

A Trade Network Game with Endogenous Partner Selection

Leigh Tesfatsion*

Economic Report Series No. 36

Original Version: April 1995

Revised Version: May 1995

375 Heady Hall/ISU/Ames, IA 50011-1070

515-294-0138

tesfatsi@iastate.edu

Abstract: This paper develops a Trade Network Game (TNG) that combines evolutionary game play with endogenous partner selection. Resource-constrained buyers and sellers choose and refuse trade partners on the basis of continually updated expected payoffs. Partner selection takes place in accordance with a “deferred choice and refusal” mechanism that is shown to have interesting stability, optimality, and uniqueness properties. The iterated prisoner’s dilemma strategies used by buyers and sellers to conduct their trades are evolved over time via a genetic algorithm that biases reproduction in favor of strategies that have been successful in past trades. The TNG is shown to encompass a variety of economic applications, such as job search games, labor markets modelled as assignment games, labor markets with endogenously determined workers and employers, and double auction games. To illustrate the way in which preferential partner selection interacts with game play in the TNG, resulting in the endogenous formation of trade networks, a 5-trader TNG is analyzed in detail.

Keywords: Trade networks; evolutionary game; optimal search; endogenous interactions; iterated prisoner’s dilemma; multi-armed bandit game; assignment game; preferential partner matching; genetic algorithm; artificial life; criterion filtering.

Copyright © 1995 by Leigh Tesfatsion. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided this copyright notice appears on all such copies.

A Trade Network Game With Endogenous Partner Selection¹

LEIGH TEFATSION

*Department of Economics and Department of Mathematics
Iowa State University, Ames, IA 50011-1070
tesfatsi@iastate.edu; <http://www.econ.iastate.edu/tesfatsi/>
June 1996*

Abstract. This paper develops an evolutionary *trade network game* (TNG) that combines evolutionary game play with endogenous partner selection. Successive generations of resource-constrained buyers and sellers choose and refuse trade partners on the basis of continually updated expected payoffs. Trade partner selection takes place in accordance with a modified Gale-Shapley matching mechanism, and trades are implemented using trade strategies evolved via a standardly specified genetic algorithm. The trade partnerships resulting from the matching mechanism are shown to be core stable and Pareto optimal in each successive trade cycle. Nevertheless, computer experiments suggest that these static optimality properties may be inadequate measures of optimality from an evolutionary perspective.

1 Introduction

Evolutionary game studies typically focus on the optimality properties of strategy configurations when agents are matched randomly or deterministically by some extraneous device. The optimality properties of the matching mechanism per se are generally not considered ([12], [15]). In contrast, optimal search studies focus on the optimality properties of preference-based agent matching mechanisms, but generally these studies are set in static contexts [19]. Since actual social interactions are often characterized both by evolutionary dynamics and by preference-based partner selection, studying both aspects together seems a logical and interesting next step to take.

This issue is addressed in Stanley et al. [21]. The standard evolutionary iterated prisoner's dilemma (IPD) is extended to an evolutionary IPD with choice and refusal (IPD/CR) by allowing players to choose and refuse game partners in each iteration on the basis of continually updated expected payoffs.²

The introduction of choice and refusal fundamentally alters the way in which players interact in the IPD and the characteristics that result in high payoff scores. Choice allows players to increase their chances of encountering other cooperative players, and refusal gives players a way to protect themselves from defections without having to defect themselves. The ostracism of defectors occurs endogenously as an increasing number of players individually refuse the defectors' game offers. Nevertheless, choice and refusal also permit opportunistic players to home in quickly on exploitable players and form parasitic relationships.

The computer experiments reported in [21] and in the subsequent studies by Ashlock et al. [1], Smucker et al. [20], and Hauk [9] indicate that the emergence of mutual cooperation in the standard evolutionary IPD is accelerated by the introduction of choice and refusal of partners. The underlying player interaction patterns induced by choice and refusal can be complex and highly path dependent, however, even when expressed play behavior is largely cooperative. Consequently, it has proved difficult to characterize the mapping from parameter configurations to evolutionary outcomes for the IPD/CR.

¹To appear in H. Amman et al. (eds.), *Advances in Computational Economics*, Kluwer Book Series, 1997. This paper is based on Tesfatsion [23]. An abbreviated version of this paper appears in the Proceedings Volume of the Fifth Annual Conference on Evolutionary Programming. The author is grateful to D. Fogel, M. Smucker, and N. Vriend for helpful comments.

²Other game theory studies focusing on the endogenous determination of player interaction patterns, e.g., by permitting avoidance of unwanted interactions, endogenously determined probabilities of interaction, or evolution of interaction lengths, include Fogel [6], Guriev and Shakhova [8], Hirshleifer and Rasmusen [10], Kitcher [13], Mailath et al. [14], and Orbell and Dawes [18]. There is also a growing body of work by economists on multi-agent systems with endogenous interactions. See, for example, Brock and Durlauf [2], De Vany [3], Duong [4], Durlauf [5], Ionnides [11], and Vriend [24].

A potentially useful way to proceed, then, is to focus on more concrete settings which impose problem-specific constraints on agent interactions. In Tesfatsion [23] an evolutionary *trade network game* (TNG) is proposed for studying the interplay between evolutionary dynamics and endogenous preference-based partner selection under alternatively specified market structures.

