Equilibrium Selection in Repeated Business-to-Business Matching Markets*

Ernan Haruvy†
University of Texas at Dallas

M. Utku Ünver‡
Koç University

This Draft: November 2004

Abstract

A two-sided matching framework is applied to repeated business-to-business procurement matches. Both static and dynamic solutions concepts— namely Gale-Shapley deferred acceptance algorithm, learning dynamics, and genetic algorithms— are used to obtain solutions. The settings under investigation include both full information and limited information settings. We show that under certain conditions the dynamic predictions refine the core of the matching market. According to the theoretical predictions, organizational buyers would be better off in buyer-proposing settings than in seller-proposing settings, whereas sellers would prefer the opposite. The salience of this theoretical prediction is tested and confirmed using experimental data under different environments.

*We would like to thank Paul Milgrom, Muriel Niederle, Peter Norman, Alvin Roth, Michael Schwartz, and other participants of 2003 SITE Workshop on Matching Markets for comments. We use zTree software (Fischbacher, 1999) in our experiments. All errors are our own responsibility. Ünver gratefully acknowledges support from the NSF.

†Address: University of Texas at Dallas, School of Management, e-mail: eharuvy@utdallas.edu, phone: 1-(972)-883-4865.

‡Address: Koç University, Department of Economics, e-mail: uunver@ku.edu.tr, phone: 90-(212)-338-1354.

1 Introduction

The term "two-sided matching" refers to settings in which members of each of two disjoint sets wish to be matched to members of the other set. Since the two-sided matching framework applies in many market settings, a significant body of research has applied insights from matching theory to economic settings[1]. However, little attention has been given to repeated matching environments. In this work, we model repeated business-to-business (B2B) interactions as decentralized matching markets with sellers periodically making offers to buyers.

Typically, the matching of two firms involved in B2B transactions originates in a request for proposals (RFP) by the buyer. The RFP posts the buyer's specific requirement of a product. These are matched with sellers' profiles and products and forwarded to them. Sellers respond with offers—which may include the product specifications, quantity available, financial terms, and delivery schedule, as well as price– and the buyer chooses the best offer, often with some negotiation. Buyers and sellers are matched repeatedly. Orders are typically large and in many cases require significant resources on the part of both buyers and sellers. Most importantly, many of these contracts require substantial collaboration, knowledge sharing, and compatibility between the two transacting firms. Therefore, many B2B exchanges are "relational" and using the two-sided matching framework is more appropriate in such settings than using a standard or a non-standard auction framework.

Characterizing equilibria in buyer-seller markets is very important. Buyers and sellers as well are not ambivalent to the equilibrium chosen. We consider the deferred acceptance algorithms (Gale and Shapley, 1962). The seller-proposing (buyer-proposing) deferred acceptance algorithm is a "direct revelation mechanism." Buyers and sellers reveal their preferences over each other, and the algorithm is executed to find a matching. It is an iterative algorithm and works as follows: In the first step, each seller (buyer) simultaneously makes an offer to his most favorite buyer (seller). Each buyer (seller) who has received at least one offer rejects all but the best offer and conditionally accepts the best offer. This proceeds in an iterative process, where in each step a seller (buyer) whose offer is not conditionally accepted makes an offer to his most preferred buyer (seller) that he has not proposed to yet. Each buyer (seller) rejects all but the best offer among the ones he received in this step and the one conditionally held from past steps (if there is one). That best offer is conditionally accepted until a better offer arrives. The seller-proposing (buyer-proposing) algorithm terminates when no offers are rejected in a step. The conditionally accepted offers by buyers (sellers) then become permanently accepted and are realized as matches.

The importance of this algorithm lies in the fact that it always finds a "stable" matching which is "optimal" for the proposing side. A matching is said to be "stable" if there is no buyer-seller pair, each of whom prefers the other to his existing partner under this matching and there is no agent who prefers being unmatched to his partner under this matching (Gale and Shapley, 1962). The stable matchings

[1] See Roth and Sotomayor, 1990, for an extensive survey of two-sided matching models and their applications in entry-level labor markets. There is also a growing literature on matching market experiments. For example, Kagel and Roth (2000) and Ünver (2000) studied the properties of centralized matching mechanisms, which are used to admit applicants in different entry-level markets. Experimental studies by Haruvy, Roth and Ünver (2001), McKinney, Niederle and Roth (2002), Niederle and Roth (2003), and Niederle, Roth and Ünver (2004) study market characteristics and aspects of demand and supply in decentralized or centralized labor markets.

can be ordered according to the sellers' preferences or according to the buyers' preferences. There is an optimal stable matching for all sellers, and this is the worse stable matching for all buyers. Similarly, the best stable matching for all buyers is the worst stable matching for all sellers. The seller-proposing deferred acceptance algorithm finds the seller-optimal stable matching, whereas the buyer-proposing deferred acceptance algorithm finds the buyer-optimal stable matching.

In this work, we propose that the buyer-proposing algorithm is a strong predictor in buyer-proposing matching markets, such as forward B2B auctions, and the seller-proposing algorithm is a strong predictor in seller-proposing matching markets, such as reverse B2B auctions.

Standard deductive equilibrium selection principles such as payoff dominance and risk dominance are not very meaningful in these settings. This leaves us with inductive selection principles—namely, learning or evolutionary dynamic models. We pursue a dynamic investigation of data in experimental repeated matching markets and find that leading models have good predictive power in such markets. Nevertheless, we caution that the parameterization of dynamic models is still prohibitive in their application to broad settings. That is, it is not trivial to generalize parameters from one setting to another.

The questions we investigate in the experimental B2B matching markets are (1) whether the observed outcomes are likely to be stable à la Gale and Shapley (1962), and if they are, (2) whether in a market where only one side (in our case, the seller) proposes bids and the other side (the buyer) accepts or rejects, the outcome is necessarily the optimal stable matching for the proposing party, (3) whether incentives alter the likelihood of reaching a stable outcome, as well as the speed of convergence to that outcome, (4) whether knowledge of others' preferences is helpful in reaching the stable outcome, and (5) whether the receiving party (the buyer) is able to behave strategically with and without information.

To examine predictions empirically, we use experimental methods. The experimental treatments we study involve four sellers and four buyers. Each seller makes a single offer per period and each buyer, when not automated, can accept a single offer in each period (multi-period one-sided proposals by sellers with incomplete information to buyers were studied by Rapoport, Erev, and Zwick, 1995, in a non-matching context). These matching decisions are repeated many times, which allows for convergence.

The experimental conditions vary along two dimensions, resulting in a 2x2 design. One dimension of manipulation is the strategic behavior of buyers. This is done by automating buyers in one set of conditions and allowing them be strategic in another set of conditions. The other dimension of manipulation is the level of information. In the low information conditions, no subject has information about the valuations of others or others' offers. In the high information conditions, all subjects know all pertinent information about valuations, offers and their outcomes in the market. With this 2x2 design we aim to separately identify the effects of information and strategic behavior by buyers.

We find that in the incomplete information treatment with automated buyers, the seller-optimal outcome is predicted by dynamic learning and is almost always reached. That means that buyers would get their worst stable matching in the seller-proposing market as opposed to a buyer-proposing market. In the high information and strategic buyers conditions, repeated interactions occasionally culminate in buyer-optimal stable matchings, but the seller-optimal matching is nevertheless reached in the vast majority of instances. A closer inspection of these results reveals that the few instances of buyer-optimal outcomes are likely due to additional information rather than to strategic behavior by buyers.

2 Experimental Design

We begin with the commonalities to the experimental conditions. In each condition, subjects make decisions in 150 matching markets. Sellers' and buyers' valuations over each other remain fixed for the first 30 markets, a second set of valuations applies to the next 60 markets, and finally a last set of valuations applies in the last 60 markets. We refer to the three valuation profiles as profiles 1, 2, and 3, respectively (see Table 1). The three profiles were chosen such that the first profile converges in the seller-proposing Gale-Shapley algorithm after five iterations, the second profile converges in six to seven iterations, and the last profile converges in eight iterations. That is, convergence becomes increasingly difficult as the experiment proceeds. Each profile is characterized by two stable matching outcomes, as shown in Table 2.

The payoff to a seller from matching to a buyer can be 1, 2, 3, or 4 tokens. Typically, the preferences are strict. [2] Staying unmatched is costly and results in the loss of 1 token for each unmatched party. Each subject begins with a 10 token initial endowment to prevent bankruptcies early on.

In each session, three preference profiles are used. The first profile is used in the first 30 markets, the second profile in the next 60 markets, and the last profile in the last 60 markets.

As discussed in the introduction, four experimental conditions are studied. The first two experimental conditions are low (incomplete) information conditions and the remaining two conditions are high (complete) information conditions. Within each information structure, one condition has automated buyers and one condition has active human buyers.

In each market there are two stages. In the first stage, each seller can extend an offer to one and only one buyer. After offers are simultaneously submitted by sellers in the first stage of the market, buyers simultaneously decide which offer to accept in the second stage of the market. Each buyer can accept only one offer in a market. When buyers are automated, each buyer accepts the best incoming offer. Each subject then observes his payoff for that market. We summarize the treatments as follows:

1. **Incomplete information matching market with strategic buyers (Condition 1):**

   Sellers and buyers know only their own value profiles. Buyers know only their own received offers in each market. Subjects do not know other subjects' value profiles or other subjects' offers, and which of these offers were accepted in previous markets.

1A. **Incomplete information matching market with automated buyers (Condition 1A):**

   Sellers know only their own value profiles. Subjects do not know other subjects' value profiles or other subjects' offers, and which of these offers were accepted in previous markets. Buyers are automated robots that accept best incoming offers.

