and (L, L) , we can define $p = -c/(a - c - d)$, the value of x_H at which average payoffs are equalized. The right-hand side of (5) is then continuous, positive when $x_H > p$, and negative when $x_H < p$, so trajectories that begin above p converge to $x_H = 1$ (an H -convention) and trajectories that begin below p converge to $x_H = 0$ (an L -convention). When the game has asymmetric pure strategy equilibria, the mixed profile at $x_H = p$ is stable and attracts all interior trajectories.

Properties of Flexible Matching The following results build intuition about the matching process under flexibility and will be useful in characterizing the model. Unless otherwise noted, the matching process will be assumed to be at its steady state in the sequel, so I will drop the superscript star and write $m(x^t, S)$ or just m where attention to the arguments is not

⁵This feature hinges on the linearity of the dynamics in payoffs. Under many other choices of dynamics (for example, most versions of best reply), this would not be true. Then there would be two effects distinguishing the system with flexible matching from one determined by aggregate average payoffs: the pure stochastic effect of basing learning on individual payoffs, and the additional effect supplied by flexible matching. An advantage of this specification is that it allows the latter to be isolated.

needed. From (1), the steady state condition is

$$y_i y_j = \sigma_{ij} m_{ij} Y \quad \text{for all } i, j \in \{H, L\} \quad (7)$$

Squaring both sides, we have

$$\begin{aligned} (y_i y_i)(y_j y_j) &= (\sigma_{ij} m_{ij})^2 Y^2 \\ (\sigma_{ii} m_{ii} Y)(\sigma_{jj} m_{jj} Y) &= (\sigma_{ij} m_{ij})^2 Y^2 \\ \sqrt{(\sigma_{HH} m_{HH})(\sigma_{LL} m_{LL})} &= \sigma_{HL} m_{HL} \text{ or} \\ m_{HL} &= \sqrt{m_{HH} m_{LL} / S} \end{aligned} \quad (8)$$

letting $i = H$, $j = L$, and as earlier, defining $S = \sigma_{HL}^2 / \sigma_{HH} \sigma_{LL}$. It is then straightforward that

$$\begin{aligned} m_{HH} + \sqrt{m_{HH} m_{LL} / S} &= x_H \\ m_{LL} + \sqrt{m_{HH} m_{LL} / S} &= x_L \end{aligned} \quad (9)$$

Depending on the exit decisions, S will take one of the three values $\{s^2, s, 1/s^2\}$. When S equals s^2 or s at least one type of agent prefers to exit from ‘‘mixed’’ matches, as in examples (A) and (B) above. In these cases, as matching becomes more flexible (as s grows larger), S grows, and the measure of agents in mixed matches shrinks. The case $S = s$ reflects a game like (B) in which agents of types H and L disagree about whether to exit mixed matches. We can see from (8) that the unilateral decision to exit dominates, but it is tempered: mixed matches decline at a slower rate with s than they do when both types avoid them. One might wonder why this is true; after all, in this model a match does not dissolve any faster if both parties are displeased than if one is. The difference lies in the fact that the type that likes mixed matches (say agents playing L) exits from every match (sometimes by choice, and sometimes not). This means that L 's flood the rematching market. Consequently, as H 's slowly exit from HH matches, they must run a gauntlet of L 's before finding another good match, driving the rate of mixed matches up somewhat. The next lemma formalizes several properties of the matching process that will be useful.

Lemma 1 *The following are true at the steady state of (MP):*

1. $m_{HL} = 0$ if $x_H = 0$ or $x_L = 0$.
2. m_{HL} is increasing in $x_H x_L$, and therefore increasing in x_H if $x_H < 1/2$ and decreasing in x_H if $x_H > 1/2$.

3. m_{HL}/x_H is decreasing in x_H and m_{HL}/x_L is decreasing in x_L (strictly decreasing when $x_H x_L > 0$). That is, the fraction of time spent playing against the other strategy decreases as an agent's own strategy becomes more prevalent.
4. $\lim_{x_i \rightarrow 0} m_{HL}/x_i = 1$. (As a strategy vanishes, it never plays against itself.)
5. m_{HL} and $m_{HL}/(x_H x_L)$ are strictly decreasing in S (when $x_H x_L > 0$).
6. $\lim_{S \rightarrow 1} m_{HL} = x_H x_L$
7. $\lim_{S \rightarrow \infty} m_{HL} = 0$
8. $\lim_{S \rightarrow 0} m_{HL} = \min\{x_H, x_L\}$

Proof. Appendix ■

Parts 5. and 7. just strengthen the intuition from the last paragraph: if either type prefers not to renew mixed matches, then the share of mixed matches vanishes as matching becomes completely flexible. Similarly, Part 8. shows that if both types prefer to renew mixed matches then their share increases with flexibility until every agent of the scarcer type is in a mixed match. Part 6. shows that matching is random at $S = 1$. Parts 3. and 4. may be thought of as regularity results: all agents play more often against a particular strategy as its population share increases and never play against it as it vanishes.

3.1 Long Run Characterization

Here I focus on the two classes of games for which the effect of flexible matching is most interesting. In *coordination games*, there are Nash equilibria at (H, H) and (L, L) , and each strategy earns its highest payoff in equilibrium: $a > c$ and $d < 0$. In *stag hunt games*, (H, H) and (L, L) are again Nash equilibria, and $a > c$, but L does better when matched with H than against itself: $d > 0$.⁶ Given our earlier assumption (without loss of generality) that $a > 0$, all games with two symmetric pure strategy equilibria have one of these two forms.

The first step is to derive $m^*(H_k, L_l | x^t, S)$. For these types of games, agents playing H always exit from mixed matches, so all mixed matches are the immediate result of random matching in the rematching pool. This means that the histories of the two agents in a mixed match are independent. Then, the probability that the agent playing H in a newly formed mixed match has history k is just equal to the fraction of H 's in the rematching pool with history k . This in turn is equal to the rate at which H 's flow out of Hk matches and into the

⁶In the literature, the term "stag hunt game" is also sometimes used to refer to coordination games in which the efficient equilibrium is not risk dominant: that is, $0 + d > a + c$, but not necessarily $d > 0$. I eagerly welcome suggestions for a more felicitous name that will avoid this confusion.

rematching pool, divided by the total flow of H 's into the rematching pool:

$$\begin{aligned}\Pr(H_k) &= \frac{\sigma_{Hk}m_{Hk}}{\sigma_{HH}m_{HH} + \sigma_{HL}m_{HL}} \\ &= \frac{y_k}{Y}\end{aligned}\tag{10}$$

where the second step uses (7). Similarly, the probability that the agent playing L has history l is equal to the fraction of L 's flowing into the rematching pool that have history l :

$$\begin{aligned}\Pr(L_l) &= \frac{\sigma_{Ll}m_{Ll}}{\sigma_{LL}m_{LL} + \sigma_{HL}m_{HL}} \\ &= \frac{y_l}{Y}\end{aligned}\tag{11}$$

The fact that the distribution of histories is the same for both types is an implication of matching being in steady state. I use $w_H = y_H/Y$ to denote the common probability of history H and $w_L = y_L/Y$ to denote the common probability of history L . Then we can write

$$m^*(H_k, L_l|x^t, S) = w_k w_l m_{HL}(x^t, S)$$

Substituting this into (AD) yields

$$\dot{x}_H = \lambda m_{HL} \sum_{k \in \{H, L\}} \sum_{l \in \{H, L\}} w_k w_l (\pi_{Hk} - \pi_{Ll})\tag{12}$$

Notice the resemblance to (4). The dynamics look like random matching, except that m_{HL} no longer equals $x_H x_L$ and payoffs are weighted with the relative proportions of H and L in the rematching pool, w_H and w_L , , not the true population shares x_H and x_L . In the same fashion as for random matching, (12) can be rewritten in the following way.

$$\dot{x}_H = \lambda m_{HL} (\tilde{\pi}_H - \tilde{\pi}_L)\tag{13}$$

where the ‘‘pseudo-average payoffs’’ $\tilde{\pi}_H$ and $\tilde{\pi}_L$ are given by

$$\begin{aligned}\tilde{\pi}_H &= w_H \pi_{HH} + w_L \pi_{HL} \\ \tilde{\pi}_L &= w_L \pi_{LL} + w_H \pi_{LH}\end{aligned}$$

In contrast to random matching, the population share of H now grows when $w_H > p$, rather than when $x_H > p$. However, as long as w_H is well-behaved – that is, if it is a continuous, strictly increasing function of x_H – then the dynamics will still be governed by a threshold value of x_H , above which trajectories converge to an H -convention and below which trajectories converge to a L -convention. The following lemma ensures that this is the case. Its proof is in

the appendix.

Lemma 2 *For coordination games and stag hunt games under flexible matching, there exists a threshold level $\bar{x}(s)$ of x_H such that trajectories that start above $\bar{x}(s)$ converge to $x_H = 1$ and trajectories that start below $\bar{x}(s)$ converge to $x_H = 0$. This threshold is determined by $\omega_s(\bar{x}(s)) = p$, where $w_H = \omega_s(x_H)$ is a continuous, strictly increasing function.*

Thus the effect of flexible matching on the long run behavior of the population is determined entirely by the bias it creates in the rematching pool. This effect is examined first for coordination games.

Coordination Games. In these games, agents of both types prefer to renew coordinated matches and exit from mixed matches. This has two effects: it increases the measure of agents in coordinated matches which tends to increase the measure of agents that have optimistic perceptions of their strategies, but it also locks those agents up, with their high perceived payoffs, in long term matches where agents playing different strategies are not exposed to them. From this standpoint, the effect on learning seems ambiguous. This is where the correspondence between the histories of agents in a position to learn and the relative shares of each strategy in the rematching pool proves valuable. The bias in the rematching pool has a very clear direction: inflows come disproportionately from mixed matches for which the ratio of H 's to L 's is one to one, so the shares of each strategy in the rematching pool are biased toward $1/2$. As a result, when agents are exposed to new strategies, the probability that they assess them based on their performance against either strategy is closer to $1/2$ than it would be under random matching. This is to the benefit of the risk dominant strategy. My first major result formally states this effect and demonstrates that it is limited.

Proposition 2 *For coordination games,*

1. $|\bar{x}(s) - 1/2|$ is strictly increasing in s . That is, the basin of attraction of the risk dominant equilibrium is increasing in flexibility.

2. $\lim_{s \rightarrow \infty} \bar{x}(s) = \bar{x}_0$, where \bar{x}_0 satisfies $\frac{\bar{x}_0}{1-\bar{x}_0} = \left(\frac{p}{1-p}\right)^2$

Proof. It is not hard to show that $\omega_s(x_H)$ coincides with x_H at 0 , $1/2$, and 1 , so w_H lies on the same side of $1/2$ as x_H , and consequently, $\bar{x}(s)$ lies on the same side of $1/2$ as p . Then, in light of Lemma 2, it suffices for part 1 to show that $\omega_s(x_H)$ is biased toward $1/2$, with the bias increasing in s . When this is the case, we can alternatively think of $x_H = \omega_s^{-1}(w_H)$ as biased away from $1/2$, with the bias increasing in s , so in particular, $\bar{x}(s) = \omega_s^{-1}(p)$ moves away from $1/2$ as s increases.

To show that $\omega_s(x_H)$ is biased toward $1/2$, first note that using (10), we can write $w_H/w_L = \sigma_{HH}m_{HH}/\sigma_{HL}m_{HL}$, or $w_H/w_L = m_{HH}/sm_{HL}$, using the fact that HH matches are renewed

and HL matches are exited. Next, noting that $S = s^2$, we use (8) to write $w_H/w_L = \sqrt{m_{HH}/m_{LL}}$, or, substituting for m_{HH} and m_{LL} ,

$$\frac{w_H}{w_L} = \sqrt{\frac{x_H - m_{HL}}{x_L - m_{HL}}} = \sqrt{1 + \frac{x_H - x_L}{x_L - m_{HL}}}$$

It is clear from this expression that w_H and x_H coincide at 0, 1/2, and 1. When $x_H - x_L$ is positive, because m_{HL} is strictly decreasing in s , w_H/w_L is strictly decreasing in s (but always greater than 1). When $x_H - x_L$ is negative, w_H/w_L is strictly increasing in s (but always less than 1). Thus, $|w_H - 1/2| = |\omega_s(x_H) - 1/2|$ is strictly decreasing in s , completing the first part of the proof.