The player set for the TNG consists of buyer and seller tradebots who choose and refuse trade partners on the basis of continuously updated expected payoffs. Buyers make trade offers to preferred sellers which the sellers either accept or refuse. A trade offer is an invitation to engage in a risky trade modelled as a two-player game. Each buyer and seller initially associates a prior expected payoff with each potential trade partner and adopts a random strategy for use in subsequent trades. The buyers and sellers then enter into a trade cycle loop consisting of successive rounds of partner matching, resource-constrained trading, and updating of expected payoffs. At the end of the trade cycle loop the buyers and sellers enter into an evolutionary step in which trade strategies successful in past trades are retained while trade strategies unsuccessful in past trades are replaced with variants of more successful strategies. A new trade cycle loop then commences.

The modular form of the TNG facilitates experimentation with alternative specifications for trade partner matching, trading, expectation updating, and trade strategy evolution. This paper presents preliminary experimental results obtained for the particular TNG module specifications outlined in Tesfatsion [23] using the C++ trade platform developed in McFadzean and Tesfatsion [17]. As will be clarified below, trade partners are determined in accordance with a *deferred choice and refusal* (DCR) mechanism, a modified Gale-Shapley matching mechanism [7] that retains the static optimality properties of the original Gale-Shapley mechanism. Expected payoffs are updated by means of a simple learning algorithm that yields consistent estimates. A trade is modelled as a prisoner’s dilemma game, and trade (IPD) strategies are evolved by means of a standardly specified genetic algorithm.

Two types of markets are considered: buyer-seller markets; and two-sided markets. In the buyer-seller market, each tradebot is both a buyer and a seller in the sense that he can both make and receive trade offers. In the two-sided market, the set of buyers (tradebots who can make offers) is disjoint from the set of sellers (tradebots who can receive offers). For each type of market, attention is focused on the average fitness scores achieved by the tradebots as the market evolves and on the degree to which the tradebots engage in mutually cooperative behavior.

One interesting finding concerns the high transactions costs that agents can suffer under Gale-Shapley type matching mechanisms due to large numbers of refused offers. Specifically, for certain parameter specifications, the negative refusal payoffs accumulated under the DCR mechanism in both buyer-seller and two-sided markets can result in lower average fitness scores for the tradebots relative to other less sophisticated matching mechanisms. Another interesting finding concerns the evolutionary effects of the bias of Gale-Shapley type matching mechanisms in favor of those who actively make offers. Under the DCR mechanism, which inherits this bias, buyers in two-sided markets with relatively large seller acceptance quotas appear to be able to form long-term parasitic relations with sellers that reduce seller fitnesses relative to buyer fitnesses and hinder the emergence of mutually cooperative behavior. Overall, these experimental findings suggest that the conventional optimality properties used to evaluate agent matching mechanisms in static market contexts may be inadequate measures of optimality from an evolutionary perspective.

The TNG module specifications are detailed in Section 2. Static optimality properties of the DCR matching mechanism are developed in Section 3, and the evolutionary implications of the DCR matching mechanism are investigated in Section 4 by means of computer experiments. Concluding comments are given in Section 5.

2 The TNG Model

The set of players for the TNG is the union $V = B \cup S$ of a nonempty subset B of *buyer tradebots* who can submit trade offers and a nonempty subset S of *seller tradebots* who can receive trade offers, where B and S may be disjoint, overlapping, or coincident. For example, the buyers and sellers might represent customers

		Player 2	
		c	d
Player 1	c	(C,C)	(L,H)
	d	(H,L)	(D,D)

Table 1. Payoff Matrix for the Prisoner's Dilemma Game

and retail store owners, workers and employers, borrowers and lenders, or auction traders.

Each generation of tradebots participates in a *trade cycle loop* consisting of a fixed number of trade cycles. In each trade cycle, each buyer m can submit up to O_m trade offers to sellers, and each seller n can accept up to A_n trade offers from buyers, where O_m and A_n are strictly positive. One interpretation for the buyer offer quota O_m is that buyer m has a limited amount of resources (credit, labor time, collateral,...) to trade in exchange for other items, and one interpretation for the seller acceptance quota A_n is that seller n has a limited amount of items (goods, job openings, loans,...) to provide.

The tradebots determine their submission, acceptance, and refusal of trade offers in each trade cycle using a modified version of the well-known Gale-Shapley deferred acceptance mechanism [7]. This modified mechanism, hereafter referred to as the *deferred choice and refusal* (DCR) mechanism, is discussed more carefully in Section 3 below. During the DCR matching process, any tradebot that has a trade offer refused is immediately penalized by receipt of a negative *refusal payoff*, R ; the tradebot who does the refusing is not penalized. If a tradebot neither submits nor accepts any trade offers during this matching process, he receives a *wallflower payoff*, W .

A trade offer is an offer by a buyer to a seller to participate in a risky trade modelled as a prisoner's dilemma (PD) game. For example, the trade may involve the exchange of a good or service of a certain promised quality in return for a loan or wage contract entailing various payment obligations. A buyer participating in a trade may either cooperate (fulfill his trade obligations) or defect (renege on his trade obligations), and similarly for a seller. The range of possible payoffs is the same for each trade in each trade cycle: namely, L (the sucker payoff) is the lowest possible payoff, received by a cooperative tradebot whose trade partner defects; D is the payoff received by a defecting tradebot whose trade partner also defects; C is the payoff received by a cooperative tradebot whose trade partner also cooperates; and H (the temptation payoff) is the highest possible payoff, received by a defecting tradebot whose trade partner cooperates. More precisely, the payoffs are assumed to satisfy $L < D < 0 < C < H$, with $(L + H)/2 < C$. The payoff matrix for the PD game is depicted in Table 1.