2. **Complete information matching market with strategic sellers (Condition 2):**

   At the beginning of each market, each seller and each buyer has complete information about the preference profiles of all others. Moreover, in the second stage, each buyer observes which sellers

[2] In Profile 2, one seller is indifferent between two buyers. This was designed to allow for six to seven iterations before convergence in the seller-proposing deferred acceptance algorithm.

Profile 1				
Seller Payoffs	**Buyer 1**	**Buyer 2**	**Buyer 3**	**Buyer 4**
Seller 1	3	1	4	2
Seller 2	4	1	3	2
Seller 3	4	3	2	1
Seller 4	1	4	3	2
Buyer Payoffs	**Seller 1**	**Seller 2**	**Seller 3**	**Seller 4**
Buyer 1	2	4	3	1
Buyer 2	4	1	2	3
Buyer 3	2	1	4	3
Buyer 4	3	1	2	4
Profile 2				
Seller Payoffs	**Buyer 1**	**Buyer 2**	**Buyer 3**	**Buyer 4**
Seller 1	2	4	3	1
Seller 2	4	3	1	3
Seller 3	3	1	2	4
Seller 4	4	2	1	3
Buyer Payoffs	**Seller 1**	**Seller 2**	**Seller 3**	**Seller 4**
Buyer 1	4	1	3	2
Buyer 2	3	1	2	4
Buyer 3	1	4	2	3
Buyer 4	4	3	2	1
Profile 3				
Seller Payoffs	**Buyer 1**	**Buyer 2**	**Buyer 3**	**Buyer 4**
Seller 1	1	3	4	2
Seller 2	1	3	2	4
Seller 3	1	4	2	3
Seller 4	2	4	1	3
Buyer Payoffs	**Seller 1**	**Seller 2**	**Seller 3**	**Seller 4**
Buyer 1	1	2	4	3
Buyer 2	3	4	1	2
Buyer 3	1	4	2	3
Buyer 4	4	2	1	3

Table 1: Value profiles used in the experiment

Stable Matchings	Seller-Optimal	Buyer-Optimal
Profile 1	$(4, 1, 3, 2)$	$(2, 1, 3, 4)$
Profile 2	$(3, 4, 1, 2)$	$(1, 4, 3, 2)$
Profile 3	$(4, 2, 3, 1)$	$(4, 2, 1, 3)$

Key: Matching (W,X,Y,Z)- The numbers in the parentheses are the buyers matched to sellers 1, 2, 3 and 4, respectively in the depicted matching.

Table 2: Stable matchings of the matching markets

made offers to which buyers in the first stage. At the end of each market, every subject observes who made offers to whom and which of these offers were accepted.

2A. **Complete information market with automated buyers (Condition 2A):**

At the beginning of each market, each seller has complete information about the preferences of all others. At the end of each market, every subject observes who made offers to whom and which of these offers were accepted. Buyers are automated robots that accept best incoming offers.

There are nine cohorts of Condition 1, ten cohorts of Condition 1A, eight cohorts of Condition 2, and ten cohorts of Condition 2A using the three value profiles.[3]

The subjects were recruited from the Harvard Business School experimental research subject pool. The subject pool consists of mostly undergraduate students from Harvard, Boston University, MIT, and surrounding schools. All the experimental sessions were conducted at the Harvard Business School Computer Lab for Experimental Research. Subjects earned a participation fee and their token earnings were converted into $US. The exchange rate was 30 tokens per $1.

3 Theoretical Analysis

In this section, we derive the equilibrium predictions in each of the game proposed settings.

3.1 Theoretical Analysis for Condition 1 and Condition 1A

Although we do not impose a belief structure, it is reasonable to assume that due to insufficient information each agent believes that all possible value profiles are equally likely to occur. We call this belief structure a **uniform belief structure**.

3.1.1 Stage-Game:

The equilibrium analysis for the one-shot game solution will form the bases for our repeated game analysis. We define a "neutral strategy" as a strategy which prescribes an action to the seller (buyer) based only on the values of buyers (sellers). Neutral strategies are defined by "making an offer to k^{th} choice" for sellers. That is, they are independent of the names of the buyers. Similarly for the buyers, when they are not automated, neutral strategies are defined in terms of the ranking of the offering agents, i.e. a neutral buyer strategy consists of following type of actions in each information set: "accept an offer from $\ell(I)^{th}$ choice in information set I." One example for a neutral buyer strategy is: "accept the best offer in every information set if it is at least as good as the ℓ^{th} choice." We assume that the space of strategies consist of neutral strategies for sellers and buyers in our theoretical analysis.

It is straightforward to see that the unique neutral Bayesian Nash equilibrium is each seller making an offer to his highest ranked choice under the uniform belief structure:

[3] Instructions are given online at http://www.utdallas.edu/~eharuvy/B2B

Proposition 1: In Condition 1A– the incomplete information condition with automated buyers– under the uniform belief structure and uniform distribution over possible profiles, a strategy profile where each seller makes an offer to his highest ranked buyer is the unique neutral Bayesian Nash equilibrium.

Proof: Suppose that all sellers believe that all profiles are equally likely to occur. Consider a seller s. Given that every other seller t is offering to his k_t^{th} choice, every offer of seller s is equally likely to be accepted by the buyers, who are automated robots accepting best incoming offers. It is best response for seller s to make an offer to his first choice. Therefore, the unique neutral Bayesian Nash equilibrium is every seller making an offer to his top buyer. ♦

Next we analyze the stage game of Condition 1, where buyers are strategic players.

Proposition 2: In Condition 1– the incomplete information condition with strategic buyers– under the uniform belief structure and uniform distribution over possible profiles, a strategy profile where each seller makes an offer to his highest ranked buyer and each buyer accepts the best offer in each information set is the unique undominated neutral Bayesian Nash equilibrium.

Proof: First consider the information sets of the buyers. The dominant strategy for each buyer is accepting the best coming offer. Therefore, under undominated Bayesian Nash equilibrium, buyers behave as automated buyers. The remainder of the proof follows from Proposition 1. ♦

3.1.2 Repeated-Game:

Note that agents play each profile repeatedly. We will refer to each stage-game in the repeated-game block as a period.

Consider the following repeated game strategy for seller s:

- In the first period, a seller s initially makes an offer to his highest ranked buyer.
- In the k^{th} period,
  - if s did not get matched in the previous period, he makes an offer to his highest ranked buyer that he has not yet made an offer to;
  - if s was matched in the previous period, he makes an offer to the choice that he was matched in $(k-1)^{th}$ period.

We name this repeated game strategy as the "going-down-the-list" strategy. Consider any preference profile where the deferred acceptance algorithm converges to the seller-optimal stable outcome in S steps. Let the game be played $K \geq S$ times using the same preference profile.

We first claim that this strategy profile will result in the seller-optimal stable matching in all of the last $K - S + 1$ periods for Condition 1A:

Lemma 1: In the last $K - S + 1$ periods of a finitely repeated game in Condition 1A, the outcome of the "going-down-the-list" strategy profile coincides with the outcome of the seller-proposing deferred acceptance algorithm (i.e., seller-optimal stable matching).

Proof: In the first period of the repeated game, each seller proposes to his first choice. The stage-game best-response is to accept the best offer and reject all others. This is equivalent to the first step of the seller-proposing deferred acceptance algorithm. In the second period, all sellers except the ones matched in the first period make offers to their second ranked choices. The matched sellers make offers to their partner in the first period again. The best-responding buyers reject all but the best offers. This period's outcome is equivalent to the outcome of the second step of the seller-proposing deferred acceptance algorithm. Similarly, subsequent periods will have equivalences in the seller-proposing deferred acceptance algorithm. When we iteratively continue in this manner, each period of the repeated game block is equivalent to a step of the seller-proposing deferred acceptance algorithm. Hence, the outcome of the dynamic will be identical to the outcome of the seller-proposing deferred acceptance algorithm, which is known to be the seller-optimal match. This will occur in the S^{th} period of the repeated game block. Since no offers are rejected in S^{th} period, no seller will go down his list in stage game $S+1$ and in later periods and each seller will continue to be matched in each of the following periods to the same choice as he did in period S. ♦

In the experiment, sellers rarely use strict going-down-the-list strategies. However, they use a delayed version of the going-down-the-list strategy more frequently. A delayed going-down-the-list strategy is defined as:

- In the first period, a seller s initially makes an offer to his highest ranked buyer.
- In the k^{th} period,
  - if s did not get matched in the previous period,
    - he makes an offer to his highest ranked buyer to whom he has not yet made an offer if he got rejected by the same choice more than ℓ times,
    - he makes an offer to the same choice to whom he made an offer in $(k-1)^{th}$ period if he got rejected by that choice less than or equal to ℓ times.
  - if s was matched in the previous period, he makes an offer to the choice to whom he was matched in $(k-1)^{th}$ period.

In this strategy, ℓ is the delay and it can be different for every choice, but should be bounded above. Lemma 1 can be trivially generalized for the delayed going-down-the-list strategies.

Corollary 1. If the number of periods in a Condition 1A repeated game block is sufficiently large then the outcome of delayed going-down-the-list strategies will converge to the seller-optimal stable matching.

Proof: Similar to the proof of Lemma 1. ♦

Next we consider Condition 1. Lemma 1 and Corollary 1 will hold if the buyers play their dominant stage game strategies in the repeated game. Since buyers have no information about other players strategies, actually it is more than natural that they accept the best offer in each stage game.

3.2 Theoretical Analysis for Condition 2 and Condition 2A

3.2.1 Stage-Game:

In Condition 2, we consider the subgame perfect equilibrium as a stage-game solution concept. In order to analyze the subgame perfect equilibrium in Condition 2, we first analyze Condition 2A in which buyers are non-strategic responders who accept the best incoming offer and sellers have all information about others' preferences. Sellers make simultaneous offers to buyers.

Lemma 2: The matching outcome of every pure strategy equilibrium of a stage game in Condition 2A must be stable à la Gale-Shapley.