For the second part, note that $\bar{x}(s)$ can be defined by

$$\sqrt{\frac{\bar{x}(s) - m_{HL}}{1 - \bar{x}(s) - m_{HL}}} = \frac{\omega_s(\bar{x}(s))}{1 - \omega_s(\bar{x}(s))} = \frac{p}{1 - p}$$

Taking limits on both sides, and using the fact that $\lim_{s \rightarrow \infty} m_{HL} = 0$, we have

$$\lim_{s \rightarrow \infty} \sqrt{\frac{\bar{x}(s) - m_{HL}}{1 - \bar{x}(s) - m_{HL}}} = \sqrt{\frac{\lim_{s \rightarrow \infty} \bar{x}(s)}{1 - \lim_{s \rightarrow \infty} \bar{x}(s)}} = \sqrt{\frac{\bar{x}_0}{1 - \bar{x}_0}} = \frac{p}{1 - p}$$

which proves part 2. ■

The most striking implication of this proposition is that when the payoff dominant equilibrium is not risk dominant, giving agents another instrument with which to coordinate (flexible match length) makes efficiency less likely in the long run, not more. Essentially, this is because the benefits of flexibility are distributed unequally over the population: while all agents coordinate better on average (which gives a relative advantage to the efficient strategy), some agents coordinate more poorly (in the short run) than under random matching, and it is these agents who are most likely to change their strategies.

However, part 2 shows that the bias toward the risk dominant strategy is limited. One way to think about this is that when both types renew own type matches, the rematching pool remains relatively balanced. (In fact it is biased toward 50 – 50.) This means that the rematching pool clears quickly, in the sense that an agent exiting a long term match does not go through many bad matches before finding another good one. As a result, agents do not have enough time to grow too discouraged about their strategies.

Stag Hunt Games Quick clearing of the rematching pool depends crucially on the fact that the desired matches of both types of agent are consistent. When agents' preferences over matches conflict, rematching can become very sluggish, and the bias in the dynamics need not be bounded. Consider a stag hunt game like (B). As matching becomes more flexible,

both agents playing H and L tend to spend most of their time playing against their own type. However, because L 's would prefer to match with H 's, they exit own-type matches quickly and flood into the rematching pool. Agents playing H only trickle into rematching slowly, but when they do, they must run a long gauntlet of agents playing L before finding another long term match. This gives them a long time to become discouraged with strategy H , making them more likely to switch to strategy L and also more likely to relay bad news about strategy H to L 's considering switching. This may dramatically reduce the long run prospects of strategy H , even though its average payoff is relatively higher under flexibility. The following proposition demonstrates that as matching becomes completely flexible, the efficient equilibrium is never played in the long run.

Proposition 3 *For stag hunt games,*

1. $\bar{x}(s)$ is strictly increasing in s .
2. $\lim_{s \rightarrow \infty} \bar{x}(s) = 1$. (All interior trajectories converge to $x_H = 0$ and payoffs converge to 0.)

Proof. The key step is to show that the bias in the rematching pool is now unbounded. Now, LL matches are exited, so we have $\sigma_{HH} = \sigma/s$, $\sigma_{HL} = \sigma_{LL} = \sigma$, and $S = s$. This time it is convenient to use (11) to write

$$w_H = \frac{\sigma_{HL}m_{HL}}{\sigma_{HL}m_{HL} + \sigma_{LL}m_{LL}} = \frac{m_{HL}}{m_{HL} + m_{LL}} = \frac{m_{HL}}{x_L}$$

Because m_{HL} is strictly decreasing in s , we have $w_H = \omega_s(x_H)$ strictly decreasing in s as well. Then using $\omega_s(\bar{x}(s)) = p$,

$$\frac{d}{ds}(\omega_s(\bar{x}(s))) = \omega'_s(\bar{x}(s))\frac{d\bar{x}(s)}{ds} + \frac{\partial \omega_s(\bar{x}(s))}{\partial s} = 0$$

The second term is negative, and ω_s is strictly increasing in x_H so $d\bar{x}(s)/ds > 0$, proving part 1.

Next, taking limits, $\lim_{s \rightarrow \infty} \omega_s(\bar{x}(s)) = p$, or using $w_H = m_{HL}/x_L$, $\lim_{s \rightarrow \infty} m_{HL}(\bar{x}(s))/(1 - \bar{x}(s)) = p$. But $m_{HL}(\bar{x}(s)) \leq m_{HL}(1/2)$, so $0 \leq \lim_{s \rightarrow \infty} m_{HL}(\bar{x}(s)) \leq \lim_{s \rightarrow \infty} m_{HL}(1/2) = 0$ (applying Lemma 1), and so the numerator of the left-hand side goes to 0. For the condition to hold, the denominator must go to 0 as well, so $\lim_{s \rightarrow \infty} \bar{x}(s) = 1$, proving part 2. ■

Here, the distribution of the benefits from flexibility is severely unequal. In the context of the earlier example of biologists in a stag hunt, when it is easy to renew matches, a biologist skilled in polymerase chain reactions does very well once she has found a fellow specialist in *PCR* because they can do project after project together. However, when her colleague finishes his postdoc and takes a faculty position across the country (for example) she will face a sea of biologists trained in the standard method who are eager to ride on her training. Sorting

through all of the standard method researchers to find another *PCR* specialist will take a long time precisely because matching is flexible – all of the *PCR* specialists will be locked up together. After a while, she will come to believe, and those who interact with her will learn, that her specialized skill set is wasted among so many agents with whom she is mismatched, and she is likely to let her expertise lapse (i.e., switch to *TT*). Because all *PCR* agents will go through a similar period of disillusionment eventually, their numbers inexorably shrink – even though their average payoffs over the full cycle of good and bad times are higher than under random matching.

It is also interesting to consider the stark difference between the results for coordination games and for stag hunt games. In most of the work on evolutionary games, the critical distinction is between strategies that are risk dominant and those that are not; whether a strategy performs better in or out of equilibrium plays no special role. However, when matching is flexible, agents (roughly speaking) can pick both rows and columns on the game form, so the specific boxes in which a strategy performs well necessarily become important. To further explore this contrast, consider the following two games.

	<i>H</i>	<i>L</i>	
<i>H</i>	1, 1	−1, −1	
<i>L</i>	−1, −1	0, 0	
	D		

	<i>H</i>	<i>L</i>
<i>H</i>	3, 3	−1, 1
<i>L</i>	1, −1	0, 0
	E	

In both games there are two (pure strategy) equilibria, (H, H) , which is both efficient and risk dominant, and (L, L) . Because H is risk dominant, the H -convention has a larger basin of attraction for both games under random matching. However, as matching becomes more flexible, the two games diverge. In D, flexibility causes the basin of H to grow, while in E, it causes the basin of H to shrink. This reflects a substantive difference between agents' incentives in the two games. An interesting parallel is to Aumann's concept of credible cheap talk. In D, an announcement of the intention to play H is credible because an agent who intends to play L would not announce H if she thought she would be believed. In E, however, the announcement would not be credible: an announcement of H that is believed will elicit play of H , so an agent has an incentive to announce H regardless of what she plans to play. The connection can be thought of in an even more direct way: insofar as an announcement is an attempt to induce one's opponent into a particular action, it is a natural analog to the attempt under flexible matching to match with an opponent who plays a particular action.

After exploring the effects of large amounts of flexibility, it is natural to ask how the model behaves when matching is close to random. The following corollary follows directly from the continuity of m_{HL} and ω_s as established earlier.

Corollary 1 *In coordination games and stag hunt games, the cutoff level $\bar{x}(s)$ between basins of attraction is continuous in s .*

In particular, this means that small perturbations of random matching in the direction of flexibility will have small effects on the cutoff level. There is a large literature that attempts to draw conclusions about the stochastic stability of different equilibria by comparing the relative sizes of their basins of attraction; this result confirms that those conclusions will generically be robust when matching is made slightly flexible.

3.2 Slow Convergence

Even when flexible matching does not change the long run behavior of the population, it can have dramatic effects on the speed of learning, and hence, on medium run behavior. When convergence is slow enough, this medium run behavior may be the most relevant description of expected play. In this model, flexibility affects learning by shifting the rate at which agents playing different strategies are exposed to each other. If agents actively seek out mixed matches, then learning will be faster than it would be under random matching. Or if, as in the cases discussed so far, at least one type of agent prefers to avoid mixed matches, then learning will be slower. Then, even if flexibility has advantageous effects in the long run, they may be balanced by welfare costs in the short run. This possibility is clearest in games for which flexibility has no effect on the selection of a limit point, as for example, in the following proposition.

Proposition 4 *Let payoffs satisfy $c/2 + d < a < c + d$ and $a > c$. Then,*

1. *H is a dominant strategy, and $\lim_{t \rightarrow \infty} x_H^t(x^0, s) = 1$ for all interior x^0 and all $s \geq 1$.*

But,

2. *$x_H^t(x^0, s') < x_H^t(x^0, s)$ at every time $t > 0$ if $s' > s$.*
3. *$\bar{\pi}^t(x^0, s') < \bar{\pi}^t(x^0, s)$ at every time $t > 0$ if $s' > s$.*

where $x_H^t(x^0, s)$ and $\bar{\pi}^t(x^0, s)$ are the population share of H and the average payoff of the entire population at time t , for a path originating at x^0 .

Proof. Appendix. ■

For this constellation of payoffs, flexible matching not only slows convergence to the dominant strategy H , it also reduces average payoffs along the entire future path emanating from any initial point. Because the conditions on payoffs are not very illuminating, it is helpful to draw intuition from examples like the following one.

	H	L
H	5, 5	4, 2
L	2, 4	0, 0

Here, all agents eventually learn to play the dominant strategy H . However, agents playing H do not internalize the value of the learning they produce when matched with agents playing L .

Random matching forces them to create a certain amount of learning through forced interaction with agents playing the dominated strategy, but given more choice over their interactions, they will interact with these agents less, slowing the rate of learning. The payoff conditions enter in two ways. First, because only HH matches are renewed, the rematching pool will be heavily biased toward strategy L (just as for the stag hunt game). This means that as matching becomes more flexible, agents who learn about H are more likely to be comparing its performance against L to their own performance against L . This comparison, with a payoff difference of 4, favors H more than the comparison when both agents have played against H recently (in which case the payoff difference is 3). Thus, although flexibility means that L 's are exposed to H less frequently, it also means that H appears relatively more attractive when they are exposed to it – making them more likely to switch when exposed. The condition $a - d > c/2$ ensures that this effect is not too large, so that on net, flexibility slows convergence. The second condition ensures that the “cost” of learning is not too high: incrementally increasing learning by splitting one HH match and one LL match and forming two mixed matches does not lower total payoffs as long as $a < c + d$. Under these conditions, flexibility lowers average welfare. Furthermore, this decline exacerbates payoff inequality: the deterioration in average payoffs is concentrated among the agents playing the dominated strategy, while the average payoff of agents who have learned to play H improves (as they isolate themselves from those who have not).

For coordination games, a clean analysis of welfare is more difficult, but the qualitative effects are similar. Because mixed matches are exited, learning will be slower under flexibility. As matching grows more and more flexible, the population will spend an arbitrarily long time stratified into two groups, one playing the payoff dominant equilibrium and one playing the other equilibrium, with little interaction between the two. Even if the ghettoized second group eventually joins the first in playing the efficient equilibrium, it may take a very long time to do so. Proposition 5 demonstrates that when matching is quite flexible, the length of this period of stratification grows at rate s .