The trade behavior of each tradebot, whether he is a pure buyer in $V - S$, a buyer-seller in $B \cap S$, or a pure seller in $V - B$, is characterized by a finite-memory pure strategy for playing a PD game with an arbitrary partner an indefinite number of times, hereafter referred to as a *trade strategy*. Each tradebot thus has a distinct trading personality even if he engages in both buying and selling activities. No tradebot knows any other tradebot's strategy *a priori*; he can only learn about it by engaging the other tradebot in repeated trades and observing the payoff history that ensues. Moreover, each tradebot's choice of an action in a current trade with a potential trade partner is determined entirely on the basis of the payoffs obtained in past trades with this same partner.

At the beginning of the initial trade cycle loop, before any actual trades have taken place, each tradebot v associates an exogenously given *prior expected payoff* $U_v^0(k)$ with each potential trade partner k . Throughout each trade cycle, tradebot v then uses a simple variable-gain criterion filter [22] to update his current expected payoffs on the basis of the new payoffs he obtains from interactions with his potential trade partners. In particular, if tradebot v receives a payoff P from an interaction with a potential trade partner k , then v

```

int main () {
    tngInit() ;           // Construct initial tradebots
                           // with prior expected payoffs.
    For (G = 0,...,GMAX-1) { // Enter the generation loop.
        For (I = 0,...,IMAX-1) { // Enter the trade cycle loop.
            MatchTraders();      // Determine trade partners,
                                // given expected payoffs,
                                // and record refusal and
                                // wallflower payoffs.
            Trade();             // Implement trades and
                                // record trade payoffs.
            UpdateExp();         // Update expected payoffs
                                // using newly recorded payoffs.
        }
        AssessFitness();        // Record fitness scores.
        If (G < GMAX-1) {
            EvolveGen();         // Genetic step: Evolve the
                                // current tradebots.
        }
    }
    Return 0 ; }

```

Table 2. TNG Pseudo Code

forms an updated expected payoff for k by taking a convex combination of this new payoff P and his previous expected payoff for k , where the inverse of the weight on P is 1 plus tradebot v 's current payoff count with k . In this way, tradebot v keeps a running tab on the payoff outcomes of his interactions with k . As explained in Tesfatsion [23, Section 5], this updating procedure guarantees that the expected payoff tradebot v associates with k converges to the true average payoff tradebot v attains from interactions with k as the number of interactions between v and k becomes arbitrarily large.

At the end of a trade cycle loop, the fitness score of each tradebot is calculated as the average of all of the trade, refusal, and wallflower payoffs he received during this loop. The tradebots then enter into a *genetic step* in which the trade strategies used by the *elite* (the fittest tradebots) are retained while the trade strategies used by the remaining tradebots are replaced with variants of more successful trade strategies. A detailed discussion of the finite state machine (FSM) and standardly specified genetic algorithm (GA) used in the genetic step to represent and evolve the tradebots' trade strategies can be found in McFadzean and Tesfatsion [17]. Briefly, the GA is applied to a population of bit strings that encode the trade strategies (FSMs) of the tradebots. The GA employs two-point crossover with bit mutation, and parent strategies are chosen probabilistically in proportion to fitness. The particular FSM and GA parameter settings used in all experiments reported in this paper are detailed in the next section.

At the end of the genetic step the memories of the tradebots are reset to zero, their associated FSMs are reset to a fixed initial state, and their expected payoffs are reset to the prior expected payoff levels. A new trade cycle loop then commences. Table 2 depicts the overall logical progression of the TNG in pseudo-code.

Before reporting on analytical and simulation findings regarding the optimality properties of the DCR matching mechanism, three special cases of the TNG will be sketched to indicate the range of economic applications it encompasses.

Case 1: A Labor Market Modelled as an Assignment Game with Choice and Refusal

The set B consists of workers and the set S consists of employers, where B and S are disjoint. Each worker m can make work offers to a maximum of O_m employers, or he can choose to be unemployed and receive the known payoff W . Each employer n can hire up to A_n workers, and employers can refuse work offers. Once matched, a worker and employer engage in work site interactions modelled as a PD game.

This TNG special case extends the usual assignment problem [19] by incorporating subsequent strategic game play between matched pairs of agents and by having game play iterated over time. Assignment problems are commonly used by economists to model job-matching in labor markets as well as other economic processes,

but the payoff outcome for each matched pair of agents is usually specified a priori.

Case 2: A Labor Market with Endogenously Determined Workers and Employers

The subsets B and S coincide, implying that each tradebot can both make and receive trade offers. Each tradebot v can make up to O_v work offers to tradebots at other work sites and receive up to A_v work offers at his own work site. The degree to which any accepted work offer results in satisfactory outcomes for the participant tradebots is determined by subsequent PD game play. Ex post, four pure types of tradebots can emerge: (1) pure workers, who work at the sites of other tradebots but have no tradebots working for them at their own sites; (2) pure employers, who have tradebots working for them at their own sites but who do not work at the sites of other tradebots; (3) unemployed tradebots, who make at least one work offer to a tradebot at another site but who end up neither working at other sites nor having tradebots working for them at their own sites; and (4) inactive (out of the work force) tradebots, who neither make nor accept any work offers.