Proof: Suppose there exists an equilibrium of the matching game, σ, whose outcome is unstable. Since all sellers and buyers are acceptable, the matching is necessarily blocked by a seller-buyer pair (s, b). This means that seller s (buyer b) prefers buyer b (seller s) to his match. Then seller s would be better off by making offer to b instead of his current match, since buyer b is an automatic best-response player. This contradicts to σ being an equilibrium. An unstable matching cannot be sustained by an equilibrium strategy. ♦

We use this result to prove our result for Condition 2. Recall that in this condition there are fully informed buyers and sellers. Moreover buyers can observe to which buyers the sellers have made offers. We will refer to this condition as Condition 2.

Lemma 3: The outcome of every pure strategy subgame perfect equilibrium of a stage game in Condition 2 must be stable à la Gale-Shapley.

Proof: Consider the subgame where buyers accept incoming offers. Any equilibrium of this subgame should involve buyers accepting the best incoming offers. By backward induction, buyers act as automated buyers who accept best incoming offers under every subgame perfect equilibrium. These together with Lemma 2 imply that every subgame perfect equilibrium should have a stable matching as its outcome.♦

Note that the seller-optimal stable matching is the payoff dominant stable matching (and payoff dominant subgame perfect equilibrium outcome) for the sellers and the buyer-optimal stable matching is the payoff dominant stable matching (and payoff dominant subgame perfect equilibrium outcome) for buyers. Moreover in the profiles we use, there are only two stable matchings. Hence, these two stable matchings cannot be ordered according to payoff dominance for the whole player set.

We can prove the following proposition about the equilibrium strategies in Condition 2A:

Proposition 3: The following strategy profiles are the only pure strategy equilibria of a stage game in Condition 2A: Let μ be a stable matching of buyers with sellers; each seller s makes an offer to $\mu(s)$.

Proof: Consider such a strategy profile defined for a stable matching μ. The outcome of this strategy profiles is μ. Under μ, there is no buyer-seller pair that can block the outcome. That is, no seller can profitably deviate by making an offer to a more preferred buyer that would accept. Therefore, this strategy profile is Nash equilibrium. Moreover, there is no other strategy profile that results with the outcome μ. Since μ is an arbitrary stable matching, by Lemma 2 there can be no other pure strategy Nash equilibria than the strategy profiles defined in the Proposition. ♦

We can prove the following proposition about the subgame perfect equilibrium strategies in Condition 2:

Proposition 4: The following strategy profiles are the only pure strategy subgame perfect equilibria of a stage game in Condition 2: Let μ be a stable matching of buyers with sellers; each seller s makes an offer to $\mu(s)$; each buyer accepts the best offer incoming in every information set in which he receives offers.

Proof: Consider such a strategy profile defined for a stable matching μ. The outcome of this strategy profile is μ. Under μ, there is no buyer-seller pair that can block the outcome. That is, no seller can profitably deviate by making an offer to a more preferred buyer that would accept. Each buyer accepts the best offer in each information set. Therefore, this strategy profile is a SPE. In the second part, the unique Nash equilibrium is every buyer accepting the best offer in any information set. When buyers accept the best incoming offer, the unique strategy profile for sellers that has the outcome μ is every seller s making an offer to $\mu(s)$. Since μ is an arbitrary stable matching, by Lemma 3 there can be no other pure strategy subgame perfect Nash equilibria than the ones defined in the proposition. ♦

3.2.2 Repeated Game and Learning of a Stage-Game-Equilibrium:

As we will mention later in the experimental results section – subjects do not necessarily use a repeated-game subgame perfect equilibrium strategy. Rather they learn to play one of the two subgame perfect equilibria of the stage-game over time.

History dependent "going-down-the-list" repeated-game strategy is an ad-hoc heuristic for sellers to use as a learning device even under complete information. From our analysis of Condition 1 and Condition 1A, we know that this strategy's outcome for each period will converge to the seller optimal stable matching if buyers always accept only the best-incoming offer. That is the subgame perfect equilibrium strategy of the stage-game for each buyer.

We know that from a buyer's perspective, the seller-optimal stable matching is dominated by the buyer-optimal stable matching. Moreover, any buyer who has a strict preference for the buyer-optimal stable matching (some are indifferent) can unilaterally use a different strategy against "going-down-the-list" sellers to obtain a convergence to a matching in which he gets matched to his buyer-optimal stable partner. We will prove this result as follows:

Let there be K periods in a repeated-game block. We will define a plausible strategy for a buyer. Let μ_B be the buyer-optimal stable matching and μ_S be the seller-optimal stable matching. We define a stage game strategy for a buyer b as an "achievable threshold strategy" if in every information set the buyer accepts the best offer unless it is coming from a strictly worse seller than a threshold seller s (say, k^{th} choice of buyer b) where s is not better than $\mu_B(b)$, the buyer-optimal stable matching partner, for buyer b. For example, "accepting the best offer in every information set" is an achievable strategy for any buyer with the threshold set as the least favorite seller. We also extend this strategy to a repeated-game strategy, if the buyer uses the same achievable threshold stage-game strategy in every period. Consider the following repeated-game strategy profiles:

- Each seller plays a delayed-going-down-the-list strategy, and each buyer plays the a repeated achievable threshold stage game strategy. We will refer to this repeated strategy profile as σ. Let ν_B (ν_S) be the buyer-(seller-) optimal stable matching of the matching market induced by buyers truncating their preferences at their thresholds under σ and sellers having their true preferences. We refer to this matching market as the induced market of σ.

- Let buyer b be such that he is strictly worse off under ν_S with respect to ν_B. Each seller plays a delayed-going-down-the-list strategy, and each buyer, but a buyer b, plays a repeated achievable threshold strategy, and buyer b plays the repeated achievable threshold strategy with threshold set at $\nu_B(b)$. We refer to this repeated-game strategy profile as σ^D. Similarly, we define the induced market of σ^D.

Note that σ^D is attained by deviation of buyer b from σ as noted above. Our next result states which strategy-profile is more desirable for buyer b between these two:

Proposition 5: In a repeated game in Condition 2 with sufficiently large K, strategy profile σ^D is going to converge to a matching in which buyer b is matched with $\nu_B(b)$, on the other hand strategy profile σ will converge to ν_S. Therefore, with sufficiently large K, buyer b will be better off under σ^D.

Proof: Let $s_1 = \nu_S(b)$ be the seller-optimal stable match partner of buyer b under the induced market of σ and let $s_2 = \nu_B(b)$ be the buyer-optimal stable match partner of buyer b under the induced market of σ. Profile σ^D will induce a matching market where all but one agent (that is buyer b) has the same value profile as the original market of σ. Note that σ will converge to ν_S by Corollary 1 and σ^D will converge to the seller-optimal stable matching of the induced market of σ^D by Corollary 1. Matching ν_B is stable in the induced market of σ^D. Since b is matched to s_2 in this stable matching, he will never be unmatched in a stable matching of the modified market of σ^D. Moreover the seller-optimal matching of the modified market is stable in the original market, since the preferences of the agents up to their seller-optimal stable partners stay intact between the two profiles. These imply that the seller-optimal match of b in the modified market is equal to s_2. ♦

Therefore with delayed-going-down-the-list sellers, as long as all other buyers play threshold strategies in every period (including the non-strategic "accepting the best offer in each information set in every period" repeated game strategy), a buyer will have an incentive to deviate from this type of a strategy if he is strictly worse off under the seller-optimal stable matching. Although these strategies seem plausible learning tools, they are ad-hoc and do not truly convey what the sellers and buyers are doing in the experiment. There are some cohorts of data in which this type of behavior is exhibited and there are many more cohorts of data where the seller-optimal stable matching is obtained as the outcome. We also observe cohorts with no apparent convergence.

We propose learning models to see how the population's average agents learn. In the next section, we delineate our methodology for learning simulations.

4 How do the Subjects Learn?

We model learning in simple strategies. This learning method captures human learning quite well. We use two alternative adaptive models to model simple strategy learning: a reinforcement learning model and a coevolutionary genetic algorithm. Note that neither model takes into consideration the "information" structure of the games. Therefore, we have a unique prediction of each learning model for both information structures. That is, for Condition 1 and Condition 2, in which buyers are players like the sellers and the information structure is different between these conditions, each learning model creates only one set of predictions. Similarly for Condition 1A and Condition 2A, in which buyers are automated and the sellers are the only players, each learning model finds only one set of predictions again.

4.1 Reinforcement Learning

We begin with the reinforcement learning model (REL) (Erev, Bereby-Meyer, and Roth, 1999). This model was found to provide nice approximations of actual behavior in more than 80 decision tasks (including n-person strategic games, team games, 2-person games in which the players could not reciprocate, and different individual decision tasks). Our adaptation of the model, in its simplest form, can be summarized by the following assumptions:

L1: Uniform Belief Propensities: The decision maker (DM) has uniform initial beliefs. As such, initially the DM will be more likely to make an offer to the most preferred buyer and least likely to make an offer to the least preferred. If we were to use expected values for initial propensities, $q_j(1)$ would equal to the value of j in the DM's value profile multiplied by the expected probability that the offer would be accepted. That expected probability would be 0.684 if all other sellers make all offers with equal probability and all buyers accept their best offer. We interpret uniform beliefs to mean that every offer is as likely to be accepted as any other offer, but we impose no expected probabilities of acceptance. Since the initial expected probabilities are not separately identified from $N(1)$, we let $q_j(1)$ simply equal the value of j in the DM's value profile. $N(1)$ thus represents the strength of the initial propensity multiplied by the initial expected probability of acceptance.