Proposition 5 *For a coordination game, suppose that trajectories beginning at some $x^0 \in (0, 1)$ converge to the efficient convention ($x_H = 1$) for all $s > 1$. For each $\phi \in (x_H^0, 1)$, define $T_\phi(s)$, the first time at which $x_H \geq \phi$. Then, $0 < k_1 < \lim_{s \rightarrow \infty} T_\phi(s)/s < k_2 < \infty$.*

Proof. Appendix. ■

To rephrase, we start the system from a state that is in the basin of attraction of the efficient strategy (for all s). As flexibility (and hence isolation) grows, the waiting time to escape any neighborhood of the miscoordinated initial state diverges at rate s . Then when matching is quite flexible, this picture of stratified populations playing different equilibria is a more reasonable prediction of behavior in the medium run than complete coordination on

either equilibrium.

3.3 Robustness

Because the model I have presented is quite specific, there is a natural interest in how general its conclusions are. I will argue that these effects will be present whenever matching is endogenized in a non-uniform way – and that endogenous matching has a natural tendency to be non-uniform. First, however, I examine how heavily the results lean on the specification of learning.

Of the constraints on learning, there are two that seem particularly restrictive. In limiting agents to learning from their opponents, I rule out learning from global statistics or from successful agents who have not been played against. By requiring agents to act on the basis of information from their most recent opponents, I rule out the possibility of cumulative learning. However, neither of these restrictions drives the results; as long as flexibility induces cyclicity in payoffs and agents are sufficiently impatient, similar results will go through. As an example, consider well-informed agents who form rational expectations of the future returns to each strategy, but are quite myopic. When one of these agents considers switching strategies, she is between matches and correctly anticipates that with probability w_H she will be rematched with an agent playing H . Her expectation of the short run payoff to choosing strategy i is just $w_H\pi_{iH} + w_L\pi_{iL} = \tilde{\pi}_i$. Based on this correct expectation, she will optimally choose to play H if $\tilde{\pi}_H - \tilde{\pi}_L > 0$, so the sign of the dynamics will be exactly the same as it is under *(AD)*. This equivalence arises because in the standard model, agents' recent past payoffs accurately predict (on average) their payoffs in the near future. Moreover, when unhappily matched agents playing different strategies compare payoffs, the payoffs they learn are accurate reflections of how they would fare in the short run in the rematching pool were they to switch strategies. The fact that they tend not to consult agents playing the other strategy who have good experiences but are locked up in long term relationships does not matter too much – even if they did observe those agents' payoffs, they would not view them as achievable in the short run because of the necessity of trolling through the rematching pool for a while before finding a good match. Similarly, recalling past observations of a strategy would not necessarily be helpful in assessing current short run prospects, which depend on the current state of the matching pool.

This line of reasoning suggests that the crucial factors leading endogenous matching to contribute to inefficient outcomes in this case are that flexibility spreads the payoffs to the efficient strategy and that new adopters of the efficient strategy disproportionately occupy the lower tail of that spread. This will tend to be true whenever endogenous matching allows incumbents of the efficient strategy to isolate themselves from the rest of the population if that isolation also makes the strategy harder to learn about, an assumption that seems reasonable in a wide array of contexts. In this model, the final factor that ensures that the efficient

strategy sometimes fails is that incumbents in the isolated group playing the efficient strategy are occasionally ejected and forced to start isolating themselves from the inefficient strategy from scratch, as if they were new adoptees of the strategy. Then, because there are outflows from the efficient strategy but few inflows because of the steep learning curve, the efficient strategy may decline. The generality of this last factor seems more limited; in many situations it is more natural to think that agents who have firmly ensconced themselves within a network of fellow H 's will have minor setbacks but will very rarely be forced to rebuild that network from scratch. In these cases, there still may effectively be barriers to switching to H , so stratification may persist indefinitely.

The next section demonstrates that some degree of myopia is essential to the inefficiency results of this model. When agents are willing to wait a long time for the efficient strategy to pay off, giving them some choice over their partners will make efficient outcomes more likely, not less.

4 The Long Memory Model

This section presents an alternative model in which the perceived payoffs that agents form reflect a longer memory than in the benchmark model. This will be the only difference from the benchmark model, but its effect will be sweeping – now flexibility will make efficient outcomes more likely in a broad class of games. Not only does this mean that the efficient equilibria are more likely to be selected in coordination games and stag hunt games, but when matching is flexible enough, efficient profiles of play will be selected even if they are not Nash equilibria. In particular, a high levels of cooperation will be selected in the prisoner's dilemma. This is possible because with longer time horizons and choice over match lengths, agents can learn to “punish” defectors by not renewing matches with them. This suggests that evolutionary models may be able to shed light on the stability of cooperative equilibria in large populations in a manner analogous to the way that they can make predictions about equilibrium selection in static games.

Now, an agent playing i will form a perceived payoff equal to her current time average payoff in the steady state of the matching process. Thus I move from one extreme (complete myopia) to another (infinitely long time horizons). Looking at these limiting cases allows contrasts in the effects of flexibility to be framed in the clearest possible way; intermediate cases will display a blend of the two sets of results. I continue to interpret the perceived payoff as an agent's assessment of her past performance using a strategy, so the use of time average payoffs indicates that agents recall a statistically large number of past matches, rather than just the most recent

one. Formally, the perceived payoff of an agent playing i is given by

$$\pi_i^t = \frac{m_{iH}^*(x^t)}{x_i^t} \pi_{iH} + \frac{m_{iL}^*(x^t)}{x_i^t} \pi_{iL}$$

Notice that the dependence on history has been dropped: over a long enough horizon, all agents playing strategy i will have the same average experience with it. Furthermore, this time average payoff is the same as the cross-sectional average payoff across agents playing i . This foreshadows the results; when agents' memories are long enough that the cyclical induced by flexible matching does not affect their assessments, then the dynamics will not depend on the distribution of payoffs. Then the fact that flexibility boosts the average payoff to the efficient strategy will come into play.

The rest of the model unfolds almost identically to Section 3.

The choice to exit or renew

As before, an agent chooses to renew a match if it gives her a flow payoff higher than her perceived payoff. That is, if she is playing i , she renews a match with an agent playing j if and only if $\pi_{ij} > \pi_i^t$. Here there are no delicate issues of ties – for interior values of x^t , π_i^t lies strictly between π_{iH} and π_{iL} , so an agent almost surely incorporates both the best and worst matches for strategy i into her perceived payoff. Her exit choices are straightforward and exactly the same as earlier: she exits from her worst rewarded match and tries to renew her best rewarded match, so once again the exit rates will be given by (2).⁷

Adaptive Dynamics

Learning is the same as before: when two agents playing different strategies are matched, the agent with the lower perceived payoff switches strategies with probability proportional to the difference in perceived payoffs. In this case, the formulation is somewhat simpler than (3) because there is no need to keep track of the histories of the matched agents, as all agents playing the same strategy have the same perceived payoff. The dynamics are given by

$$\dot{x}_H^t = \lambda m_{HL}^*(x^t, S)(\pi_H^t - \pi_L^t) \tag{AD'}$$

The intuition developed for the matching process still applies. For example, in coordination games and stag hunt games where at least one type of agent prefers to exit from mixed matches, the steady state measure of mixed matches declines with flexibility. For stag hunt games, this has the effects discussed in the introduction: the average payoffs of agents playing H improve as they isolate themselves from agents playing L , while agents playing L do worse when they

⁷There is actually one subtle issue. An agent's perceived payoff is shaped by the steady state of the matching process, which in turn depends on past exit rates. Thus, the choice of which matches to exit is potentially path dependent. For 2×2 games this is not a problem because exit decisions are constant with respect to m^t , but it must be dealt with more carefully in extensions to more than two strategies.

are forced to play against each other. Although agents playing H still experience long spells of bad matches after a good match ends, those bad spells never last long enough to dominate their assessments of H . As a result, flexibility tends to benefit the efficient strategy.

The situation is a bit more complicated for coordination games, but the result is the same. In this case both strategies benefit from the isolation provided by flexibility. Generally, the efficient strategy benefits more, but under some circumstances it will not. For example, when the population share of H is very close to 1, agents playing H are close to receiving their highest payoff under random matching, so greater flexibility cannot benefit them very much. In this case, the minority of agents playing L have relatively more to gain from isolation. However, this case can arise only when H is so prevalent that it substantially outperforms L , so that the relative advantage to L slows, but does not reverse, convergence to an H -convention. In all other cases, flexibility makes H relatively more attractive, so for coordination games as well, flexibility tends to favor the efficient equilibrium.

The next result establishes that this intuition applies more broadly than just to coordination games and stag hunt games; when agents have long memories, flexible matching weakly improves the efficiency of long run play in *all* 2×2 games. First, though, I introduce some terminology. The greatest lower payoff g of the underlying game is defined by $g = \max_{i,j} < \min < \pi_{ij}, \pi_{ji} >>$. For example, in (A), the greatest lower payoff is 1, while in (C) it is 2. For a vector of population shares x^0 , define the limit point of the trajectory beginning at x^0 to be $x^*(x^0, s)$. Let $\pi^*(x^0, s)$ be the payoff earned at that limit point by all of the strategies in its support. (If all of the surviving strategies did not earn identical payoffs then $x^*(x^0, s)$ would not be a limit point.)

Theorem 1 *For any symmetric 2×2 game G that satisfies (GP), when agents have long time horizons,*

1. *For all x^0 , $\pi^*(x^0, s)$ is increasing in s . (Long run payoffs increase with flexibility.)*
2. *If G has a symmetric Nash equilibrium with payoff lower than g , then for s large enough, there is an open set $X^0(s)$ that is strictly increasing in s with $\pi^*(x^0, s) > \pi^*(x^0, 1)$ for all $x^0 \in X^0(s)$.*

Proof. Appendix. ■

Games that fit the description of the second part of the theorem are: games with two symmetric pure strategy NEs, games with asymmetric NEs, and the prisoner's dilemma. For coordination games and stag hunt games, the theorem means that the basin of attraction of the payoff dominant equilibrium grows with flexibility. For the prisoner's dilemma, the theorem implies that payoffs outside of the static Nash set can be sustained in the long run under flexibility. Ruled out by the second part of the theorem are games with dominant strategy equilibria in which g is earned; flexible matching has no effect on equilibrium selection in these games. The proof proceeds case by case and is relegated to the appendix, but the intuition

is easily sketched. The long run behavior of the system is determined entirely by the zeros of $\Delta\pi(x_H|s) = \pi_H - \pi_L$, which, of course, are rest points of the dynamics; they will be either stable limit points or unstable points that define the boundaries of basins of attraction. I show that the difference in payoffs is increasing in s at these zeros, so as flexibility grows, the subset of $[0, 1]$ on which x_H grows is increasing. Roughly, this means that the efficient strategy is played more often in the long run, which tends to increase long run payoffs.

When (H, H) and (L, L) are not both Nash equilibria of the underlying game, the dynamics are qualitatively different than for coordination games and stag hunt games. Often there will be stable interior rest points that do not correspond to a Nash equilibrium of the underlying game, and these rest points may appear discontinuously when matching becomes flexible enough. The following example helps to illustrate these issues.

Eg. The Prisoner's Dilemma

The underlying game is presented below. For comparison, a normalized version of the stag hunt game from the introduction is presented to its side.