Case 3: Intermediation with Choice and Refusal

The buyer subset B and the seller subset S overlap but do not coincide. The pure buyers in $V - S$ are the depositors (lenders), the buyer-sellers in $B \cap S$ are the intermediaries (banks), and the pure sellers in $V - B$ are the capital investors (borrowers). The depositors offer funds to the intermediaries in return for deposit accounts, and the intermediaries offer loan contracts to the capital investors in return for a share of earnings. The degree to which accepted offers are satisfactorily fulfilled is determined by subsequent PD game play.

3 Determination of TNG Trade Partners

The original Gale-Shapley matching mechanism [7] was designed for two-sided markets in which one subset of agents makes offers to a second disjoint subset of agents, with all offer quotas equal to 1. As detailed in Roth and Sotomayor [19], this mechanism continues to be the subject of numerous theoretical studies, and it has found practical application as well—most notably to the labor market for medical interns and residents.

The DCR mechanism extends the original Gale-Shapley mechanism to the more general TNG context in which the subsets of buyers (makers of offers) and sellers (receivers of offers) can be coincident, overlapping, or disjoint, with arbitrarily specified offer and acceptance quotas. The TNG tradebots use the DCR mechanism to determine their submission, acceptance, and refusal of trade offers conditional on their current preference rankings over potential trade partners. This section first provides a more careful description of DCR mechanism. It is then shown that the DCR mechanism possesses the usual static optimality properties associated with Gale-Shapley type matching mechanisms: namely, core stability, and Pareto optimality for the subset of agents who actively make offers. Although trades in the TNG are currently modelled as prisoner’s dilemma games, neither the definition of the DCR mechanism nor its static optimality properties depend on this specification.

For any tradebot v in V , let τ_v denote tradebot v ’s *minimum tolerance level*, and let $U_v(k)$ denote the expected payoff that tradebot v currently associates with a potential trade partner k . Tradebot k will be called *tolerable* for tradebot v if and only if $U_v(k) \geq \tau_v$. Also, given any two potential trade partners k' and k'' , tradebot v will be said to *strictly prefer* to trade with k' rather than with k'' if $U_v(k') > U_v(k'')$ and to *equally prefer* trades with k' and k'' if $U_v(k') = U_v(k'')$.

The DCR Mechanism In Any Trade Cycle T:

Step 1: Each buyer m submits trade offers to a maximum of O_m most preferred sellers that he finds tolerable, with at most one trade offer going to any one seller. Each seller n then selects up to A_n most preferred trade offers received from tolerable buyers and places these offers on a waiting list. Seller n refuses all other trade offers received in Step 1. Buyers and sellers use random selection to break ties among equally-preferred trades.

Step K ($K \geq 2$): If r sellers refused trade offers from buyer m in Step $K - 1$, then buyer m submits up to r replacement trade offers in Step K to the most preferred and tolerable sellers who have not refused a trade offer from him during Steps $1, \dots, K - 1$. Each seller n then selects up to A_n most preferred and tolerable trade offers from among both the new trade offers he receives in Step K and the trade offers already on his waiting list from Step $K - 1$, and he refuses the rest. These selected trade offers constitute seller n 's waiting list for Step K . Buyers and sellers use random selection to break ties among tolerable equally-preferred trades.

Final Step K^ :* Step K^* is the final step if and only if no new trade offers are submitted in Step $K^* + 1$. If Step K^* is the final step, each seller accepts all trade offers on his Step K^* waiting list.

Under the DCR mechanism, two tradebots v and k can engage in zero, one, or two trades with each other in each trade cycle. Zero trades occur when neither tradebot submits an accepted trade offer to the other; one trade occurs when only one of the tradebots submits an accepted trade offer to the other; and two trades occur when each tradebot submits an accepted trade offer to the other. Note that this determination of trade partners requires the tradebots to pass messages back and forth to each other at event-driven times, a feature not present in standard economic models. This requirement is easily implemented by an object-oriented programming language such as C++.

Step K^* is the last step of the DCR mechanism in trade cycle T if and only if each buyer m either has had no trade offer refused in Step K^* , and hence has no desire to submit a new trade offer in Step $K^* + 1$, or is unable to submit a new trade offer in Step $K^* + 1$ because each seller that buyer m finds tolerable has already received a trade offer from him sometime during Steps $1, \dots, K^*$. The first proposition provides an upper bound for K^* . The proofs for this proposition, and for all subsequent propositions, are detailed in Tesfatsion [23, Section 4].

PROPOSITION 3.1. *Let M denote the number of buyers in the buyer set B , and let N denote the number of sellers in the seller set S . In each trade cycle T , the DCR mechanism always ends at some Step K^* occurring on or before Step MN .*

DEFINITION 3.1. A buyer m and seller n will be said to form a *match* in trade cycle T , denoted by (m, n) , if buyer m is tolerable for seller n , seller n is tolerable for buyer m , and buyer m and seller n engage in a trade during trade cycle T . A *matching outcome for trade cycle T* is then any listing of distinct matches (m, n) for trade cycle T such that each buyer m appears at most O_m times in the first position and each seller n appears at most A_n times in the second position. By construction, then, a matching outcome for trade cycle T is a list of trades that could feasibly take place during this cycle.