L2: Average Updating: The propensity to play strategy j in period $t+1$ is a weighted average of the initial propensity ($q_j(1)$) and the average payoff obtained from playing j in the first t periods ($AVE_j(t)$). The weight of the initial propensity is a function of a "strength of initial propensities" parameter $N(1)$. The weight of the average past payoff is a function of the number of times strategy j was actually chosen in the past ($C_j(t)$). Specifically,

$$q_j(t+1) = q_j(1)\frac{N(1)}{C_j(t)+N(1)} + AVE_j(t)\frac{C_j(t)}{C_j(t)+N(1)}. \tag{1}$$

L3: Exponential Response Rule: The probability $p_j(t)$ that a DM selects the j^{th} pure strategy at time t is given by

$$p_j(t) = \frac{e^{\frac{q_j(t)\lambda}{S(t)}}}{\sum_{k=1}^{m} e^{\frac{q_k(t)\lambda}{S(t)}}}, \tag{2}$$

(where m is the number of pure strategies, in the current case 4), and is a parameter that determines reinforcement sensitivity, and $S(t)$ is a measure of the standard deviation of the payoffs that the DM has experienced up to time t. Thus, the probability of selecting a given strategy increases with the propensity to select it (which increases with the average payoff from past selections). The division by the standard deviation measure, $S(t)$, implies that noisy reinforcements reduce reinforcement sensitivity (and thus lead toward more uniform choice probabilities).

The standard deviation, $S(t)$, is estimated by the average absolute difference between the recent payoff (x at trial t) and the accumulated average payoff in the first t trials, $A(t)$. Following the logic of equation 2, $S(t)$ is updated as follows:

$$S(t+1) = S(t)W'(t) + |A(t) - x|\,(1 - W'(t)) \tag{2'}$$

where

$$W'(t) = \frac{t + N(1)}{t + N(1) + 1}.$$

For the initial value $S(1)$ we take the expected absolute difference between the payoff from random choice and "the expected payoff given random choice" ($S(1) = 1$ in the current setting).

The average payoff $A(t)$ is calculated in a similar manner:

$$A(t+1) = A(t)W'(t) + x(1 - W'(t)) \tag{3}$$

where $A(1)$ is the expected payoff from random choice ($A(1) = 2.5$).[4]

Altogether, the model has two initial propensity parameters (the value of $q_j(1)$ for each strategy j) and two free parameters, and $N(1)$, that represent the shape of the learning model. We find that the parameters that best fit our data are: $N(1) = 3$ and $\lambda = 2.5$.

We adopt this learning model for the treatments with strategic buyers and for the treatments with automated buyers.

4.2 Coevolutionary Learning

The second model we use to capture learning dynamics is a genetic algorithm model. A genetic algorithm is based on a coevolutionary learning environment. It regards actions as living beings which are coevolving in a closed environment (Holland, 1975). The extant literature on genetic learning of economic agents in mainly concerned with the way computational agents evolve in long-run economic environments. However, there are not many studies that focus on modeling human subject learning in experiments using genetic algorithms and other genetic learning approaches. Arifovic (1996), Ünver (2000), Haruvy, Roth, and Ünver (2001), Chen, Duffy, and Yeh (2002) are a few of these studies. We calibrate genetic algorithm parameters to capture dynamics of the experimental data.

A genetic algorithm uses computational operators inspired by evolution of biological species. Each action is represented by a binary code in a genetic algorithm. The code itself bears importance, since this is treated similar to the DNA of living beings in evolution.

[4] For seller 2 in profile 2, since his first ranked choice brings him payoff 4, his second and third ranked choices bring him payoff 3 and his lowest ranked choice brings him payoff 1, we have $A(1) = 2.75$, $S(1) = 0.875$.

We code the actions with a binary representation. In all conditions, the strategy of each seller is an integer between 1 and 4 representing the ranking of the offer's recipient (we code rank 1 with 11, rank 2 with 10, rank 3 with 01, and rank 4 with 00). In Condition 1 and Condition 2, buyers are strategic players and choose one of the four actions (1, 2, 3, or 4) represented by the minimal ranking of an acceptable seller's offer (we use the same coding as seller actions). If they receive multiple offers, which are at least as good as their strategic cut-off ranking, then they accept the best offer. Therefore, we model the buyer strategies as "threshold strategies." We will refer to each action as a *gene* in our genetic algorithm.

There are 4 agents in Condition 1A and Condition 2A (Seller 1, Seller 2, Seller 3, and Seller 4) and 8 agents in Condition 1 and Condition 2 (Seller 1, Seller 2, Seller 3, Seller 4, Buyer 1, Buyer 2, Buyer 3, and Buyer 4). Each agent has a pool of *DNA strings* of actions that evolves over time. Each DNA string consists of two genes. The first gene is the *dominant gene*, which determines how the agent will act if this string is his DNA. The action that the agent who uses this string will use is known as the *phenotype* of the string. Hence, the phenotype of a string is the action dictated in the first gene of the string. The second gene is the recessive gene, which is inherited from the parents and does not affect the current behavior of the agent. However, both dominant and recessive genes have equal chance to be dominant (and recessive) when they are transmitted to the offspring. A pool of DNA strings is called a *generation*. An example of a seller string is given below:

$$\text{DNA string: 1101} \quad \overbrace{11}^{\substack{\text{Action: Offer to } 1^{\text{st}}\text{ranked buyer} \\ \textit{dominant gene}}} \quad \overbrace{01}^{\substack{\text{Action: Offer to } 3^{\text{rd}}\text{ranked buyer} \\ \textit{recessive gene}}}$$

The dominant gene dictates making an offer to the 1^{st} rank buyer and the recessive gene dictates making an offer to the 3^{rd} rank buyer. The agent who has this DNA acts as the dominant gene dictates, hence its phenotype is 1.

At this point it will be useful to interpret what each DNA string represents in the game theoretic jargon. Although the phenotype of a seller DNA string is the same as a seller strategy in all conditions, this interpretation is not correct for a buyer DNA string in Condition 1 and Condition 2 (recall that in Condition 1A and Condition 2A buyers are not players). The phenotype of a buyer DNA string is more subtle: naively we assume that buyers only keep track of a threshold ranking to accept seller offers instead of modeling all information sets of buyers. This interpretation is consistent with the usage of other individual learning rules in the experimental economics literature (i.e. see Erev, Bereby-Meyer, and Roth, 1999). Moreover, as pointed in Proposition 5, the most beneficial strategies that buyers can use are threshold strategies in certain situations.

There are three fundamental operators, which derive new DNA strings using current generation of strings:

1. Elitist Selection: This operator works like the asexual reproduction operator, which clones the DNA string of a parent to the next generation as a child. The selection of a parent is described below.

2. Crossover: This operator works like the sexual reproduction operator, which derives two offspring DNA strings from two parent DNA strings. For each offspring DNA string, it combines one gene of the first parent DNA string with one gene of the code of the second parent DNA string. Each gene of each parent has equal probability of being chosen as the dominant (recessive) gene. The first offspring string receives its dominant gene from the first parent and recessive gene from the second parent. The second offspring string receives its dominant gene from the second parent and its recessive gene from the first parent. We give below an example of crossover where genetic material of parent 1 is printed in bold and genetic material of parent 2 is printed in normal characters:

`Parent 1`: **0110**	Phenotype: 3			`Child 1,1`: **01**10	Phenotype: 3
		↘	↗	`Child 1,2`: **01**11	Phenotype: 3
				`Child 1,3`: **10**10	Phenotype: 2
		Possible		`Child 1,4`: **10**11	Phenotype: 2
		Crossover		`Child 2,1`: 10**01**	Phenotype: 2
				`Child 2,2`: 10**10**	Phenotype: 2
		↗	↘	`Child 2,3`: 11**01**	Phenotype: 1
`Parent 2`: 1011	Phenotype: 2			`Child 2,4`: 11**10**	Phenotype: 1

In the above example, the first offspring string will be chosen among the first 4 candidates with equal probability and the second offspring string will be chosen among the last 4 candidates with equal probability. The selection of the parents are explained below.

3. Mutation: This operator works like the genetic mutation operator, which can modify the DNA of the offspring during transmission from the parents. Each binary digit in an offspring strategy is changed (from 0 to 1 or from 1 to 0) with a small probability p_m. An example is given below:

`Child`: 0011 Phenotype: 4 $\rightrightarrows$ Mutation in 1$^{\text{st}}$digit $\rightrightarrows$ `Child`: **1**011 Phenotype: 2

In a genetic algorithm, economic learning is initiated by the following evolutionary idea. Fitter individuals have a higher chance of survival to reproduce and to transmit their genetic material to the offspring generation. According the this principle, we define a selection operator for a genetic algorithm.[5]

The strings play a number of games against the strings of other agents. The average payoff each string gets in this tournament is called the *fitness* of that string.[6] **Selection operators** determine parent strings for crossover and elitist selection using the fitness scores:

i. Selection of parents for elitist selection: This operator first orders DNA strings according to their fitness scores. Then it chooses top α% of the current generation as parents for elitist selection starting from the fittest DNA strings, where α is a parameter. Hence, it generates parents of α% of the next generation.

[5] Mutation is the stochastic aspect of the learning model and it is also used in many other economic learning dynamics. Moreover, the elitist selection aspect is also common with many learning models such as reinforcement learning dynamics and replicator dynamics. On the other hand, the crossover operator is a genuinely new idea captured by a genetic algorithm, it creates new DNA strings using existing ones. Although this operator is usually coding dependent, the current usage adopted here using recessive and dominant genes makes it coding independent.

[6] In the simulations, there are 20 DNA strings in each pool. Each agent has a separate pool of DNA strings. This brings the "co-evolutionary" feature: each agent is treated like a separate species. In Condition 1, there are 80 strings in a single generation, in Condition 2, there are 160 strings in a single generation (20 per agent). Each generation of strings plays a total of 200 games (i.e., each string code plays 10 games on average). Then average payoff of each string determines its fitness. Considering a large number, such as 200 tournament games, ensures that each string code plays a number of games against different opponents and the average fitness it accumulates is a good indicator of its probability of survival as a parent in the current generation.

ii. Selection of parents for crossover: This operator first randomly chooses two parent candidates with discrete uniform density among the current generation. The fitter DNA string of the two becomes the first parent. After these two strings are returned back to the generation pool, the second parent is similarly determined. Then crossover is applied to the selected parents. This operator generates parents of $100 - \alpha$ % of the next generation.