	<i>H</i>	<i>L</i>	
<i>H</i>	2, 2	-1, 3	
<i>L</i>	3, -1	0, 0	
	(PD)		

	<i>H</i>	<i>L</i>	
<i>H</i>	2, 2	-1, 1	
<i>L</i>	1, -1	0, 0	
	(G)		

From the perspective of Nash equilibrium, the two games are quite different. In (PD), strategy L is dominant and (L, L) is the unique equilibrium, while in contrast, (G) has a second pure strategy equilibrium at (H, H) . However, viewed from the standpoint of matching, they are similar: in both games, agents playing H prefer to match with each other, while agents playing L would like to match with agents playing H . That is, $\sigma_{HH} = \sigma/s$ and $\sigma_{HL} = \sigma_{LL} = \sigma$ for both games. Consequently, the effect of more flexible matching on both games should be similar. It was argued earlier for the stag hunt that flexibility allows agents playing the efficient action H to isolate themselves, which benefits the payoff to H more than the payoff to L . For (PD), the same logic means that when cooperators playing H can isolate themselves from defectors playing L , the relative payoff to cooperation rises. When matching is sufficiently flexible, cooperation will be better rewarded than defecting for at least some population shares. To see just how flexible matching must be, I will examine the difference in payoffs, which for the game (PD) is given by $\Delta\pi(x_H|s) = 2 - 3(m_{HL}/x_H + m_{HL}/x_L) = 2 - 3m_{HL}/(x_H x_L)$. Now, $S = s$, and m_{HL} is given by

$$m_{HL} = \frac{1}{2} \frac{\sqrt{1 + 4(s - 1)x_H x_L} - 1}{s - 1}$$

which reflects the diminished impact of flexibility when the two types do not agree on which

matches to exit early. Notice that $m_{HL}/(x_H x_L)$ is symmetric about $x_H = 1/2$ and recall from Lemma 1 that it is strictly decreasing in $x_H x_L$, so $\Delta\pi(x_H|s)$ is symmetric with a single peak at $x_H = 1/2$. We are interested in whether this payoff difference is ever positive; let us look at the most promising case, when $x_H = 1/2$. With a bit of algebra we find

$$\Delta\pi(1/2|s) = 2\left(1 - \frac{3}{\sqrt{s+1}}\right)$$

so with an equally split population, cooperating will outperform defecting whenever $s > 4$. When this condition is met, $\Delta\pi(x_H|s)$ will split the simplex into three regions. When x_H is small, cooperation is not worthwhile; there are too many defectors to screen out. For middle ranges of x_H , cooperation attains a critical mass for which effective isolation is possible; in this region, cooperation will be optimal. When x_H is large, defectors cannot be ghettoized – there simply aren't enough of them. Defectors will face cooperators most of the time, thus earning higher payoffs. This situation is illustrated for $s = 5$ in Figure 1

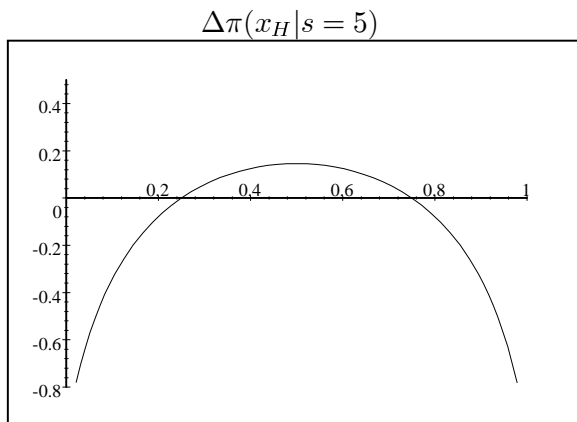


Figure 1

The system has two interior rest points at $x_H = 1/4$ and $x_H = 3/4$ (as well as the rest points at the boundaries, $x_H = 0$ and $x_H = 1$). The first is unstable and the second is stable, so trajectories beginning in $[0, 1/4)$ will converge to the noncooperative static Nash outcome of $x_H = 0$, but trajectories beginning in $(1/4, 1)$ will converge to a profile of play in which $3/4$ of the agents cooperate. Cooperation is sustained at this limit point because $m_{HL}/(x_H x_L) = 2/3$: cooperators and defectors spend $2/3$ as much time interacting as they would under random matching. At this limit point, every agent's payoff is $3/2$ – less than under full cooperation, but more than the static Nash payoff of 0. Notice that the emergence of cooperation is not gradual. When s is near its critical value of 4, a small increase in flexibility makes a large degree of cooperation suddenly possible.⁸

⁸Of course, this is just parallel to the sudden emergence of cooperation in standard repeated games when

Intuition suggests that if boundedly rational agents converge toward a profile of play in which they earn payoffs outside of the set of static Nash equilibria, then they may be behaving in a way that approximates the strategies that patient, rational agents would use to sustain cooperation, and indeed this is the case. Roughly speaking, these agents learn to play in a way that mimics an equilibrium in which most agents cooperate and punish defectors by severing ties with them.⁹ Without friction in matching, full cooperation cannot be sustained because the threat of ending a match is empty: a rogue defector will immediately find another cooperator to victimize. Therefore in any such cooperative equilibrium, there must be enough defectors that the threat of ending a match has teeth. Furthermore, any such equilibrium with rational agents requires those agents to be sufficiently patient, so if boundedly rational agents are to mimic such an equilibrium, they must at least act as though they were patient. Because agents here form their opinions of different strategies on the basis of their performance over a long sequence of matches, their choices have the flavor of those that patient, forward-looking agents would make.

One could take the emergence of cooperation in this example as either a positive or a negative result. To argue the former, it is impressive that the exit option, which makes use of only the most local information and is exercised by boundedly rational agents with only their outside options (and not more elaborate strategic issues) in mind, suffices not only to guarantee the possibility of cooperation, but also convergence to it (from a wide range of initial conditions). To defend the latter perspective, this example demonstrates that matching may need to be quite flexible before cooperation becomes possible. How realistic it is to assume that agents can identify and curtail bad matches in one quarter the time of a normal match, will of course depend on the context. The point, though, is that there is no free lunch: if local information (essentially the costly screening of bad matches) is to substitute for free global information in sustaining cooperation, then that local information must be relatively cheap.

Returning to the general 2×2 case, it is natural to wonder how far the effects of flexible matching will extend. As matching becomes completely flexible, will efficient outcomes always emerge? The following theorem delimits the extent to which this is true.

Theorem 2 *For any symmetric 2×2 game G that satisfies (GP) and with greatest lower payoff g , when agents have long horizons there is some \hat{x} such that $\lim_{s \rightarrow \infty} x^*(x^0, s) = \hat{x}$ and $\lim_{s \rightarrow \infty} \pi^*(x^0, s) = g$ for all interior x^0 .*

Proof. The theorem will be proved here for games with $\sigma_{HL} = \sigma$. The case with $\sigma_{HL} = \sigma/s$ has a similar flavor; for the proof, see the appendix. Suppose that $a > c$ or $d < 0$ so that the discount factor grows large enough; here it is lower hemicontinuity of the correspondence mapping s to limit points that fails.

⁹Ghosh and Ray (1996) present a model with rational, forward-looking agents that has an equilibrium of this sort. Also, Eshel, Samuelson, and Shaked (1998) demonstrate the emergence of cooperation in a evolutionary model with a different brand of imitation dynamics than that used here.

$\sigma_{HL} = \sigma$. Note that in either case, a is the greatest lower payoff. We will show that $\hat{x}_H = 1$: in the limit play converges to strategy H . Fix some interior x^0 and some $\varepsilon > 0$. We must show that for s large enough, $x_H^*(x^0, s) > 1 - \varepsilon$. We do this by demonstrating that for s large enough, $\Delta\pi(x_H, s)$ is strictly positive for $x_H \in [\bar{\varepsilon}, 1 - \bar{\varepsilon}]$ where $\bar{\varepsilon} = \min\{x_H^0, \varepsilon\}$. Then either x_H starts above $1 - \bar{\varepsilon}$ and stays there, or it starts within $[\bar{\varepsilon}, 1 - \bar{\varepsilon}]$ and eventually grows (and remains) greater than $1 - \bar{\varepsilon}$. Toward this end, recall that $\Delta\pi(x_H, s) = a - \frac{m_{HL}}{x_H}(a - c) - \frac{m_{HL}}{x_L}d$ and

$$m_{HL} = \frac{\sqrt{1 + 4(r-1)x_H(1-x_H)} - 1}{2(r-1)}$$

where $r = s^2$ if $\sigma_{HH} = \sigma_{LL} = \sigma/s$ and $r = s$ otherwise. Because m_{HL} has a single peak at $x_H = 1/2$, $m_{HL}(x_H, s) \leq m_{HL}(1/2, s) = \frac{1}{2(1+\sqrt{r})}$. Now choose $\bar{s}(\bar{\varepsilon})$ so that $1 + \sqrt{r} > (|a - c| + |d|)/(2a\bar{\varepsilon})$. Then, if $x_H \in [\bar{\varepsilon}, 1 - \bar{\varepsilon}]$,

$$\begin{aligned} \left| \frac{m_{HL}}{x_H}(a - c) + \frac{m_{HL}}{x_L}d \right| &\leq \frac{1}{2(1 + \sqrt{r})} \left(\frac{|a - c|}{x_H} + \frac{|d|}{x_L} \right) \\ &\leq \frac{1}{2(1 + \sqrt{r})} \left(\frac{|a - c|}{\bar{\varepsilon}} + \frac{|d|}{\bar{\varepsilon}} \right) \\ &< a \end{aligned}$$

so for all $s > \bar{s}(\bar{\varepsilon})$, $\Delta\pi(x_H, s)$ is strictly positive for $x_H \in [\bar{\varepsilon}, 1 - \bar{\varepsilon}]$. Then, for all $s > \bar{s}(\bar{\varepsilon})$, $x_H^*(x^0, s) > 1 - \bar{\varepsilon} \geq 1 - \varepsilon$. Because ε was arbitrary, $\lim_{s \rightarrow \infty} x_H^*(x^0, s) = 1$. It follows directly that $\lim_{s \rightarrow \infty} \pi^*(x^0, s) = a = g$. ■

Corollary 2 *Under the conditions of the theorem, if $g = a$, then $\hat{x}_H = 1$.*

Proof. This was demonstrated above for the case when $\sigma_{HL} = \sigma$. For the proof when $\sigma_{HL} = \sigma/s$, see the appendix. ■

It may be more illuminating to look at the implications of this theorem for various subclasses of 2×2 games. As matching becomes completely flexible, the following are true.

- Coordination games and stag hunt games converge to the payoff dominant equilibrium.
- The prisoner's dilemma converges to full cooperation.
- Hawk-dove games converge to everyone playing "dove."

Hawk-dove games have the form of (HD). The best outcome is to be a hawk (L) facing a dove (H), but if one's opponent is a hawk, it is better to "back down" by being a dove.¹⁰

¹⁰Apologies for the unfortunate labeling – the desire for consistency with the rest of the paper won out.

Thus, (H, L) and (L, H) are Nash equilibria, and under random matching, the population will converge to $2/3$ hawk, $1/3$ dove. This game has something of the flavor of the prisoner’s dilemma; here it is doves who can benefit by isolating themselves from hawks. As in the prisoner’s dilemma, for finite s there is no “all-dove” limit point: hawks will always do well in such a population. However, as the ability to isolate improves, the population can move arbitrarily close to “all-dove.”

	H	L
H	2, 2	1, 4
L	4, 1	0, 0

(HD)

This game also illustrates the limit of flexible matching’s impetus toward efficiency. First of all, it is unclear what efficiency should mean here because there is not a unique Pareto optimal strategy profile. However, if we assume that agents’ payoffs are comparable, then we may be interested in profiles that maximize the sum of payoffs, which in this game occurs along the off-diagonal. Because the game is symmetric, any reasonable equilibrium concept in which players are anonymous (or limit point of an evolutionary game) should have all players earning the same expected payoff (otherwise some agents could do better by switching). The largest payoff consistent with this requirement is 2.5, which is generated from the correlated profile in which (H, L) is played half the time and (L, H) is played the other half of the time. The dynamics here have no hope of generating behavior of this sort because of the assumption that strategy changes happen slowly. It seems likely that in a different model, with quick strategy changes and some common information for agents to condition their behavior on, that play resembling the correlated equilibrium could be reached by some adaptive dynamics. Here it is clear that the connection of flexible matching to efficiency is weak. Flexible matching only implies that outcomes in which at least one agent earns a low payoff are rare. This tends to increase average payoffs when the outcome that maximizes the sum of payoffs treats the two players equally. However, when that outcome favors one player, flexibility encourages the other player to either change strategies or exit the match, both of which reduce the average payoff in the population..