Note that the trades determined in any trade cycle T by the DCR mechanism always constitute a matching outcome for trade cycle T in the sense of Definition 3.1.

DEFINITION 3.2. (cf. Gale and Shapley [7, p. 10]) A matching outcome G for trade cycle T will be called *unstable* if the following conditions hold for some buyer m and seller n :

- (a) m and n are *not* matched under G ;
- (b) m would prefer to be matched with n , in the sense that either there exists a match (m, n') under G such that $U_m(n) > U_m(n')$, or m is matched with fewer than O_m sellers under G and $U_m(n) \geq \tau_m$.
- (c) n would prefer to be matched with m , in the sense that either there exists a match (m', n) under G such that $U_n(m) > U_n(m')$, or n is matched with fewer than A_n buyers under G and $U_n(m) \geq \tau_n$.

A matching outcome G for trade cycle T will be called (*pairwise*) *stable* if no such buyer m and seller n exist.

Intuitively, given conditions (a), (b), and (c), the buyer m and seller n could block the implementation of the matching outcome G by agreeing to play each other instead. Roth and Sotomayor [19] are careful to

differentiate between group stable and pairwise stable matching outcomes, whereby the former they mean a matching outcome that is not blocked by *any* coalition of agents. In the TNG, each trader is assumed to have a preference ranking only over pairwise matches with other traders, and this preference ranking is unaffected by who else might be trading with these potential trade partners. Consequently, a matching outcome in the TNG is group stable if and only if it is also pairwise stable. Without risk of confusion, then, the qualifier “pairwise” will be omitted in all subsequent references to stable matching outcomes.

PROPOSITION 3.2. *Any matching outcome generated via the DCR mechanism is stable.*

The final proposition shows that the matching outcome generated by the DCR mechanism in each trade cycle is a best possible matching outcome from the viewpoint of the buyers, i.e., the tradebots who actively make offers, as long as buyers and sellers have strict preferences over their potential trade partners. The proof is a generalization of the optimality proof provided in Gale and Shapley [7, Thm. 2, p. 14] for their college admissions problem with B disjoint from S and offer quotas $O_m = 1$ for all agents m in B ; see also Roth and Sotomayor [19, Chapter 5].

DEFINITION 3.3. A matching outcome G will be called *B-optimal* if: (a) G is a stable matching outcome; and (b) each buyer matched under G is at least as well off under G as under any other stable matching outcome.

PROPOSITION 3.3. *Suppose each buyer (seller) in some trade cycle T has a strict preference ranking over the sellers (buyers) he finds tolerable. Then the DCR mechanism yields the unique B-optimal matching outcome for trade cycle T .*

4 TNG Computer Experiments

In the previous section it is shown that the trade partnerships generated via the DCR mechanism exhibit the usual static optimality properties associated with Gale-Shapley type matching mechanisms. Still to be determined, however, is the extent to which the DCR mechanism induces a high average payoff for the tradebots over evolutionary time. In particular, to what extent does the DCR mechanism encourage the emergence and persistence of mutually cooperative behavior?

Four types of computer experiments are discussed in the present section: (a) buyer-seller market experiments with sellers unconstrained by acceptance quotas; (b) buyer-seller market experiments with seller acceptance quotas set to 1; (c) two-sided market experiments with sellers unconstrained by acceptance quotas; and (d) two-sided market experiments with seller acceptance quotas set to 1. All experimental findings reported below were obtained using *TNG*, a C++ trade platform developed by McFadzean and Tesfatsion [17] which in turn is supported by the C++ abstract base classes developed by McFadzean [16] for a general artificial life platform, *SimBioSys*. These findings are preliminary in the sense that only average fitness scores are considered. A more thorough understanding of these findings will require delving more deeply into the underlying trade networks and the trade strategies that support these trade networks.

For each type of experiment, multiple runs from different initial random seeds are reported. The following features are set commonly across all of these experimental runs. The wallflower payoff W is set at 0, the refusal payoff R is set at -0.6 , the PD trade payoffs are set at $L = -1.6$, $D = -0.6$, $C = 1.4$, and $H = 3.4$, and each tradebot’s minimum tolerance level is set at $\tau = 0$. Each tradebot assigns the same prior expected payoff, $U^0 = C$, to each other tradebot, implying that he is initially indifferent concerning which trade partners he interacts with; and each tradebot assigns a negative prior expected payoff to himself, thus ensuring that he never trades with himself. Each buyer tradebot has an offer quota of 1, meaning that he can have at most one trade offer outstanding to sellers at any given time. The total number of tradebots is set at 24, and the 16 most fit tradebots in each generation are taken to be the elite. The number of trade cycles in each trade cycle loop is set at 150, and the number of generations is set at 50.

Each trade strategy is represented by a 16-state FSM with a fixed initial state and with memory 1. At the beginning of the first trade cycle loop, a bit string coding for each FSM is randomly generated. At

the end of each trade cycle loop, the current population of trade strategies (FSMs coded as bit strings) is evolved by means of a genetic algorithm employing two-point crossover and bit mutation. The probability of crossover is set at 1.0 and the probability of a bit mutation is set at 0.005.