Initial conditions: Researchers commonly use random initial propensities in genetic algorithm research. We also follow this practice. After the adoption of each profile, the initial DNA string generations are randomly determined.

The outline: In the initial generation for each preference profile, we randomly determine the initial pool of DNA strings for each agent. In each generation, the fitness scores of each DNA string is determined in the tournament. Each offspring string is determined via one of the reproduction operators after selection operations based on fitness scores. Each digit of a child DNA is mutated by probability p_m. The next generation consists of the possibly mutated offspring pools of the current generation. This procedure is repeated until the end of the intended number of generations.[7]

Each generation corresponds to a single period in the experimental session.

In the learning jargon, *crossover* is the innovation in learning using experience, *elitist selection* is the reinforcement of previously successful strategies, and *mutation* is the experimentation in learning.

The important question is whether the genetic algorithm learning will be similar to our experimental subjects' behavior. To answer these questions, we calibrate the mutation probability p_m and the elitist reproduction size α for the first profile used. We use the simulated method of moments (SMM) to minimize the squared distance between the experimental welfare of sellers and the genetic algorithm welfare of sellers (for profile 1 in Condition 1A) with respect to α and p_m. Then we use the same parameter values for the other profiles and the other condition.[8]

We determined $\alpha = 10\%$ and $p_m = 0.0075$ by the SMM optimization done on a grid of parameters. These are speed parameters in learning. The calibrated choices show how the genetic algorithms can be successful in mapping the speed of convergence of computational agents to human learning speed.

5 Experimental Results

Figures 1-3 describe the simulated and laboratory dynamics observed in Condition 1 and Condition 2, with strategic buyers. Figures 4-6 correspond to the simulated and laboratory dynamics observed in Condition 1A and Condition 2A, with automated buyers. The first two columns of each figure depicts the dynamics observed in the laboratory experiments. The next two columns display the simulated dynamics under reinforcement learning and genetic algorithms, respectively.

In Figure 1 and Figure 4, we see the average proportion of sellers who are making offers to their seller-optimal stable matching partner over time with strategic buyers and automated buyers, respectively. In

[7] See Holland (1975) and Goldberg (1989) for detailed discussion on genetic algorithms. See Dawid (1999) and Duffy (2004) for detailed discussion of genetic algorithms in economics and in experimental economics, respectively.

[8] We ran 50 parallel evolution simulations for each profile in each treatment and used the averages. A greater number will reduce the standard error of our statistics further down.

$p-$values		Seller 1	Seller 2	Seller 3	Seller 4	Buyer 1	Buyer 2	Buyer 3	Buyer 4
Profile 1	**Period 1-5**	0.81	0.064	0.16	0.61	0.083	0.78	0.0498	0.68
	Period 26-30	0.69	1	0.78	0.68	1	0.61	0.70	0.86
Profile 2	**Period 1-10**	0.66	0.0037	0.60	0.55	0.0012	0.13	0.50	0.16
	Period 51-60	0.88	0.18	0.37	0.55	0.41	0.32	0.099	1
Profile 3	**Period 1-10**	0.056	0.88	0.0068	0.072	0.45	0.017	0.192	0.024
	Period 51-60	0.72	0.18	0.82	0.095	0.47	0.62	0.78	0.68

Table 3: Block-by-block comparison of welfare of agents in Condition 1 and Condition 2 – two-tailed two-sample homoskedastic t-test p-values

$p-$values		Seller 1	Seller 2	Seller 3	Seller 4	Buyer 1	Buyer 2	Buyer 3	Buyer 4
Profile 1	**Period 1-5**	0.44	0.46	0.15	0.13	0.18	0.84	0.01	<0.01
	Period 26-30	0.015	1	0.38	0.47	0.66	0.73	0.38	0.11
Profile 2	**Period 1-10**	0.06	0.01	0.82	0.01	<0.01	0.02	<0.01	0.01
	Period 51-60	0.82	0.60	0.15	0.95	0.01	0.06	0.02	0.19
Profile 3	**Period 1-10**	0.55	0.10	<0.01	0.01	<0.01	<0.01	<0.01	<0.01
	Period 51-60	0.02	<0.01	<0.01	0.07	0.04	0.24	0.06	0.04

Table 4: Block-by-block comparison of welfare of agents in Condition 1A and Condition 2A – two-tailed two-sample homoskedastic t-test p-values

Figure 2 and Figure 5 we see the average welfare of the sellers over time with and without strategic buyers respectively. In Figure 3 and Figure 6 we see the average welfare of the buyers over time with and without strategic buyers respectively. In what follows we analyze the effect of information and the effect of strategic buyers.

5.0.1 The Effect of Information

We compare Condition 1 with Condition 2 and Condition 1A with Condition 2A, to see the effect of having information on the outcomes.

Table 3 and Table 4 present the p-values of block-by-block t-test comparison of Condition 1 with Condition 2 and Condition 1A with Condition 2A, respectively, for t-test comparison of the welfare of sellers and welfare of buyers at the beginning of the repeated game and at the end of the repeated game. Though it appears that occasionally at the beginning information matters, and agents do obtain different welfare levels, at the end of the conditions with strategic buyers (in Condition 1 and Condition 2). Welfare levels in the two information conditions converge and are, for the most part, no longer significantly different. The speed of convergence, as shown by the figures, seems to be the main difference between the treatments. In conditions without strategic buyers (in Condition 1A and Condition 2A), information makes a statistically significant difference at the end, especially in profile 3. However, the greater differences are still early on, indicating convergence.

In the majority of the sessions we observe that the seller-optimal stable matching emerges as an outcome in the latter markets of each profile. In Table 5, we give the convergence results. We say that a cohort's outcome **converges strongly** to a matching if in at least five periods out of the last ten periods of the cohort, that matching is observed as the outcome and no other matching has been observed five times out of the last ten periods. We say that a cohort's outcome **converges weakly** to a matching if in at least three/four periods out of the last ten periods of the cohort, that matching is observed and the majority of the remaining periods out of the last ten have matchings that differ with only one match from the limit matching.[9] In Condition 1 and Condition 2 we observe occasional convergence to the buyer-optimal stable outcome. The number of convergences to the buyer optimal outcome is 4/27 in condition 1 versus 4/24 in Condition 2. There is no convergence to stable outcomes in 2/27 instances in Condition 1 and 5/24 instances in Condition 2. In all other occasions, we observe convergence to the seller-optimal stable matching. Therefore, the convergence results are very much comparable.

Condition 1A and Condition 2A also produce comparable convergence results. 20/30 instances in Condition 1A, 21/30 instances in Condition 2A we observe convergence to the seller-optimal stable matching. Only difference is about the emergence of the buyer-optimal stable matching. The buyer-optimal stable matching never emerges in Condition 1A, although it emerges in Condition 2A in 4/30 instances. Therefore, even with automated buyers, full information causes emergence of buyer-optimal stable matching more often.

5.0.2 The Effect of Strategic Buyers

We compare Condition 1 with Condition 1A and Condition 2 with Condition 2A to understand the effect of strategic buyers in our results. In comparing the effect of strategic buyers (Tables 6 and 7), most buyers and sellers in most periods are not significantly different in average welfare. That is, buyers do not appear to behave strategically in a way that alters welfare for either buyers or sellers (except seller 3 in profile 3).

Therefore, strategic buyers almost behave like non-strategic automated buyers. This can be seen in Table 8 where we give the proportion of buyers accepting the best incoming offer in Conditions 1 and 2. Indeed, buyers almost always accept the best incoming offer.

By Proposition 5, each buyer can divert his partner from the seller-optimal matching partner to the buyer-optimal matching partner under full information, if sellers play going-down-the-list strategies: A buyer can reject all offers below his buyer-optimal stable match to divert the outcome. This happens very infrequently as seen in Table 9.

However, these small differences in buyer behavior are enough to induce different payoff levels for seller 3 under profile 3 as the only apparent effect of having strategic buyers.

5.0.3 Analysis Combining Both Effects

In this subsection we look at the combined effects of information and strategic buyers. Specifically, we compare Condition 1A (automated buyers and incomplete information) to Condition 2 (strategic buyers and complete information).

Convergence is significantly faster in Condition 2 than in Condition 1A. In Table 10, we show p-values of block-by-block t-test comparisons of Condition 1A and Condition 2 for welfare levels of each agent. In

[9] Actually we need the definition of weak convergence only for Condition 1 and Condition 1A. All the convergences are strong in Condition 2 and Condition 2A.