5 Interpreting the Results

In order to constructively assess the assumptions and results of this model, it is helpful to look at them in the context of related work. The model’s primary lineage is the literature on learning and evolutionary game theory. The models in this literature usually describe a world

in which a large number of agents repeatedly interact with each other, and the outcomes of these interactions generate some dynamic process on the strategies used by the agents. Within this rubric, there is a great deal of diversity in how the dynamics are modeled. For example, at one extreme, in the paradigm introduced from evolutionary biology by Maynard Smith, agents are taken to be genetically programmed to play different strategies, and the dynamics select for better performing agents. At the other extreme are models of rational learning like Fudenberg and Levine (1993), in which agents have complete flexibility in choosing strategies and the dynamics describe agents' best responses to evolving beliefs about how other agents will play. My conception of strategy change lies between these two benchmarks; my agents change strategies, but there are implicit adjustment costs, so change is slow. The way that agents interact with each other has also been modeled in a variety of ways, but with a few exceptions, the interaction structure has been taken to be exogenous. Because these models tend to predict that the equilibrium converged to (when there are more than one) will depend on initial conditions when the dynamics are deterministic, a body of work following Kandori, Mailath, and Rob (1993) and Young (1992) focuses on identifying which equilibria are stable when noise is added to the dynamics. In a broad selection of models, the risk dominant equilibrium can be shown to be uniquely stable in coordination games with vanishing noise, even if it is not efficient.¹¹ Roughly, this is because noise periodically knocks the system out of equilibrium, and the risk dominant strategy performs well more often in these disequilibrium situations than the other strategy does.

However, this result hinges on the exogeneity of the interaction structure (or on the paucity of information, depending on how one wishes to think about it). An agent is assumed to be incapable of identifying and selectively interacting with subgroups of agents who are more likely to be playing strategies that she considers favorable. This may be either because she cannot control her interactions or because she has no reliable information with which to distinguish potential opponents. More recently, these assumptions have been challenged by Kim and Sobel (1995), Ely (1996), and others. Kim and Sobel have a model in which strategy change is quick and note that when the game has common interests and agents can communicate sufficiently well, announcements of one's intentions to play the efficient strategy can be credible, so that using and responding to these announcements can provide an adaptive advantage. When strategy change is slow (so that one's strategy cannot be tailored to each individual opponent), this sort of signaling is only helpful if agents can isolate their interactions on the basis of these signals. Ely and Oeschler (1997) both give agents an instrument with which they can influence who they interact with (in both cases this is thought of as a choice of location, although the interpretation need not be geographic) and show that agents can use this instrument to coordinate on the efficient strategy.

¹¹For contrarian positions, see Bergin and Lipman (1996), and Robson and Vega-Redondo (1996)..

However, in these models there is a tension between the ability of agents to isolate themselves, which must be relatively good if an efficient but fragile strategy is to be nurtured, and the information flow from that isolated group to the larger population, which must also be good if that strategy is to grow. Both Ely and Oeschler assume that information about the isolated pocket of efficiency is good, not necessarily in the sense that it arrives quickly (which it may not), but more in the sense that it is complete – when an agent learns about this pocket of efficiency, she in one fell swoop learns everything necessary to change strategies and isolate herself in the same way. My model follows the spirit of these two in endogenizing interaction, but differs in two principal ways. First, the instrument that my agents use to influence interaction, extending match length, is not costless because agents face the implicit cost of screening out bad matches before reaping the benefits of isolation. This implicit cost is endogenous to the model and as we have seen, it is sometimes quite high. Second, in my model, information is hard to come by, and learning a new strategy can be difficult. This is captured in the fact that the rate of strategy change depends on agents’ direct exposure to different strategies. As the last section demonstrated, the latter factor alone will not prevent efficient long run outcomes when agents assess the costs and benefits presented by flexibility from a long term perspective.¹² However, when agents are shortsighted, their perceptions of an efficient but fragile strategy under flexibility may be colored by the fact that the costs using the efficient strategy are frontloaded (in the form of screening bad matches) and the benefits are backloaded. In this case, even though the underlying game may have common interests, the larger game incorporating agents’ decisions about who to match with does not: for example, when two agents playing H renew a match, their payoffs improve, but those of other agents playing H drop, as their chances of rematching successfully have just fallen.

While in this model it is the horizon over which agents assess their payoffs that determines how flexibility will affect long run behavior, one can easily imagine similar models in which the assessment of payoffs could be given different interpretations. To take one simple example, suppose that agents are shortsighted but simultaneously engage in several relationships rather than just one. They must use the same strategy in each relationship, and they have the same ability to exit or renew each of these matches as in the standard model. If these multiple relationships are independent of each other, they will tend to smooth an agent’s payoff toward its time average. For example, it will be rare for an agent playing the efficient strategy to be completely discouraged with it because that would require all of her matches to have ended more or less simultaneously. Then the dynamics will collapse to the standard model when the number of concurrent matches is one, and they will approach the long horizons model when the number of matches is large. Consequently, flexibility will enhance efficiency in populations that

¹²Although this may be just an artifact of the restriction to 2×2 games, where, because there is only one state variable, only the sign of the dynamics (and not the speed) matters for long run convergence. With more strategies, the relative rates of growth of various strategies will typically matter as well.

are densely connected, but possibly hurt it in populations that are more sparsely connected.¹³

In contrast with a preponderance of the work in this area, I concentrate on deterministic dynamics.¹⁴ I do so principally because I am interested in describing long run behavior rather than what Binmore, Samuelson, and Vaughan (1995) have termed the “ultra-long run,” i.e., the time frame over which a system with very small perturbations realizes its ergodic distribution, which as Ellison (1993) has demonstrated, may be longer than any reasonable period of observation. Of course, a stability analysis akin to stochastic stability could be applied more or less directly to this model, but it would not add very much; all of the new intuition lies in the deterministic dynamics.¹⁵

In designing my flexible matching process, I have followed the ample literatures on search and matching in some ways and diverged in others. I will highlight the differences because they are of the most interest. First of all, all of the friction in matching is supplied entirely by the possibility of entering bad matches; my agents are never “unemployed.” This is not necessary, but it is desirable if the model is to be thought of as game theory. By introducing unemployment, and a value associated with it, I would effectively normalize payoffs by something external to the game form and potentially make the source of the results more difficult to identify. In a similar spirit, my agents exercise control over matching *ex post* rather than *ex ante*: they decide how quickly to leave matches, but they cannot identify and refuse bad matches before entering them. This is more consistent with a world in which exigencies force pairs of agents to work together for a while than with one in which those exigencies only determine which agents meet. The feature that agents decide which matches to renew with an eye toward a (possibly adversely) selected rematching pool is developed much more elaborately in Montgomery (1999) and Burdett and Coles (1998), who both demonstrate the possibility that cycles may be generated. In order to manage both strategy change and match endogeneity, I have restricted agents to a particularly simple set of exit options, but the qualitative conclusions would not change dramatically with a more complex matching framework. The current framework extends directly to more than two strategies, but the ranking of the outside option relative to flow payoffs will no longer be degenerate. With four or more strategies, this produces new coordination problems: there may be multiple sets of consistent exit decisions, because the matches that one type considers to be better than average depend on its perceived

¹³Of course, a truly interesting model would endogenize this number of concurrent matches. One can foresee that the number of matches would probably be suboptimal for standard reasons (agents do not internalize the contagion effects when their bad experiences with the efficient strategy lead them to switch, so they do not “insure” enough against those bad experiences), but there are all sorts of other issues to explore as well.

¹⁴That is, deterministic at the aggregate level. Of course, individual agents would not consider anything about this model deterministic.

¹⁵More or less, because there would be a few, surmountable technical obstacles to overcome. In order to apply stochastic stability directly, one would have to reformulate the model with a finite number of agents (but all of the intuition from the continuum case would carry through). Alternatively, one could inject noise into the continuum population directly, in the style of Fudenberg and Harris (1992).

payoff, which in turn depends on the exit decisions made by other agents.

Finally, my specification of learning is motivated in part by recent work, both theoretical and empirical, that emphasizes how agents' choices are influenced by the choices of those they interact with. This influence may operate by increasing the salience of certain strategies, by introducing information about the rewards to those strategies (as I have assumed), by providing an accessible example from which to literally learn how to play a complex strategy, or through a number of other channels (some economic, and some not).¹⁶ Durlauf (2000) provides a good survey, particularly of the sociological evidence, while Bala and Goyal (1998) and Ellison and Fudenberg (1993 and 1995) give theoretical characterizations of this sort of local social learning. When local learning is important, any self-sorting that agents do will affect what they learn, and through this, the actions they choose. The possibility that this sort of feedback could lead to the entrenchment of some agents in a suboptimal strategy is explored by Piketty (1995) and Fernandez and Rogerson (2000), among others. My model could be interpreted roughly as a model of local learning in which learning, because it occurs through play of the game, imposes an externality on the agent one learns from. If agents playing L drag down the payoffs of agents playing H when they learn from them, then allowing agents more flexibility to sort themselves helps to insulate those already playing the efficient strategy, robbing the others of the opportunity to learn.

6 Conclusion

Allowing agents to bring their experience in a larger population to bear on how they play a game against a specific opponent has proved to be a valuable and persuasive way to explain how equilibrium beliefs may form. This model stems from my conviction that there is an equally interesting story in how agents use this broader experience to decide *who* to play. This choice is endogenized in an extremely simple way – agents use their perceived outside options to decide whether to continue playing against a current opponent. This simple formulation interacts with the process of belief formation to generate results that are different from the standard model, sometimes subtly so. Endogenous matching can improve returns to strategies that are efficient in equilibrium, but fragile out of it, by permitting agents using those strategies to isolate themselves from those who are not. However, when this isolation also makes the success of efficient strategies harder to learn about, the net effect of endogenous matching may be to make the long run survival of efficient strategies less likely. When agents have more patience with the cyclical nature of their payoffs, efficiency is more likely to survive, but it may be at the expense of a persistent “ghetto” of agents playing inefficiently who are discouraged from

¹⁶This last possibility seems underexplored; Sethi (1998) (which incorporates the difficulty of learning a strategy in an otherwise standard evolutionary model) is to my knowledge the only paper to address it.

adopting the efficient strategy by the very isolation that makes it profitable. By illustrating a situation in which learning remains endogenously incomplete for a long time, this model suggests the directions in which a full theory of endogenously incomplete social learning might go. A rapidly increasing body of empirical evidence attests that such a theory is sorely needed; its development will have to await future work.