4.1 Buyer-Seller Markets

Each tradebot in these experiments was both a buyer and a seller, implying that he could both make and receive trade offers.

In the first batch of buyer-seller experiments, the acceptance quota of each tradebot was set at 24, the total number of tradebots. Since offer quotas in these experiments were set at 1, the tradebots were then effectively unconstrained with regard to the number of trade offers they could have on their waiting lists at any given time.

As a benchmark, experiments were first run with random partner matching in place of the DCR matching mechanism. Random partner matching was effected by preventing the updating of the prior expected payoff $U^0 = C$ that each tradebot initially assigned to each potential trade partner, so that all tradebots remained indifferent concerning their potential trade partners and matching was accomplished by the default mechanism of a random draw. Although occasionally the average fitness score achieved by the tradebots under random matching rose to the mutual cooperation level, 1.4, a more typical outcome was a steady decline to the mutual defection level, -0.6 ; see Fig. 1.³ The size of the refusal payoff is irrelevant for this finding, since refusals never occur in TNG experiments with random matching and nonbinding acceptance quotas.

When the DCR matching mechanism was then restored, the average fitness score achieved by the tradebots typically evolved to the mutual cooperation level 1.4; see Fig. 2. These TNG experiments reinforce the previous IPD/CR findings of Stanley et al. [21] and Ashlock et al. [1] that a preference-based matching mechanism tends to accelerate the emergence of mutual cooperation in the IPD when each agent is permitted both to make and to refuse game offers, is unconstrained with regard to the number of received offers he can accept, and is permitted to have at most one offer outstanding at any given time.

In the second batch of buyer-seller experiments, the acceptance quotas were reduced from 24 to 1. Each tradebot could thus retain at most one trade offer on his waiting list at any one time; all other received trade offers had to be refused. Under random partner matching, the typical outcome was again the emergence of an average fitness score close to the mutual defection payoff level, -0.6 . This same outcome obtained even when refusal payoffs were omitted from fitness scores, implying that refusal payoffs resulting from limited waiting lists were not a determining factor.

When the DCR matching mechanism was restored, however, the average fitness score typically leveled out at about 1.25 instead of evolving to the mutual cooperation payoff level 1.4, the outcome for the first batch of buyer-seller experiments. The explanation for this difference appears to lie in the changed nature of the refusal payoffs.

In the first batch of buyer-seller experiments, the acceptance quota (24) was large relative to the offer quota (1). In these circumstances, tradebots are generally refused by other tradebots only if the latter find them to be intolerable because of past defections. Negative refusal payoffs received in response to defections should rightly count against the fitness of the trade strategies generating the defections, for this induces changes in these strategies in the genetic step that tend to lead to higher future fitness scores. In the second batch of buyer-seller experiments, however, the acceptance quota (1) was much smaller in relation to the offer quota (1), implying that many more received trade offers had to be refused regardless of their desirability. In these circumstances, tradebots tend to accumulate large numbers of negative refusal payoffs purely as a consequence of the relatively small acceptance quota and the nature of the DCR mechanism, regardless of their trade strategies. Since neither the quotas nor the DCR mechanism are evolved in the

³As previously noted, the initial population of strategies for each TNG experiment was randomly generated and each run consisted of only 50 generations with interactions between any two players limited to at most 150 separate encounters per generation. Thus, this evolution of mutual defection is not in conflict with the evolution of mutual cooperation observed for the IPD with random or round-robin matching by previous researchers when the initial strategy population was sufficiently seeded with cooperatively inclined strategies such as Tit-for-Tat or when much longer player interaction lengths were permitted.

current implementation of the TNG, penalizing the tradebots for these quota and DCR effects by including refusal payoffs in their individual fitness scores tends to lower their current average fitness score without inducing a higher average fitness score in the future.

As expected, the average fitness scores attained by the tradebots in the second batch of buyer-seller experiments markedly improved when refusal payoffs were removed from the calculation of the tradebots' individual fitness scores; see Fig. 3. Improvement continued to occur when, in addition, the refusal payoffs were reduced in magnitude from -0.60 to -0.30 ; but a further reduction in magnitude to -0.06 and then to 0 resulted in increasingly volatile maximum and minimum average fitness scores with no discernible improvement in average fitness scores.

The probable cause of this increased volatility is that tradebots receiving refusals during initial trade cycles have little incentive to direct their offers elsewhere in subsequent trade cycles when the magnitude of the negative refusal payoff is small. A negative refusal payoff guarantees that the continually updated expected payoff that a tradebot associates with another tradebot who repeatedly refuses him eventually falls below 0 , the minimum tolerance level, at which point he ceases making offers to this other tradebot. Nevertheless, this learning process is slow when the magnitude of the negative refusal payoff is small, and it is non-existent when the refusal payoff is simply set to 0 .

4.2 Two-Sided Markets

In each two-sided market experiment, the 24 tradebots were evenly divided into 12 pure buyers (makers of trade offers) and 12 pure sellers (receivers of trade offers).