	Condition 1			Condition 2		
Cohort	**Profile 1**	**Profile 2**	**Profile 3**	**Profile 1**	**Profile 2**	**Profile 3**
1	Seller Opt	Seller Opt	other (4,2,3,1)	Seller Opt	Seller Opt	Seller Opt
2	Seller Opt	Seller Opt	*Buyer Opt*	Seller Opt	Seller Opt	Seller Opt
3	Seller Opt	Seller Opt*	other (2,4,3,1)	No Converge	*Buyer Opt*	No Converge
4	Seller Opt	Seller Opt	Seller Opt*	Seller Opt	Seller Opt	*Buyer Opt*
5	Seller Opt	*Buyer Opt*	*Buyer Opt*	Seller Opt	No Converge	Seller Opt
6	Seller Opt	Seller Opt	Seller Opt	Seller Opt	other (2,4,1,3)	Seller Opt
7	Seller Opt	Seller Opt	Seller Opt	No Converge	Seller Opt	Seller Opt
8	*Buyer Opt*	Seller Opt	Seller Opt	Seller Opt	*Buyer Opt*	*Buyer Opt*
9	Seller Opt**	Seller Opt	Seller Opt			

	Condition 1A			Condition 2A		
Cohort	**Profile 1**	**Profile 2**	**Profile 3**	**Profile 1**	**Profile 2**	**Profile 3**
1	Seller Opt	Seller Opt	Seller Opt*	Seller Opt	*Buyer Opt*	Seller Opt
2	No Converge	Seller Opt	No Converge	No Converge	*Buyer Opt*	Seller Opt
3	No Converge	other (2,4,3,1)	Seller Opt	Seller Opt	Seller Opt	Seller Opt
4	Seller Opt	Seller Opt	Seller Opt*	Seller Opt	Seller Opt	Seller Opt
5	other (0,1,3,2)	Seller Opt	Seller Opt	Seller Opt	*Buyer Opt*	No Converge
6	No Converge	Seller Opt	other (3,2,4,1)	Seller Opt	*Buyer Opt*	Seller Opt
7	Seller Opt*	Seller Opt	Seller Opt	Seller Opt	Seller Opt	Seller Opt
8	Seller Opt*	other (0,4,1,2)	other (2,0,3,4)	other (3,4,1,2)	Seller Opt	Seller Opt
9	Seller Opt	Seller Opt	Seller Opt*	Seller Opt	other (2,4,1,3)	other (1,2,3,4)
10	other (0,1,3,2)	Seller Opt	Seller Opt	Seller Opt	Seller Opt	Seller Opt

Key:

Seller (Buyer) Opt – In the last 10 periods of the profile, at least 5 matches were the seller (buyer)-optimal stable outcome and no other matching was observed 5 times.

Seller (Buyer) Opt* – In the last 10 periods of the profile, 4 matches were the seller (buyer)-optimal stable outcome; the majority of the remainder involve one match difference from the limit matching.

Seller (Buyer) Opt** – In the last 10 periods of the profile, 3 matches were the seller (buyer)-optimal stable outcome; the majority of the remainder involve one match difference from the limit matching.

other (W,X,Y,Z)- Convergence occurred to an outcome other than a stable match. The numbers in the parentheses are the buyers matched to sellers 1, 2, 3 and 4, respectively, in the converged outcome.

No Converge – There was no convergence as defined above.

Table 5: Convergence in the experimental sessions

$p-$values		**Seller 1**	**Seller 2**	**Seller 3**	**Seller 4**	**Buyer 1**	**Buyer 2**	**Buyer 3**	**Buyer 4**
Profile 1	**Period 1-5**	0.59	0.39	0.18	0.47	0.71	0.68	0.24	0.02
	Period 26-30	0.15	0.71	0.54	0.67	0.71	0.29	0.70	0.65
Profile 2	**Period 1-10**	0.03	0.36	0.57	0.03	0.07	0.61	0.72	0.25
	Period 51-60	0.95	0.13	0.56	0.85	0.10	0.37	0.62	0.53
Profile 3	**Period 1-10**	0.66	0.14	0.57	0.19	0.06	0.54	0.13	0.94
	Period 51-60	0.82	0.15	0.07	0.42	0.10	0.58	0.52	0.21

Table 6: Block-by-block comparison of welfare of agents in Condition 1 and Condition 1A – two-tailed two-sample homoskedastic t-test p-values

$p-$values		Seller 1	Seller 2	Seller 3	Seller 4	Buyer 1	Buyer 2	Buyer 3	Buyer 4
Profile 1	**Period 1-5**	0.41	0.55	0.82	0.80	0.96	0.11	0.67	0.02
	Period 26-30	0.11	0.39	0.93	0.34	0.39	0.14	0.46	0.81
Profile 2	**Period 1-10**	0.08	0.69	0.81	0.76	0.14	0.37	<0.01	0.99
	Period 51-60	0.97	1.00	0.87	0.66	0.45	0.36	0.82	1.00
Profile 3	**Period 1-10**	0.15	0.72	0.45	0.49	<0.01	0.91	0.12	0.37
	Period 51-60	0.70	0.38	0.04	0.17	0.87	0.48	0.24	0.81

Table 7: Block-by-block comparison of welfare of agents in Condition 2 and Condition 2A – two-tailed two-sample homoskedastic t-test p-values

	Incomplete Info (Condition 1)	Complete Info (Condition 2)
Profile 1	0.99	0.98
Profile 2	0.96	0.98
Profile 3	0.95	0.98

Table 8: Proportion of buyers accepting the best incoming offer

the last few periods of each profile (in periods 26-30 for profile 1, and in periods 51-60 for profiles 2 and 3), seller welfare converges to levels that are not significantly different at the 10% level of significance between Condition 1A and Condition 2 in all profiles and for all sellers. However, buyer welfare tends to be higher in Condition 2. In 5 out of the 12 cells of Table 10, the two conditions converged to different welfare levels, statistically significant at the 10% level. Figures 3 and 6 confirm that Condition 2 indeed results in higher welfare for buyers. In the initial periods of each profile there is generally a significant difference between agents' welfare in Condition 1A and Condition 2. We see that 7 out of 12 seller cells and 6 out of 12 buyer cells have a statistically significant initial welfare difference (10% level of significance). The impact of the buyer-optimal stable matching on Condition 2's seller welfare is clearly evident in the welfare graphs. In profile 1, sellers 2 and 4 are matched to their highest ranked buyer in the seller-optimal stable matching. As such, they are the top two curves in Figure 2, corresponding to welfare in Condition 1A. However, only seller 2 gets his highest ranked buyer in the buyer-optimal stable matching for profile 1. As such, seller 2 pulls away from seller 4 in Condition 2 relative to Condition 1A, as Figures 2 and 5 demonstrate. Similarly, in profile 2, seller 2 gets his second choice under both stable matchings whereas sellers 1 and 3 get their second choice in the seller-optimal matching but their third choice in the buyer optimal matching. As such, seller 2 pulls away from sellers 1 and 3 in Condition 2 relative to Condition 1A.

	Complete Info (Condition 2)
Profile 1	0.98
Profile 2	0.99
Profile 3	0.98

Table 9: Proportion of buyers accepting the best incoming offer when the best offer is worse than the buyer-optimal stable matching partner

$p-$values		Seller 1	Seller 2	Seller 3	Seller 4	Buyer 1	Buyer 2	Buyer 3	Buyer 4
Profile 1	**Period 1-5**	0.13	0.22	0.069	0.11	0.23	0.18	0.040	0.17
	Period 26-30	0.34	0.39	0.11	0.93	0.39	0.14	0.96	0.21
Profile 2	**Period 1-10**	0.0030	0.0060	0.98	0.024	0.010	0.25	0.74	0.012
	Period 51-60	0.86	0.64	0.21	0.86	0.051	0.37	0.052	0.25
Profile 3	**Period 1-10**	0.060	0.16	0.017	0.0050	0.025	0.0010	0.28	0.0030
	Period 51-60	0.96	0.68	0.10	0.37	0.058	0.61	0.050	0.035

Table 10: Block-by-block comparison of welfare of agents in Condition 1A and Condition 2 – two-tailed two-sample homoskedastic t-test p-values

Condition 1					Condition 2				
	1^{st}offer	2^{nd}offer	3^{rd}offer	4^{th}offer		1^{st}offer	2^{nd}offer	3^{rd}offer	4^{th}offer
Profile 1	0.64	0.64	0.50	0.39	**Profile 1**	0.72	0.53	0.44	0.44
Profile 2	0.86	0.81	0.36	0.14	**Profile 2**	0.47	0.22	0.19	0.16
Profile 3	0.97	0.94	0.47	0.17	**Profile 3**	0.38	0.28	0.22	0.19
Condition 1A					**Condition 2A**				
	1^{st}offer	2^{nd}offer	3^{rd}offer	4^{th}offer		1^{st}offer	2^{nd}offer	3^{rd}offer	4^{th}offer
Profile 1	0.60	0.58	0.45	0.45	**Profile 1**	0.73	0.58	0.43	0.40
Profile 2	0.83	0.68	0.25	0.20	**Profile 2**	0.55	0.40	0.18	0.13
Profile 3	0.83	0.75	0.38	0.20	**Profile 3**	0.43	0.35	0.23	0.10

Table 11: Proportion of sellers who play consistent with a delayed going-down-the-list strategy

5.0.4 Learning Predictions and Experimental Results

The learning simulations capture learning dynamics quite well in general. What the graphs show is that (1) learning simulations – reinforcement learning and genetic algorithms– characterize the behavior in the laboratory reasonably well, [10] and (2) the seller optimal stable match is most likely to be the converged outcome of the dynamic– in simulation and in the laboratory as seen in Table 5.

Table 11 shows that a significant portion of sellers applied strategies strictly (without error or deviation) consistent with a delayed going-down-the-list strategy. If we allow for trembling hand errors, calculation errors, or random experimentation the proportion of down-the-list players will increase substantially. In this table, columns titled "Consistent with delayed down-the-list by n^{th} offer" give the number of sellers whose actions were strictly consistent with a delayed going-down-the-list strategy by the n^{th} distinct offer made in the profile. If a seller did not make n distinct offers, but he was consistent by $(n-1)^{th}$ offer with the delayed down-the list then he is still consistent by n^{th} offer with the delayed down-the-list. The majority of sellers initially offer to their first choice in both conditions. While the proportion of delayed down-the-list players increases in Condition 1 over time, it falls in Condition 2 after profile 1. Hence the lack of information forces sellers to use delayed going-down-the-list more frequently. However, agents do not purely use this strategy; rather they experiment. Hence, the ratio of sellers who behave consistently with a delayed down-the-list strategy gets smaller over time. This result suggests

[10]Initially, genetic algorithm works as a slightly worse indicator of agent learning than reinforcement learning model. That is because calibration of a genetic algorithm is very difficult due to randomness and tournament based fitness measures, especially for initial parameters.

that agents use stochastic learning dynamics instead of pure repeated-game strategies.