7 Proofs

Proof of Proposition 1

We have

$$\dot{m}_{HL} = \frac{R(m_{HL})}{D} = \frac{(1-S)m_{HL}^2 - m_{HL} + x_H x_L}{D} \quad (MP)$$

When $x_H x_L = 0$, we have $m_{HL} = 0$ and $\dot{m}_{HL} = 0$. Consider the case where $x_H x_L > 0$. Then $R(0) = x_H x_L > 0$. Furthermore,

$$\begin{aligned} R(\min\{x_H, x_L\}) &= \min\{x_H, x_L\}(\min\{x_H, x_L\} - 1) - S \min\{x_H, x_L\}^2 + x_H x_L \\ &= -\min\{x_H, x_L\} \max\{x_H, x_L\} + x_H x_L - S \min\{x_H, x_L\}^2 \\ &= -S \min\{x_H, x_L\}^2 \\ &< 0 \end{aligned}$$

Then because $R(m_{HL})$ is quadratic in m_{HL} , there exists $m_{HL}^* \in (0, \min\{x_H, x_L\})$ such that \dot{m}_{HL} is positive on $[0, m_{HL}^*)$, negative on $(m_{HL}^*, \min\{x_H, x_L\})$, and zero at m_{HL}^* ; that is, there exists a globally stable steady state at m_{HL}^* . Because $R'(m_{HL}^*) < 0$, m_{HL}^* is the smaller root of $R(m_{HL})$ when $(1-S)$ is positive, and the larger root when $(1-S)$ is negative. In both cases the correct root is

$$\frac{\sqrt{1 + 4(S-1)x_H x_L} - 1}{2(S-1)}$$

Proof of Lemma 1

1. This was shown in the last proof.
2. This is obvious if $S > 1$ and follows upon only a little bit more reflection if $S < 1$. If $S = 1$, then $m_{HL} = x_H x_L$, which is clearly increasing in itself.
3. Differentiate m_{HL}/x_H . (There is a cleaner way when $S > 1$, but not, as far as I can tell, when $S < 1$.)

$$\frac{d}{dx_H} \left(\frac{m_{HL}}{x_H} \right) = \frac{\sqrt{Z^2 - 4S(S-1)x_H^2} - Z}{2(S-1)x_H^2 \sqrt{Z^2 - 4S(S-1)x_H^2}}$$

where $Z = 2(S-1)x_H + 1$. When $S > 1$, $Z > 0$, the denominator is positive, and the second term under the square root is positive, so the numerator is negative. When $S < 1$, the denominator is negative, and the second term under the square root is positive, so the

numerator is greater than $|Z| - Z \geq 0$. In each case, the expression is negative (strictly negative when $x_H > 0$). At $S = 1$, $m_{HL}/x_H = x_L$ which is decreasing in x_H . The derivation for m_{HL}/x_L is identical.

4. For x_i small, $\sqrt{1 + 4(S-1)x_i(1-x_i)} = 1 + 2(S-1)x_i + o(x_i^2)$. (Take a Taylor expansion.) Then $m_{HL}/x_i = 1 + o(x_i)$, which goes to 1 as x_i goes to 0.

5. Again, differentiate. Letting $k = 4x_Hx_L$, we have

$$\begin{aligned} \frac{dm_{HL}}{dS} &= \frac{\sqrt{4 + 4k(S-1)} - (2 + k(S-1))}{4(S-1)^2\sqrt{1 + k(S-1)}} \\ &< \frac{\sqrt{4 + 4k(S-1) + k^2(S-1)^2} - (2 + k(S-1))}{4(S-1)^2\sqrt{1 + k(S-1)}} \\ &< \frac{(2 + k(S-1)) - (2 + k(S-1))}{4(S-1)^2\sqrt{1 + k(S-1)}} = 0 \end{aligned}$$

The first inequality holds whenever x_H and x_L are positive, and the next line follows because $S > 0$ and $x_Hx_L \leq 1/4$ imply that $2 + k(S-1)$ is positive.

6. Take a Taylor expansion in $S-1$: $\sqrt{1 + 4(S-1)x_Hx_L} = 1 + 2(S-1)x_Hx_L + o((S-1)^2)$. Then $m_{HL}/(x_Hx_L) = 1 + o(S-1)$ which goes to 1 as S goes to 1.

7. Obvious.

8. At $S = 0$,

$$\begin{aligned} m_{HL} &= (1 - \sqrt{1 - 4x_Hx_L})/2 \\ &= (1 - \sqrt{(x_H + x_L)^2 - 4x_Hx_L})/2 \\ &= (x_H + x_L - \sqrt{(x_H - x_L)^2})/2 \\ &= ((x_H + x_L) - |x_H - x_L|)/2 \\ &= (\min\{x_H, x_L\} + \max\{x_H, x_L\} - (\max\{x_H, x_L\} - \min\{x_H, x_L\}))/2 \\ &= \min\{x_H, x_L\} \end{aligned}$$

Proof of Lemma 2

First we demonstrate that w_H is strictly increasing in x_H . For stag hunt games, we can write $w_H = m_{HL}/x_L$, which is strictly decreasing in x_L (and hence increasing in x_H) by part 3

of Lemma 1. For coordination games, use (10) to write

$$\begin{aligned}
w_H/w_L &= \sigma_{HH}m_{HH}/(\sigma_{HL}m_{HL}) \\
&= m_{HH}/(sm_{HL}) \\
&= (x_H - m_{HL})/(sm_{HL}) \\
&= \frac{1}{s} \frac{x_H}{m_{HL}} - \frac{1}{s}
\end{aligned}$$

Again, using part 3 of Lemma 1, note that m_{HL}/x_H is strictly decreasing in x_H , so w_H/w_L , and hence w_H , are strictly increasing in x_H . Continuity follows directly from the continuity of m_{HL} . Finally, note that because $m_{HL} = 0$ if $x_H = 0$ or $x_L = 0$, we have $\omega_s(0) = 0$ and $\omega_s(1) = 1$ for both coordination games and stag hunt games, so ω_s maps onto $[0, 1]$. Then for any $p \in [0, 1]$, there is a unique $\bar{x}(s)$ such that $\omega_s(\bar{x}(s)) = p$. Then $\omega_s(x_H) \geq p$ for x_H greater or less than $\bar{x}(s)$ respectively, so by (13), \dot{x}_H grows whenever $x_H > \bar{x}(s)$ and shrinks when $x_H < \bar{x}(s)$.

Proof of Proposition 4

Because $a > c$ and $a < c + d$, we have $d > 0$, so $S = s$. Furthermore, $a > c/2 + d > d$, so H is a dominant strategy. Using (13), the dynamics are given by

$$\dot{x}_H = \lambda m_{HL}(w_H(a - d) + w_L c)$$

which is everywhere positive, which suffices for part 1. From the analysis of the stag hunt, we know that when $S = s$, $w_H = m_{HL}/x_L$, so we can write

$$\dot{x}_H/\lambda = \frac{m_{HL}^2}{x_L}(a - c - d) + m_{HL}c$$

We are interested in how the right-hand side changes with s . Differentiating this expression, we have

$$\frac{1}{\lambda} \frac{d}{ds}(\dot{x}_H) = \frac{\partial}{\partial m_{HL}} \left(\frac{m_{HL}^2}{x_L}(a - c - d) + m_{HL}c \right) \frac{dm_{HL}}{ds}$$

In this case, m_{HL} is decreasing in s , so

$$\text{sgn}\left(\frac{d}{ds}(\dot{x}_H)\right) = -\text{sgn}\left(\frac{2m_{HL}}{x_L}(a - c - d) + c\right)$$

But $m_{HL}/x_L < 1$, so $\frac{2m_{HL}}{x_L}(a - c - d) > 2(a - c - d)$ (because $a < c + d$). Then, $\frac{2m_{HL}}{x_L}(a - c - d) + c > 2(a - c/2 - d)$ which is positive by assumption. Consequently, \dot{x}_H is decreasing in s . This demonstrates part 2. Finally, the population average payoff at time t is just

$m_{HH}^t a + m_{HL}^t(c + d) + m_{LL}^t 0$ (payoffs weighted by the measure of agents earning each one). We can write this as

$$\bar{\pi}^t(x^0, s) = x_H^t(x^0, s)a + m_{HL}(x_H^t(x^0, s), s)(c + d - a)$$

Note that because $c + d - a > 0$, average payoffs are increasing in m_{HL} , holding x_H fixed. They are also increasing in x_H , holding s fixed. To see this, differentiate $x_H a + m_{HL}(x_H, s)(c + d - a)$ with respect to x_H to get

$$a + \frac{\partial m_{HL}}{\partial x_H}(c + d - a)$$

But $\partial m_{HL}/\partial x_H$ is given by

$$\frac{\partial m_{HL}}{\partial x_H} = \frac{1 - 2x_H}{\sqrt{1 + 4(s - 1)x_H x_L}}$$

which is minimized at $\partial m_{HL}/\partial x_H|_{x_H=1} = -1$. Then if $x_H < 1$, the derivative of average payoffs with respect to x_H is greater than $a - (c + d - a) > 2a - c - 2d$, which is positive by assumption, so $x_H a + m_{HL}(x_H, s)(c + d - a)$ is increasing in x_H . Then if $s' > s$, we have

$$\begin{aligned} \bar{\pi}^t(x^0, s') &= x_H^t(x^0, s')a + m_{HL}(x_H^t(x^0, s'), s')(c + d - a) \\ &< x_H^t(x^0, s')a + m_{HL}(x_H^t(x^0, s'), s)(c + d - a) \\ &< x_H^t(x^0, s)a + m_{HL}(x_H^t(x^0, s), s)(c + d - a) \\ &= \bar{\pi}^t(x^0, s) \end{aligned}$$

where the second line follows because m_{HL} is decreasing in s and average payoffs are increasing in m_{HL} , and the third line uses part 2 of the proposition and the fact that average payoffs are increasing in x_H . This completes the proof.

Proof of Proposition 5

The idea will be to put upper and lower bounds on the rate of growth of x_H and show that these bounds shrink at rate $1/s$ as mixed matches grow rarer. First we find an upper bound, which involves two steps: bounding $\tilde{\pi}_H - \tilde{\pi}_L$ and bounding $m_{HL}/x_H x_L$. To bound the pseudo-payoffs, note that for a coordination game, $\tilde{\pi}_H < a$ and $\tilde{\pi}_L > d$ for all interior x_H , so $\tilde{\pi}_H - \tilde{\pi}_L < L_1 = a - d$. Then write $\varepsilon = 1 - \phi$, $\bar{\varepsilon} = \min\{\varepsilon, x_H^0\}$, and $\delta = \bar{\varepsilon}(1 - \bar{\varepsilon})$. Because $m_{HL}/x_H x_L$ symmetric about $x_H = 1/2$ and decreasing in $(x_H x_L)$, we have $m_{HL}/x_H x_L < m_{HL}(\delta)/\delta$ for all $x_H \in (\bar{\varepsilon}, 1 - \bar{\varepsilon})$ (and so, *a fortiori*, for all $x_H \in (x_H^0, \phi)$). Noting that $S = s^2$,

we can write $m_{HL}(\delta)/\delta$ as follows.

$$\begin{aligned}
\frac{m_{HL}(x_H x_L = \delta)}{\delta} &= \frac{\sqrt{1 + 4\delta(s^2 - 1)} - 1}{2\delta(s^2 - 1)} \\
&= \frac{2s\sqrt{\delta}\sqrt{1 + \frac{1-4\delta}{4\delta}\frac{1}{s^2}} - 1}{2(s^2 - 1)} \\
&= \frac{s\sqrt{\delta}}{s^2 - 1}(1 + o(1/s)) - \frac{1}{2(s^2 - 1)} \\
&= \sqrt{\delta}\frac{1}{s} + o(1/s)
\end{aligned}$$

Then for $x_H \in (x_H^0, \phi)$ we have

$$\begin{aligned}
\dot{x}_H &= \lambda(m_{HL}/x_H x_L)x_H x_L(\tilde{\pi}_H - \tilde{\pi}_L) \\
&< \lambda m_{HL}/x_H x_L x_H x_L L_1 \\
&< \lambda L_1 \frac{m_{HL}(\delta)}{\delta} x_H x_L \\
&< \frac{1}{s} \frac{\lambda L_1}{\sqrt{\delta}} x_H x_L + o(1/s)
\end{aligned}$$