In the first batch of experiments, the acceptance quota of each seller was set to 12. Since the offer quota of each buyer was set to 1 in all of these experiments, sellers were thus effectively unconstrained regarding the number of trade offers they could have on their waiting lists at any one time. Experiments were first run with random partner matching in place of the DCR matching mechanism to obtain a benchmark for comparison. Interestingly, in contrast to buyer-seller experiments with nonbinding acceptance quotas and random matching, the average fitness score attained by the tradebots tended to fall to a level between -0.4 and the wallflower payoff 0 rather than dropping all the way down to the mutual defection payoff level -0.6 ; compare Fig. 4 with Fig. 1.

When the DCR matching mechanism was restored, the average fitness score of the tradebots typically evolved to about 1.2 , a payoff level markedly below the mutual cooperation level 1.4 obtained in buyer-seller experiments with nonbinding acceptance quotas and DCR matching. Moreover, the maximum fitness score, the average fitness score, and the minimum fitness score attained by the successive tradebot generations persistently deviated from one another. Compare Fig. 5 with Fig. 2.

As seen in Proposition 3.3, the DCR mechanism is only guaranteed to be Pareto optimal for buyers, that is, for the active makers of trade offers. The effects of this bias on average fitness scores is hidden in buyer-seller markets, where each tradebot is both a buyer and a seller. However, for two-sided markets with buyer offer quotas set equal to 1 and with nonbinding seller acceptance quotas, the DCR mechanism appears to result in a “separating equilibrium” in which the buyers are generally achieving high fitness scores and the sellers are generally achieving low fitness scores. In particular, the extreme pickiness of buyers combined with the acceptance by sellers of all tolerable received trade offers appears to allow buyers to form long-run parasitic relations with sellers, i.e., relations characterized by successful defections within the limits permitted by the sellers' 0 minimum tolerance levels.

In the second batch of two-sided market experiments, the seller acceptance quotas were decreased from 12 to 1. Thus, instead of accepting all tolerable received trade offers, the sellers now accepted at most one received trade offer per trade cycle.

When benchmark experiments were first run with random partner matching in place of the DCR mechanism, the typical outcome was the emergence of an average attained fitness score close to the mutual defection payoff, -0.6 . This result obtained whether or not refusal payoffs were counted in the calculation of individual fitness scores. When the DCR matching mechanism was then restored, with refusal payoffs counted in the calculation of individual fitness scores, the accumulation of refusal payoffs tended to result in an average

attained fitness score that was markedly below the mutual cooperation payoff level. When refusal payoffs were then omitted from the calculation of individual fitness scores, the average attained fitness score tended to evolve to the mutual cooperation level and to be close to the maximum attained fitness scores; see Fig. 6.

Given the structuring of the two-sided market experiments currently under discussion, with equal numbers of buyers and sellers, common acceptance and offer quotas, and tradebots who are initially indifferent among their potential trade partners, the buyer offer quota roughly predicts the number of trade offers that each seller will receive in the initial few trade cycles. The larger this number is, the more chance there is that mutually cooperative matches between buyers and sellers will be quickly discovered. On the other hand, if the seller acceptance quota is small relative to the buyer offer quota, then many buyers will accumulate negative refusal payoffs unrelated to the nature of their trade strategies; and if the seller acceptance quota is large relative to the buyer offer quota, opportunistic buyers have a greater chance to seek out and exploit sellers within the limits allowed by their minimum tolerance levels. Either circumstance could slow or even prevent the sustained emergence of mutually cooperative behavior.

It might therefore be conjectured that mutually cooperative behavior will best be induced in the current experimental setting when the seller acceptance quota and the buyer offer quota are equal and sufficiently large. Indeed, this turns out to be the case. In various two-sided market experiments with 12 pure buyers, 12 pure sellers, and equal seller and buyer quotas ranging from 3 to 12, the average fitness score attained by the tradebots tended to evolve to the mutual cooperation payoff level and to be close to the maximum attained fitness score even when refusal payoffs were included in the calculation of individual fitness scores.

5 Concluding Remarks

As seen in Section 3, the DCR mechanism inherits the usual static optimality properties associated with Gale-Shapley type matching mechanisms: namely, core stability, and Pareto optimality for the subset of agents who actively make offers. However, the TNG context also permits an examination of the DCR mechanism's optimality properties over evolutionary time; and the computer experiments reported in Section 4 suggest that the DCR mechanism has at least two potential problems from this evolutionary perspective.

First, if seller acceptance quotas are small relative to buyer offer quotas, the DCR mechanism tends to impose heavy transactions costs on buyers in the form of accumulated negative refusal payoffs that are unrelated to the social desirability of their trade strategies. Second, if seller acceptance quotas are large relative to buyer offer quotas, the inherent bias of the DCR mechanism in favor of buyers appears to allow buyers in two-sided markets to parasitize sellers, resulting in an average population fitness score that remains markedly below the mutually cooperative payoff level.

Systematic experimentation is needed to determine the true extent and seriousness of these problems. For example, the exogenously specified offer and acceptance quotas should more realistically be allowed to change over time to reflect both changing strategic considerations and changing budget and production relations. Also, the trade strategies for pure buyers and for pure sellers should be separately evolved since these tradebots face intrinsically asymmetric choice problems. These modifications may alleviate some of the difficulties with the DCR mechanism that arose in the Section 4 computer experiments with exogenously given quotas and pooled trade strategy evolution.

Finally, the mechanism by which trade partners are determined should itself be allowed to evolve over time along with the quotas and trade strategies. It will be interesting to see how Gale-Shapley type matching mechanisms fare in these less structured evolutionary contexts.