6 Conclusions

Two-sided buyer-seller matching environments have not hitherto been studied as repeated games. This is partly because of the added computational complexity dynamic matching requires. The current work is an attempt to fill this void in the literature in order to address a wide variety of repeated interactions, particularly repeated buyer-seller matching. Stable buyer-seller matchings evolve over time as a result of complex preference structures that depend on compatibility and complementarity over numerous dimensions (see e.g., Dwyer, Schurr, and Oh 1987; Anderson and Naurus, 1990). Given the importance of preferences and the heterogeneity in preferences that is likely to characterize B2B markets, a matching analysis is called for. Stability is clearly crucial since in the cooperative market the core is the set of stable matchings, and stable matchings characterize the outcome set of subgame perfect Nash equilibria. However, predicting which stable outcome is likely to emerge is a task for dynamic models. In this work, both learning dynamics and genetic algorithms were used to characterize the convergence probabilities. To the extent that the buyer optimal stable outcome can emerge in a seller-proposing market and the seller optimal stable outcome can emerge in a buyer-proposing market, or that non-stable outcomes can emerge, learning dynamics and genetic algorithms are suitable for characterizing a distribution over possible outcomes. The results suggest that in a seller-proposing market, the sellers will most often get their optimal stable outcome and the buyers will most often get their worst stable outcome. Likewise, in a buyer-proposing market, buyers will get their optimal stable outcome and the sellers will get their worst stable outcome.

An important question raised in this paper pertains to the effect of information on the outcome and convergence speed. First, information about others' preferences and choices may speed up convergence and improve welfare by allowing agents to avoid costly mistakes. Second, adding more information can alter the outcomes in the buyers' favor since buyers who strictly prefer the buyer-optimal outcome can unilaterally deviate, forcing convergence to the buyer-optimal outcome. We find evidence to back the first assertion. That is, in many instances, adding information made a statistically significant difference in average welfare for both buyers and sellers. However, we find less evidence for the unilateral deviation hypothesis. Comparing settings, under both information conditions, with strategic buyers as opposed to automated best-responding buyers, we find little or no significant difference in welfare of either buyers or sellers or in the number of instances of convergence to buyer-optimal outcomes, suggesting that buyers did not behave strategically. Instead, information improves welfare through more informed seller proposals.

Firms must carefully consider the matching environment they participate in and strive to be on the proposing side whenever possible. This choice is most notable in the choice between forward and reverse B2B auctions. Relative to the mainstream auction literature, where price is the key dimension, B2B auctions are in many cases auctions only by name since the initiator of the auction (which then receives proposals) is not obliged to select the best price offer but rather considers the entire array of relevant characteristic and remains engaged throughout the bidding process. In reverse auctions, buyers pre-approve the sellers who will participate in the auction and choose the (undisclosed) award criteria, which are functions of the bid price, quality, quantity, and seller characteristics and desirability. The literature has touted reverse auctions as a tremendous potential for cost saving for buyers (Jap, 2002; Tully, 2000; Cohn, 2000). The current work suggests, however, that a reverse auction of the kind typified in B2B markets is likely to result in the absolute worst possible equilibrium for buyers.

There are several directions for future research. The analysis here assumed a one-to-one matching environment, or deal exclusivity. That is, each seller could only make a single bid in a period. For settings where the orders are large relative to each seller's available (and not yet committed) capacity this is often a reasonable assumption. That is, an order would require full commitment of resources from the typical seller. If sellers are large and can accommodate more than one contractual relationship for the same product or task, the analysis would need some modification (for a one-to-many matching settings, see Roth and Sotomayor, 1990). Other extensions could extend the assumptions on the information available to participants and map the information set to the outcome set. Lastly, when several market mechanisms coexist, the choice of market mechanism becomes endogenous to the firm and will affect participation in each market and consequently the equilibrium outcome. Future research should address these issues.

References

[1] Anderson, J. C. and Naurus, J. A. (1990), "A model of distributor firm and manufacturer firm working partnerships," Journal of Marketing, 54 (1), 42-58.

[2] Arifovic, J. (1996), "The behavior of exchange rate in the genetic algorithm and experimental economics," Journal of Political Economy, 104, 510-541.

[3] Cohn, L. (2000), "B2B: The hottest net bet yet?" Business Week, January 17, 36-37.

[4] Chen, S.-H., Duffy, J., and Yeh, C.-H. (2002), "Equilibrium selection via adaptation: Using genetic programming to model learning in a coordination game," in The Electronic Journal of Evolutionary Modeling and Economic Dynamics, issue 1, article 1002.

[5] Dawid, H. (1999), Adaptive learning by genetic algorithms. Second Edition. Berlin-Heidelberg, Germany: Springer-Verlag.

[6] Duffy, J. (2004), "Agent-based models and human subject experiments," Judd, K. L., and Tesfatsion, L. (eds) Handbook of Computational Economics vol. 2. Amsterdam, The Netherlands: Elsevier, forthcoming.

[7] Dwyer, F. R., Schurr, P. H., and Oh, S. (1987), "Developing buyer-seller relations," Journal of Marketing, 51 (2), 11-28.

[8] Erev, I., Bereby-Meyer, Y, and Roth, A. E. (1999), "The effect of adding a constant to all payoffs: Experimental investigation, and a reinforcement learning model with self-adjusting speed of learning," Journal of Economic Behavior and Organization, 39, 111-128.

[9] Fischbacher, U. (1999), "Zurich toolbox for ready-made economic experiments," mimeo.

[10] Gale, D. and Shapley, L. S. (1962), “College admissions and the stability of marriage,” American Mathematical Monthly, 69, 9-15.

[11] Goldberg, D. E., (1989), Genetic algorithms in search, optimization and machine learning. Addison-Wesley Publishing Company.

[12] Haruvy, E., Roth, A. E., and Ünver, M. U. (2001), “Dynamics of law-clerk matching: An experimental and computational investigation of proposals for reform of the market,” mimeo.

[13] Holland, J. H. (1975), Adaptation in natural and artificial systems. Ann-Arbor, MI: The University of Michigan Press.

[14] Jap, S. D. (2002), “Online reverse auctions: Issues, themes, and prospects for the future,” Journal of the Academy of Marketing Science, 30, 506-525.

[15] Kagel, J. H. (1995), “Auctions: A survey of experimental research,” J. H. Kagel and A. Roth (eds.) Handbook of Experimental Economics, NJ: Princeton University Press, 501-580.

[16] Kagel, J. H. and Roth, A. E. (2000) “The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment,” Quarterly Journal of Economics, 115, 201-235.

[17] McKinney, C. N., Niederle, M., and Roth, A. E., (2002) “The collapse of a medical labor clearing-house (and why such failures are rare),” mimeo.

[18] Niederle, M. and Roth, A. E. (2003), “Market culture: How norms governing exploding offers affect market performance,” mimeo.

[19] Niederle, M., Roth, A. E., and Ünver, M. U. (2004), “Unraveling results from comparable demand and supply,” working paper in preparation.

[20] Rapoport, A., Erev, I., and Zwick, R. (1995), “An experimental study of buyer-seller negotiation with one-sided incomplete information and time discounting,” Management Science 41(3), 377-394.

[21] Roth, A. E. and Sotomayor, M. (1990), Two-sided matching: A study in game-theoretic modeling and analysis, Econometric Society Monographs No. 18, Cambridge, The United Kingdom: Cambridge University Press.

[22] Smith, N. (2002), “Dynamic pricing effects on strategic sourcing and supplier relations,” Leaders for Manufacturing Thesis, Massachusetts Institute of Technology.

[23] Tully, S. (2000), “The B2B tool that really is changing the world,” Fortune, March 20, 132-145.

[24] Ünver, M. U., (2000), “On the survival of some unstable two-sided matching mechanisms,” International Journal of Game Theory, forthcoming.