Dividing both sides by $x_H x_L$ and integrating between $t = 0$ and $t = T_\phi(s)$,

$$\begin{aligned}
\int_{x_H^0}^{\phi} \left(\frac{1}{x_H} + \frac{1}{1 - x_H}\right) dx_H &< \int_0^{T_\phi(s)} \left(\frac{1}{s} \frac{\lambda L_1}{\sqrt{\delta}} + o(1/s)\right) dt \\
\ln M &< \frac{T_\phi(s)}{s} \frac{\lambda L_1}{\sqrt{\delta}} + T_\phi(s) o(1/s) \quad \text{or} \\
\frac{T_\phi(s)}{s} + o(T_\phi(s)/s) &> \frac{\sqrt{\delta} \ln M}{\lambda L_1} = k_1
\end{aligned}$$

where

$$M = \frac{\phi}{1 - \phi} / \frac{x_H^0}{1 - x_H^0}$$

But then for s large enough, we have

$$\frac{T_\phi(s)}{s} > k_1$$

as desired. Now we put a lower bound on the rate of growth of x_H (and hence an upper bound on $T_\phi(s)/s$). Recall that w_H is increasing in x_H and $\tilde{\pi}_H - \tilde{\pi}_L = c + w_H(a - c - d)$ is increasing in w_H , so $\tilde{\pi}_H - \tilde{\pi}_L > c + \omega_s(x_H^0)(a - c - d)$ everywhere along a path beginning at x_H^0 . Also recall that for coordination games, $\omega_s(x_H^0)$ lies between $\omega_1(x_H^0) = x_H^0$ and $\omega_\infty(x_H^0)$ (which lies

between x_H^0 and $1/2$). Let $w_{\min} = \min\{x_H^0, \omega_{\infty}(x_H^0)\}$. Then, regardless of s , along any path starting at x_H^0 we have $\tilde{\pi}_H - \tilde{\pi}_L > L_2 = c + w_{\min}(a - c - d)$. Next we bound $m_{HL}/x_H x_L$. Because $m_{HL}/x_H x_L$ has its minimum over x_H at $x_H = 1/2$, we have

$$\begin{aligned} \frac{m_{HL}}{x_H x_L} &> 4m_{HL}(x_H x_L = 1/4) \\ &= \frac{2\sqrt{1 + 4(s^2 - 1)/4} - 1}{s^2 - 1} \\ &= \frac{2}{s + 1} \end{aligned}$$

Then,

$$\dot{x}_H > \frac{2\lambda L_2}{s + 1} x_H x_L$$

Integrating from $t = 0$ to $t = T_{\phi}(s)$ as before, we have

$$\begin{aligned} \ln M &> 2\lambda L_2 \frac{T_{\phi}(s)}{s + 1} \\ \frac{T_{\phi}(s)}{s} &< \frac{s + 1}{s} \frac{\ln M}{2\lambda L_2} \end{aligned}$$

Let $k_2 = (1 + \kappa) \frac{\ln M}{2\lambda L_2}$ for any $\kappa > 0$. Then for s large enough, we have

$$\frac{T_{\phi}(s)}{s} < k_2$$

so for s large enough

$$0 < k_1 < \frac{T_{\phi}(s)}{s} < k_2 < \infty$$

which establishes the result.

Proof of Theorem 1

Using the definition of m_{HL} we can look for solutions to $\Delta\pi(x_H|s) = 0$ in terms of a , c , d , and s . With a bit of algebra, the zeros can be expressed as solutions to a quadratic equation, so they can be written in closed form. Because these closed form solutions are not very illuminating, I resort to them only when necessary. For now, we will just make use of the fact that $\Delta\pi(x_H|s)$ crosses zero at most two times on the interval $[0, 1]$. By examining $\Delta\pi(x_H|s)$ at its endpoints, we can restrict its behavior further. Using the fact that the share of time spent playing the other type, m_{HL}/x_i , goes to 1 as x_i goes to 0 and goes to 0 as x_i goes to 1, we have $\Delta\pi(0|s) = c$ and $\Delta\pi(1|s) = a - d$. There are several cases to consider; first I look

at the four possibilities when $\sigma_{HL} = \sigma$ (that is, $a > c$ or $d < 0$) so that increasing flexibility reduces off-diagonal play.

Case 1: $a > d$ and $c < 0$

Then $\Delta\pi(x_H|s)$ crosses zero at least once, and generically only once. Later I will show that it is safe to ignore the possibility of a second zero. Then, for some $\bar{x}(s)$, $\Delta\pi(x_H|s) < 0$ if $x_H < \bar{x}(s)$ and $\Delta\pi(x_H|s) > 0$ if $x_H > \bar{x}(s)$. Note that at $\bar{x}(s)$, $\Delta\pi(x_H|s) = 0$ implies that

$$a = m_{HL}(\bar{x}(s), s) \left(\frac{a-c}{\bar{x}(s)} + \frac{d}{1-\bar{x}(s)} \right)$$

Pick $s' > s$. Then m_{HL} decreasing in s and $a > 0$ mean that

$$a > m_{HL}(\bar{x}(s), s') \left(\frac{a-c}{\bar{x}(s)} + \frac{d}{1-\bar{x}(s)} \right)$$

But then, $\Delta\pi(\bar{x}(s)|s') > 0$, so we must have $\bar{x}(s') < \bar{x}(s)$. Because x_H is continuous and positive exactly when $\Delta\pi(x_H|s) > 0$, $x_H^*(x^0, s) = 1$ if $x^0 > \bar{x}(s)$ and $x_H(x^0, s) = 0$ if $x^0 < \bar{x}(s)$. Therefore,

$$\begin{aligned} \pi^*(x^0, s) &= a && \text{if } x_H^0 > \bar{x}(s) \\ &= 0 && \text{if } x_H^0 < \bar{x}(s) \end{aligned}$$

Because $\bar{x}(s)$ is decreasing in s , $\pi^*(x^0, s)$ is increasing in s . Let $X^0(s) = \{x^0 | x_H^0 \in (\bar{x}(s), \bar{x}(1))\}$. This set is growing in s , and all initial points in $X^0(s)$ converge to strictly higher payoffs under s than they would have under random matching: $\pi^*(x^0, s) = a > \pi^*(x^0, 1) = 0$ for all $x^0 \in X^0(s)$.

Now suppose that there is some s such that $\Delta\pi(x_H|s)$ is tangent to the x_H -axis at x'_H , with $x'_H > \bar{x}(s)$. Because $\Delta\pi(x_H|1)$ is linear, we can safely assume that $s > 0$. Because $x'_H > \bar{x}(s)$, $\Delta\pi(x_H|s)$ is positive in a neighborhood of x'_H . But then we can take $s' = s - \varepsilon$, and by the same logic as above, $\Delta\pi(x'_H|s') < 0$. For ε small enough, there will still be points between $\bar{x}(s')$ and x'_H for which $\Delta\pi(x_H|s')$ is positive. But then $\Delta\pi(x_H|s')$ changes sign three times and must have three zeros, a contradiction. Similarly, there can be no second zero to the left of $\bar{x}(s)$, so the assumption of one zero is valid.

Case 2: $a > d$ and $c > 0$

$\Delta\pi(0|s)$ and $\Delta\pi(1|s)$ are both positive, so $\Delta\pi(x_H|s)$ generically crosses zero either never or twice. Under random matching, because H is a dominant strategy, the payoff difference is always positive. Toward a contradiction, let s' be the least s such that the payoff difference

has a zero, and let $\Delta\pi(x'_H|s') = 0$. Then take $s'' < s'$. We have

$$\begin{aligned} a &= m_{HL}(x'_H, s') \left(\frac{a-c}{x'_H} + \frac{d}{1-x'_H} \right) \\ &< m_{HL}(x'_H, s'') \left(\frac{a-c}{x'_H} + \frac{d}{1-x'_H} \right) \end{aligned}$$

so $\Delta\pi(x'_H, s'') < 0$. But then, $\Delta\pi(x_H, s'')$ changes sign twice and must have two zeros, contradicting the minimality of s . Thus $\Delta\pi(x_H|s) > 0$ for all x_H and s , so all trajectories converge to $x_H = 1$. In this case, because the system always converges to the greatest lower payoff under random matching, flexible matching has no additional effect.

Case 3: $a < d$ and $c > 0$

Here $\Delta\pi(0|s) > 0 > \Delta\pi(1|s)$ so $\Delta\pi(x_H|s)$ crosses zero exactly once. (Additional nongeneric zeros can be ruled out in the same way as in Case 1.) Call the zero $\bar{x}(s)$ and observe that $\Delta\pi(x_H|s)$ is positive when $x_H < \bar{x}(s)$ and negative when $x_H > \bar{x}(s)$, so all interior paths converge to $x_H = \bar{x}(s)$. In exactly the same way as for Case 1, we can show that $s' > s$ implies that $\Delta\pi(\bar{x}(s)|s') > 0$, so $\bar{x}(s') > \bar{x}(s)$ and this time $\bar{x}(s)$ is strictly increasing in s . The long run payoff is $\pi^*(x^0, s) = \pi_H(\bar{x}(s), s) = a - (a-c)m_{HL}/\bar{x}(s)$. By assumption $s_{12} = s$ and $d > a > 0$, so we must have $a-c > 0$. Therefore, $\pi^*(x^0, s)$ will be increasing in s if $m_{HL}/\bar{x}(s)$ is decreasing in s . Using the fact that m_{HL}/x_H is decreasing in x_H holding s fixed and the fact that m_{HL} is decreasing in s , for $s' > s$ we have

$$\frac{m_{HL}(\bar{x}(s'), s')}{\bar{x}(s')} < \frac{m_{HL}(\bar{x}(s'), s)}{\bar{x}(s')} < \frac{m_{HL}(\bar{x}(s), s)}{\bar{x}(s)}$$

so $m_{HL}/\bar{x}(s)$ is in fact decreasing in s , and the long run payoff increases with flexibility. In this case, all increases in s raise the long run payoff for all interior trajectories: we can take $X^0(s) = (0, 1)$.

Case 4: $a < d$ and $c < 0$

Now $\Delta\pi(0|s)$ and $\Delta\pi(1|s)$ are both negative, so there are generically zero or two points where the payoff difference is zero. In particular, under random matching, $\Delta\pi(x_H|s)$ is everywhere negative. When there are two zeros, call them $x^l(s)$ and $x^h(s)$, with $x^l(s) < x^h(s)$. Then the payoff difference will be negative, positive, and negative respectively on $[0, x^l(s))$, $(x^l(s), x^h(s))$, and $(x^h(s), 1]$. As in the other cases, we can show that if $s' > s$, then $\Delta\pi(x^l(s)|s') > 0$ and $\Delta\pi(x^h(s)|s') > 0$. Then it must be true that $x^l(s') < x^l(s)$ and $x^h(s') > x^h(s)$. Under the dynamics, paths starting in $[0, x^l(s))$ will converge to $x_H = 0$, and paths starting in $(x^l(s), 1)$ will converge to $x^h(s)$. This means that if $x_H^0 < x^l(s)$, then $\pi^*(x^0, s) = 0$. If $x_H^0 > x^l(s)$, then

the long run payoff is given by

$$\begin{aligned}\pi^*(x^0, s) &= \pi_H(\bar{x}(s), s) = a - (a - c)m_{HL}/x^h(s) \quad \text{and also} \\ \pi^*(x^0, s) &= \pi_L(\bar{x}(s), s) = dm_{HL}/(1 - x^h(s))\end{aligned}$$

This payoff is positive since $d > a > 0$. Furthermore, because $a > 0 > c$, it is increasing in s by exactly the same reasoning as in Case 3. Thus, whenever $x^l(s)$ and $x^h(s)$ exist, $\pi^*(x^0, s)$ is strictly increasing in s for all x^0 greater than $x^l(s)$. Furthermore, the set $X^0(s) = (x^l(s), 1)$ for which long run payoffs strictly increase in s is growing in s .

Now examine the situation when $\sigma_{HL} = \sigma/s$, so $a < c$ and $d > 0$. Still taking $a > 0$, we then have $\Delta\pi(0|s) = c > 0$. This time there are two cases to consider.

Case 1: $a > d$

Then $c > a > d > 0$, so $\Delta\pi(x_H|s)$ is everywhere positive for all s , and the population converges to $x_H = 1$ and a long run payoff of a which is the greatest lower payoff of this game.