References

- [1] Ashlock, D., Smucker, M. D., Stanley, E. A., and Tesfatsion, L., Preferential partner selection in an evolutionary study of prisoner's dilemma, *BioSystems* 37 (1996), 99–125.

- [2] Brock, W. A., and Durlauf, S. N., Discrete choice with social interactions, Working Paper, Department of Economics, University of Wisconsin-Madison, 1995.
- [3] De Vany, A., The emergence and evolution of self-organized coalitions, in M. Gilli (ed.), *Computational Methods in Economics and Finance*, Kluwer Academic Publishers, Norwell, MA, 1996.
- [4] Duong, D., Symbolic interactionist modeling: The coevolution of symbols and institutions, in J. A. Meystel and R. Quintero (eds.), *Intelligent Systems: A Semiotic Perspective*, Proceedings Volume, National Institute for Standards and Technology, 1996.
- [5] Durlauf, S. N., Neighborhood feedbacks, endogenous stratification, and income inequality, in W. A. Barnett, G. Gandolfo, and C. Hillinger (eds.), *Disequilibrium Dynamics: Theory and Applications*, Cambridge University Press, to appear.
- [6] Fogel, D. B. On the relationship between the duration of an encounter and the evolution of cooperation in the iterated prisoner's dilemma, *Evolutionary Computation* 3 (1995), 349–363.
- [7] Gale, D. and Shapley, L., College admissions and the stability of marriage, *American Mathematical Monthly* 69 (1962), 9–15.
- [8] Guriev, G., and Shakhova, M., Self-organization of trade networks in an economy with imperfect infrastructure, Proceedings of the International Workshop on Agents, Cooperation, and Communication, San Marco di Castellabate, Italy, 1995.
- [9] Hauk, E., Leaving the prison: a discussion of the iterated prisoner's dilemma under preferential partner selection, Working Paper, European University Institute, Florence, 1995.
- [10] Hirshleifer, D., and Rasmusen, E., Cooperation in a repeated prisoner's dilemma with ostracism, *Journal of Economic Behavior and Organization* 12 (1989), 87–106.
- [11] Ionnides, Y. M., Evolution of trading structures, Working Paper, Department of Economics, Tufts University, Boston, 1995.
- [12] Kirman, A., Economies with interacting agents, Working Paper No. 94-05-030, Santa Fe Institute, NM, 1994.
- [13] Kitcher, P., Evolution of altruism in repeated optional games, Working Paper, Department of Philosophy, UC San Diego, 1992.
- [14] Mailath, G., Samuelson, L., and Shaked, A., Evolution and endogenous interactions, SSRI Working Paper 9426, University of Wisconsin-Madison, 1994.
- [15] Lindgren, K., and Nordahl, M. G., Cooperation and community structure in artificial ecosystems, *Artificial Life* 1 (1994), 15–37.
- [16] McFadzean, D., *SimBioSys: A class framework for evolutionary simulations*, Master's Thesis, Department of Computer Science, University of Calgary, Alberta, Canada, 1995.
- [17] McFadzean, D., and Tesfatsion, L., A C++ platform for the evolution of trade networks, to appear in *Computational Economics*, Special Issue on Programming Languages.
- [18] Orbell, J. M., and Dawes, R. M., Social welfare, cooperators advantage, and the option of not playing the game, *American Sociological Review* 58 (1993), 787–800.
- [19] Roth, A., and Sotomayor, M. A. O., *Two-sided matching: A study in game-theoretic modeling and analysis*, Cambridge University Press, Cambridge, 1990.

- [20] Smucker, M. D., Stanley, E. A., and Ashlock, D., Analyzing social network structures in the iterated prisoner's dilemma with choice and refusal, Department of Computer Sciences Technical Report CS-TR-94-1259, University of Wisconsin-Madison, 1994.
- [21] Stanley, E. A., Ashlock, D., and Tesfatsion, L., Iterated prisoner's dilemma with choice and refusal of partners, in C. Langton (ed.), *Artificial Life III*, Proceedings Volume 17, Santa Fe Institute Studies in the Sciences of Complexity, Addison Wesley, 1994, pp. 131–175.
- [22] Tesfatsion, T., Direct updating of intertemporal criterion functions for a class of adaptive control problems, *IEEE Transactions on Systems, Man, and Cybernetics* SMC-9 (1979), 143–151.
- [23] Tesfatsion, L., A trade network game with endogenous partner selection, Economic Report No. 36, Iowa State University, Ames, 1995.
- [24] Vriend, N., Self-organization of markets: An example of a computational approach, *Computational Economics* 8 (1995), 205–231.

Figure 1: Buyer-seller average fitness with random matching and seller quotas equal to 24.

Figure 2: Buyer-seller average fitness with DCR matching and seller quotas equal to 24.

Figure 3: Buyer-seller average fitness with DCR matching, seller quotas equal to 1, and refusal payoffs omitted from fitness scores.

Figure 4: Two-sided market average fitness with random matching and seller quotas equal to 12.

Figure 5: Two-sided market average fitness with DCR matching and seller quotas equal to 12.

Figure 6: Two-sided market average fitness with DCR matching, seller quotas equal to 1, and refusal payoffs omitted from fitness scores.