Case 2: $a < d$

$\Delta\pi(x_H|s)$ is positive on $[0, \bar{x}(s))$ and negative on $(\bar{x}(s), 1]$ for some $\bar{x}(s)$, so all interior paths converge to $\bar{x}(s)$. Now, because $\sigma_{HL} = \sigma/s$ and $\sigma_{HH} = \sigma_{LL} = \sigma$, $S = 1/s^2$, so m_{HL} is *increasing* in s . Thus, if $s' > s$, we have

$$a < m_{HL}(\bar{x}(s), s') \left(\frac{a - c}{\bar{x}(s)} + \frac{d}{1 - \bar{x}(s)} \right)$$

so $\Delta\pi(\bar{x}(s)|s') < 0$, which means that $\bar{x}(s') < \bar{x}(s)$. Then

$$\frac{m_{HL}(\bar{x}(s'), s')}{\bar{x}(s')} > \frac{m_{HL}(\bar{x}(s'), s)}{\bar{x}(s')} > \frac{m_{HL}(\bar{x}(s), s)}{\bar{x}(s)}$$

This means that $m_{HL}/\bar{x}(s)$ is increasing in s , so the long run payoff $\pi^*(x^0, s) = \pi_H(\bar{x}(s), s) = a + (c - a)m_{HL}/\bar{x}(s)$ is increasing in s as well, for all interior paths. This completes the proof.

Proof of Theorem 2

It remains to be shown that the result holds for games with $\sigma_{HL} = \sigma/s$. For such games, $a < c$ and $d > 0$, so we also have $\sigma_{HH} = \sigma_{LL} = \sigma$. There are three subcases.

Case 1: $d > a$ and $d > c$

Then c is the greatest lower payoff. I will show that $\hat{x} = (\frac{c}{c+d}, \frac{d}{c+d})$. Here the idea will be that m_{HL} can be made arbitrarily close to $\min\{x_H, x_L\}$ by taking s sufficiently large. This will mean that trajectories can be made to approach a neighborhood of \hat{x} that can be made arbitrarily small. Now m_{HL} is given by

$$m_{HL} = \frac{1 - \sqrt{1 - 4(1 - 1/s^2)x_H x_L}}{2(1 - 1/s^2)}$$

We want to bound the ratio $m_{HL}/\min\{x_H, x_L\}$. Notice that this ratio is symmetric about $x_H = 1/2$, so it suffices to look at the interval $0 \leq x_H \leq 1/2$ for which $\min\{x_H, x_L\} = x_H$. Differentiating m_{HL}/x_H and arranging terms produces

$$\frac{d(m_{HL}/x_H)}{dx_H} = \frac{(1 - \alpha x_H) - \sqrt{(1 - \alpha x_H)^2 + \alpha(2 - \alpha)x_H^2}}{\alpha x_H^2 \sqrt{(1 - \alpha x_H)^2 + \alpha(2 - \alpha)x_H^2}}$$

where $\alpha = 2(1 - 1/s^2)$. Because $s > 1$ (here we ignore $s = 1$ because the focus is on large s), $0 < \alpha < 2$, so the second term under the square root is strictly positive (for $x_H > 0$). Therefore, m_{HL}/x_H is strictly decreasing, so $m_{HL}/\min\{x_H, x_L\}$ is strictly decreasing on $(0, 1/2)$ and reaches its minimum at $x_H = 1/2$. At this point, the ratio is equal to $2m_{HL}(1/2, 1/s^2) = s/(1 + s)$, so $1 - m_{HL}/\min\{x_H, x_L\} \leq \gamma = 1/(1 + s)$.

Then for $x_H \leq 1/2$,

$$\begin{aligned} \Delta\pi(x_H, s) &= a + \frac{m_{HL}}{x_H}(c - a) - \frac{m_{HL}}{x_L}d \\ &= a + \frac{m_{HL}}{x_H}(c - a - \frac{x_H}{x_L}d) \\ &= c - \frac{x_H}{x_L}d - (1 - \frac{m_{HL}}{x_H})(c - a - \frac{x_H}{x_L}d) \\ &\geq c - \frac{x_H}{x_L}d - \gamma(|c - a| + |\frac{x_H}{x_L}d|) \\ &\geq c - \frac{x_H}{x_L}d - \gamma(c - a + d) \end{aligned}$$

The last step follows because $c > a$ and $d > 0$. Similarly

$$\Delta\pi(x_H, s) \leq c - \frac{x_H}{x_L}d + \gamma(c - a + d)$$

To complete the picture, note that for $x_H > 1/2$, where $\min\{x_H, x_L\} = x_L$,

$$\begin{aligned}
\Delta\pi(x_H, s) &= a + \frac{x_L}{x_H}(c-a) - d - \left(1 - \frac{m_{HL}}{x_L}\right)\left(\frac{x_L}{x_H}(c-a) - d\right) \\
&\leq a + \frac{x_L}{x_H}(c-a) - d + \gamma\left(\left|\frac{x_L}{x_H}(c-a)\right| + |d|\right) \\
&\leq c - d + \gamma(c-a+d)
\end{aligned}$$

Now pick some $x_H^0 \in (0, 1)$ and some arbitrary $\varepsilon > 0$. (To avoid unnecessary work, take ε small enough that $\hat{x}_H + \varepsilon < 1/2$.) Using the inequalities above, $\Delta\pi(x_H, s)$ will be positive on $(0, \hat{x}_H - \varepsilon]$ and negative on $[\hat{x}_H + \varepsilon, 1/2]$ whenever $\gamma < \frac{c+d}{c-a+d}\varepsilon$, and $\Delta\pi(x_H, s)$ will be negative on $(1/2, 1)$ whenever $\gamma < \frac{d-c}{c-a+d}$. Take s large enough that $\gamma < \min\{\frac{c+d}{c-a+d}\varepsilon, \frac{d-c}{c-a+d}\}$. Then $|x_H^*(x^0, s) - \hat{x}_H| < \varepsilon$. Thus $\lim_{s \rightarrow \infty} x^*(x^0, s) = \hat{x}$ for all interior x^0 . Then $\lim_{s \rightarrow \infty} \pi^*(x^0, s) = \lim_{s \rightarrow \infty} \pi_i(\hat{x}, s)$ for $i = 1, 2$. The payoff functions are continuous, so taking the limits inside,

$$\begin{aligned}
\lim_{s \rightarrow \infty} \pi_H(\hat{x}, s) &= a - (a-c) \lim_{s \rightarrow \infty} \frac{m_{HL}(\hat{x}, s)}{\hat{x}_H} \\
&= c
\end{aligned}$$

where the limit follows by using the fact that $m_{HL}/x_H \leq 1$ for all x and s along with the lower bound derived earlier:

$$\begin{aligned}
\lim_{s \rightarrow \infty} \frac{m_{HL}(\hat{x}, s)}{\hat{x}_H} &\geq \lim_{s \rightarrow \infty} \frac{m_{HL}(1/2, s)}{1/2} \\
&= \lim_{s \rightarrow \infty} \frac{s}{1+s} \\
&= 1
\end{aligned}$$

Similarly

$$\begin{aligned}
\lim_{s \rightarrow \infty} \pi_L(\hat{x}, s) &= d \frac{\hat{x}_H}{\hat{x}_L} \lim_{s \rightarrow \infty} \frac{m_{HL}(\hat{x}, s)}{\hat{x}_H} \\
&= c
\end{aligned}$$

so payoffs converge to the greatest lower payoff in the game.

Case 2: $d > a$ and $c > d$

The proof for this case is essentially the same as for Case 1, so I omit it.

Case 3: $a > d$

In this case, the greatest lower payoff is a and we have $0 < d < a < c$, so H is a dominant strategy. For any x_H and s , $\pi_H(x_H, s) \geq a > d \geq \pi_H(x_H, s)$, so x_H is everywhere positive.

Thus $x_H^*(x^0, s) = 1$ and $\pi^*(x^0, s) = a$ for all x^0 and s .

References

- [1] BALA, V. AND S. GOYAL (1998): “Learning from Neighbors,” *Review of Economic Studies*, 65, 595-621.
- [2] BERGIN, J. AND B. LIPMAN (1996): “Evolution with State-dependent Mutations,” *Econometrica*, 64, 943-56.
- [3] BHASKAR, V. AND F. VEGA-REDONDO (1998): “Migration and the Evolution of Conventions,” mimeo.
- [4] BINMORE, K., L. SAMUELSON, AND R. VAUGHAN (1995): “Musical Chairs: Modeling Noisy Evolution,” *Games and Economic Behavior*, 11, 1-35.
- [5] BURDETT, K. AND M. COLES (1998): “Separation Cycles,” *Journal of Economic Dynamics and Control*, 22, 1069-90.
- [6] DURLAUF, S. (2000): “The Memberships Theory of Poverty: the Role of Group Associations in Determining Socioeconomic Outcomes,” mimeo.
- [7] ELLISON, G. (1993): “Learning, Local Interaction, and Coordination,” *Econometrica*, 61, 1047-71.
- [8] ELLISON, G. AND D. FUDENBERG (1993): “Rules of Thumb for Social Learning,” *Journal of Political Economy*, 101, 612-43.
- [9] ELLISON, G. AND D. FUDENBERG (1995): “Word of Mouth Communication and Social Learning,” *Quarterly Journal of Economics*, 110, 93-125.
- [10] ELY, J. (1996): “Local Conventions,” mimeo.
- [11] ESHEL, I., L. SAMUELSON, AND A. SHAKED (1998): “Altruists, Egoists, and Hooligans in a Local Interaction Model,” *American Economic Review*, 88, 157-79.
- [12] FERNANDEZ, R. AND R. ROGERSON (2000): “Sorting and Long-run Inequality,” NBER Working Paper 7508.
- [13] FUDENBERG, D. AND D. LEVINE (1993): “Steady State Learning and Nash Equilibrium,” *Econometrica*, 61, 547-573.

- [14] FUDENBERG, D. AND E. MASKIN (1990): “Evolution and Cooperation in Noisy Repeated Games,” *American Economic Review*, 80, 274-79.
- [15] GHOSH, P. AND D. RAY (1996): “Cooperation in Community Interaction without Information Flows,” *Review of Economics Studies*, 63, 491-519.
- [16] JOHNSON, P., D. LEVINE, AND W. PESENDORFER (1998): “Evolution and Information in a Prisoner’s Dilemma Game,” mimeo.
- [17] KANDORI, M., G. MAILATH, AND R. ROB (1993): “Learning, Mutation, and Long Run Equilibria in Games,” *Econometrica*, 61, 383-414.
- [18] KIM, Y. AND J. SOBEL (1995): “An Evolutionary Approach to Pre-Play Communication,” *Econometrica*, 63, 1181-93.
- [19] KREMER, M. (1997): “How Much does Sorting Increase Inequality?,” *Quarterly Journal of Economics*, 112, 115-39.
- [20] MAYNARD SMITH, J. (1982): *Evolution and the Theory of Games*, Cambridge University Press.
- [21] MONTGOMERY (1999): “Adverse Selection and Employment Cycles,” *Journal of Labor Economics*, 17, 281-97.
- [22] OECHSSLER, J. (1997): “Decentralization and the Coordination Problem,” *Journal of Economic Behavior and Organization*, 32, 119-35.
- [23] PIKETTY, T. (1995): “Social Mobility and Redistributive Politics,” *Quarterly Journal of Economics*, 110, 551-84.
- [24] MAILATH, G., L. SAMUELSON, AND A. SHAKED (2000): “Endogenous Inequality in Integrated Labor Markets with Two-Sided Search,” *American Economic Review*, 90, 46-72.
- [25] ROBSON, A. AND F. VEGA-REDONDO (1996): “Long Run Efficiency in Evolutionary Games with Random Matching,” *Journal of Economic Theory*, 70, 65-92.
- [26] SETHI, R. (1998): “Strategy-specific Barriers to Learning and Nonmonotonic Selection Dynamics,” *Games and Economic Behavior*, 23, 284-304.
- [27] WEIBULL, J. (1995): *Evolutionary Game Theory*, MIT Press.
- [28] YOUNG, P. (1993): “The Evolution of Conventions,” *Econometrica*, 61, 57-84.