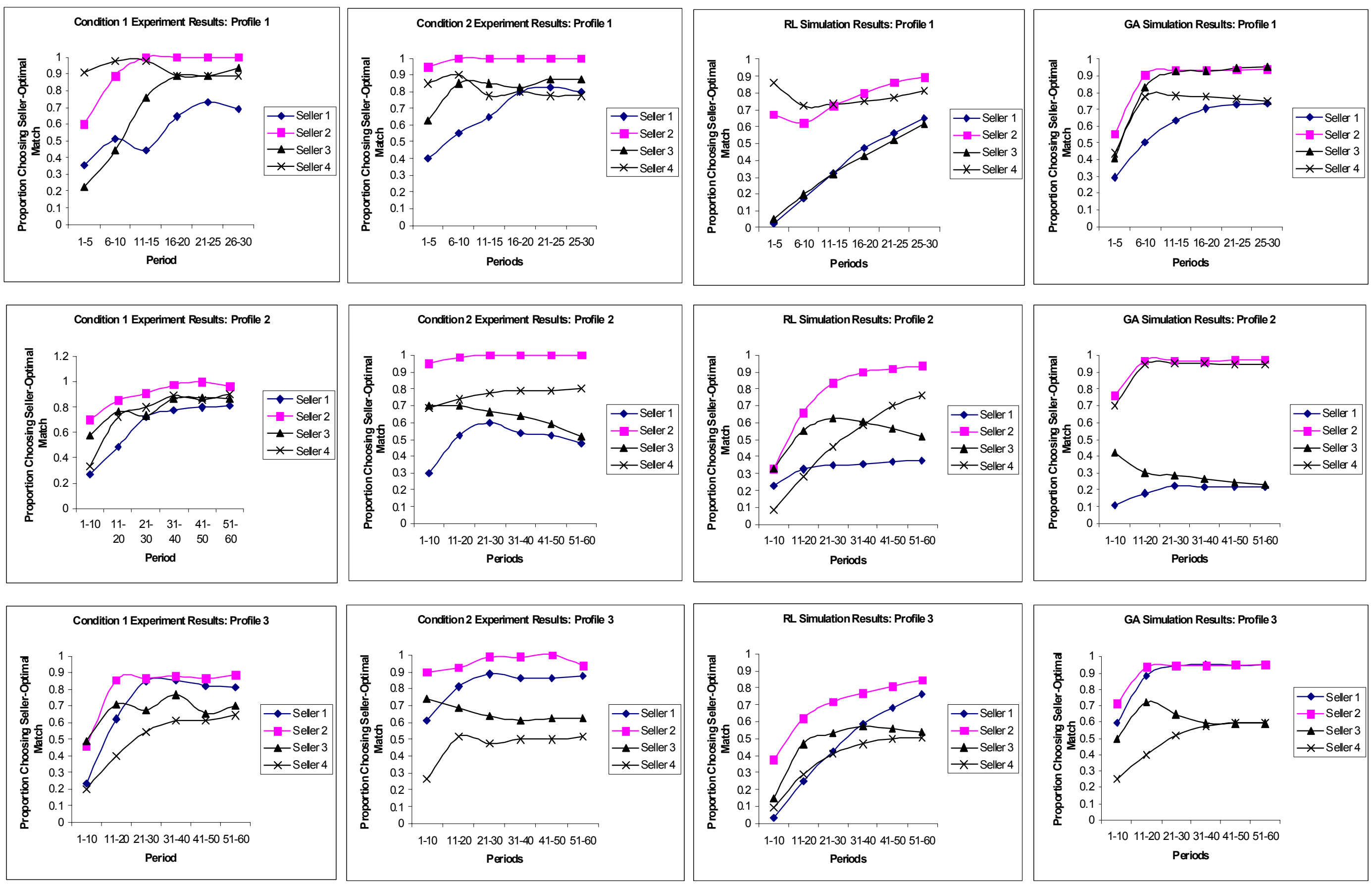

Figure 1: **Conditions with Strategic Buyers:** Proportion of Sellers Proposing to Their Seller-Optimal Stable Matching Partners

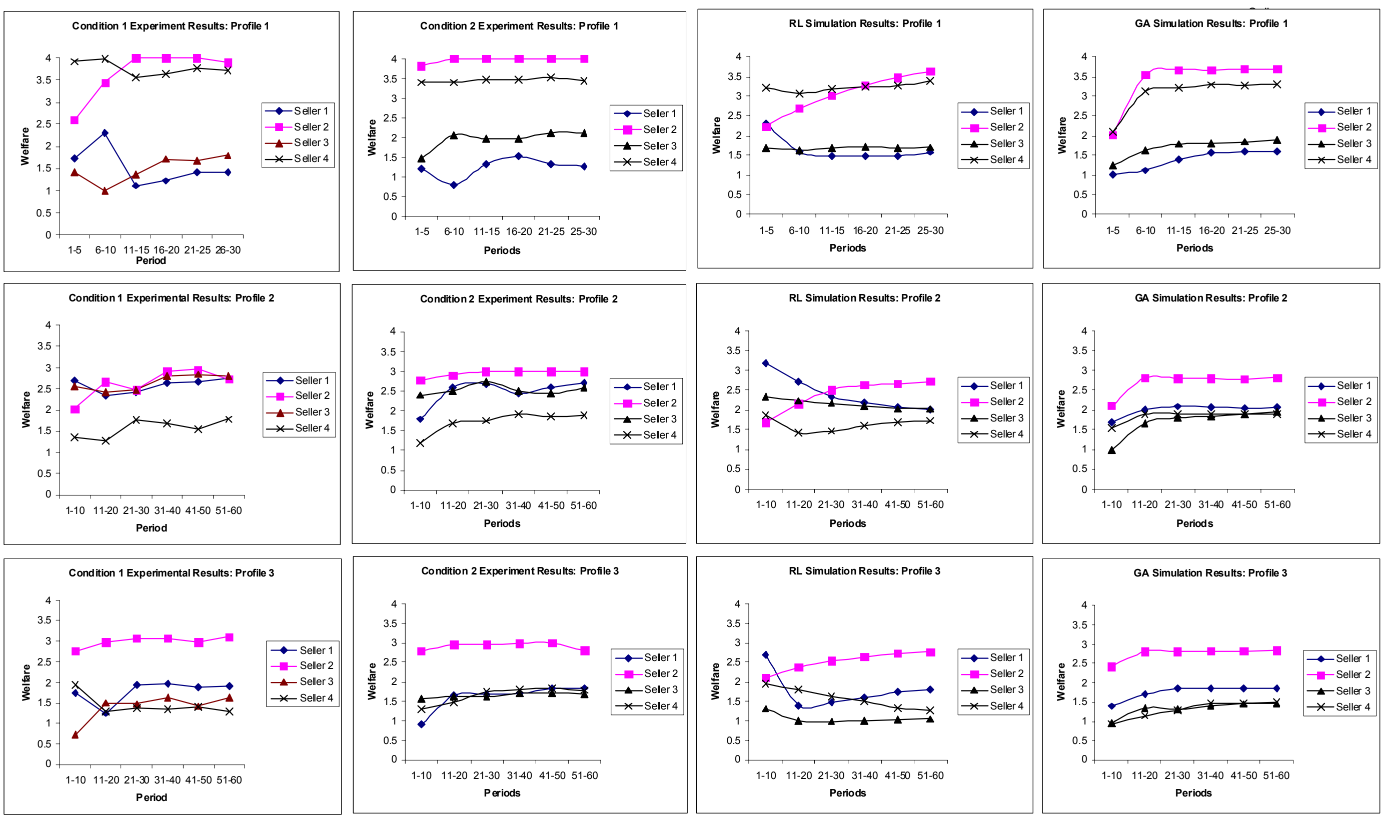

Figure 2: **Conditions with Strategic Buyers:** Welfare of Sellers

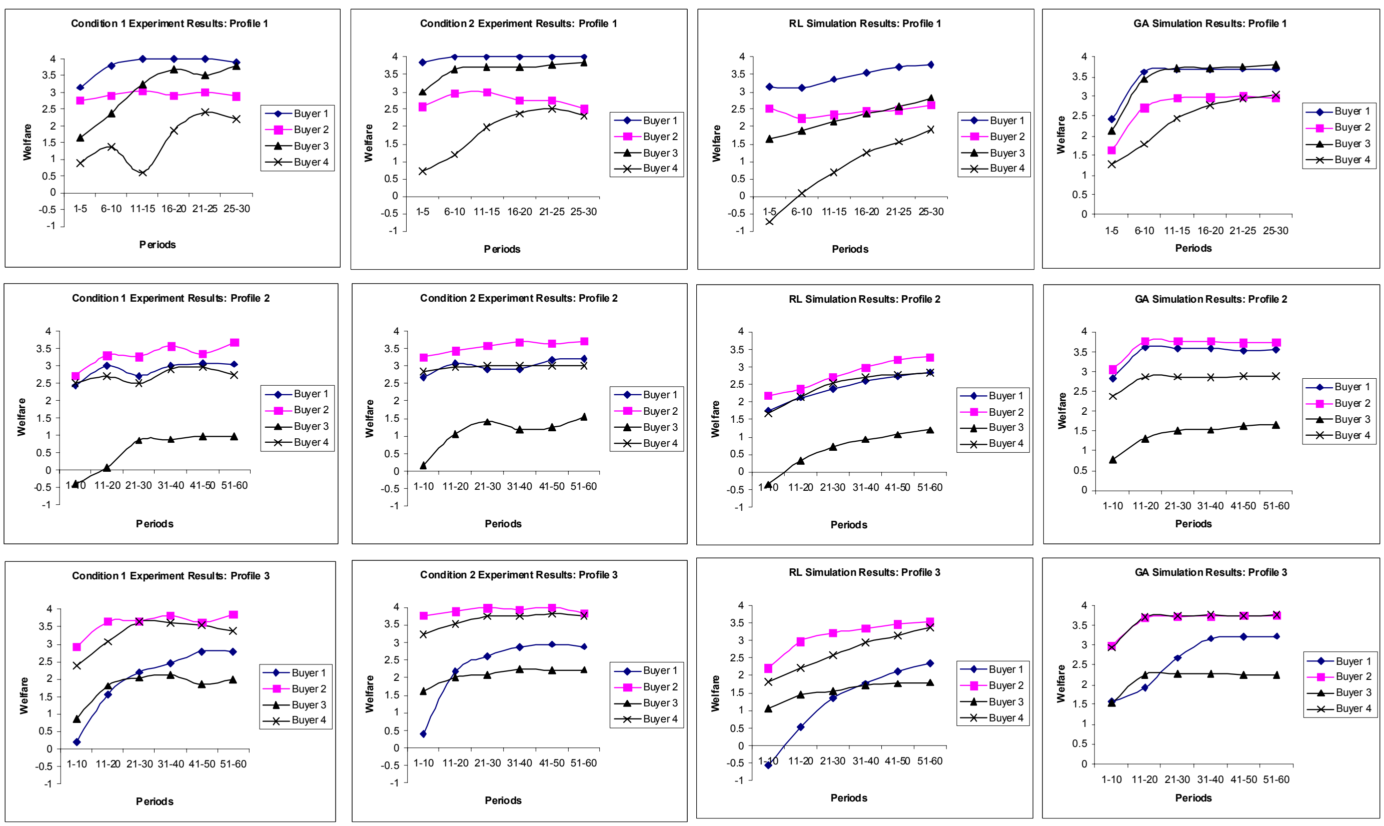

Figure 3: **Conditions with Strategic Buyers:** Welfare of Buyers

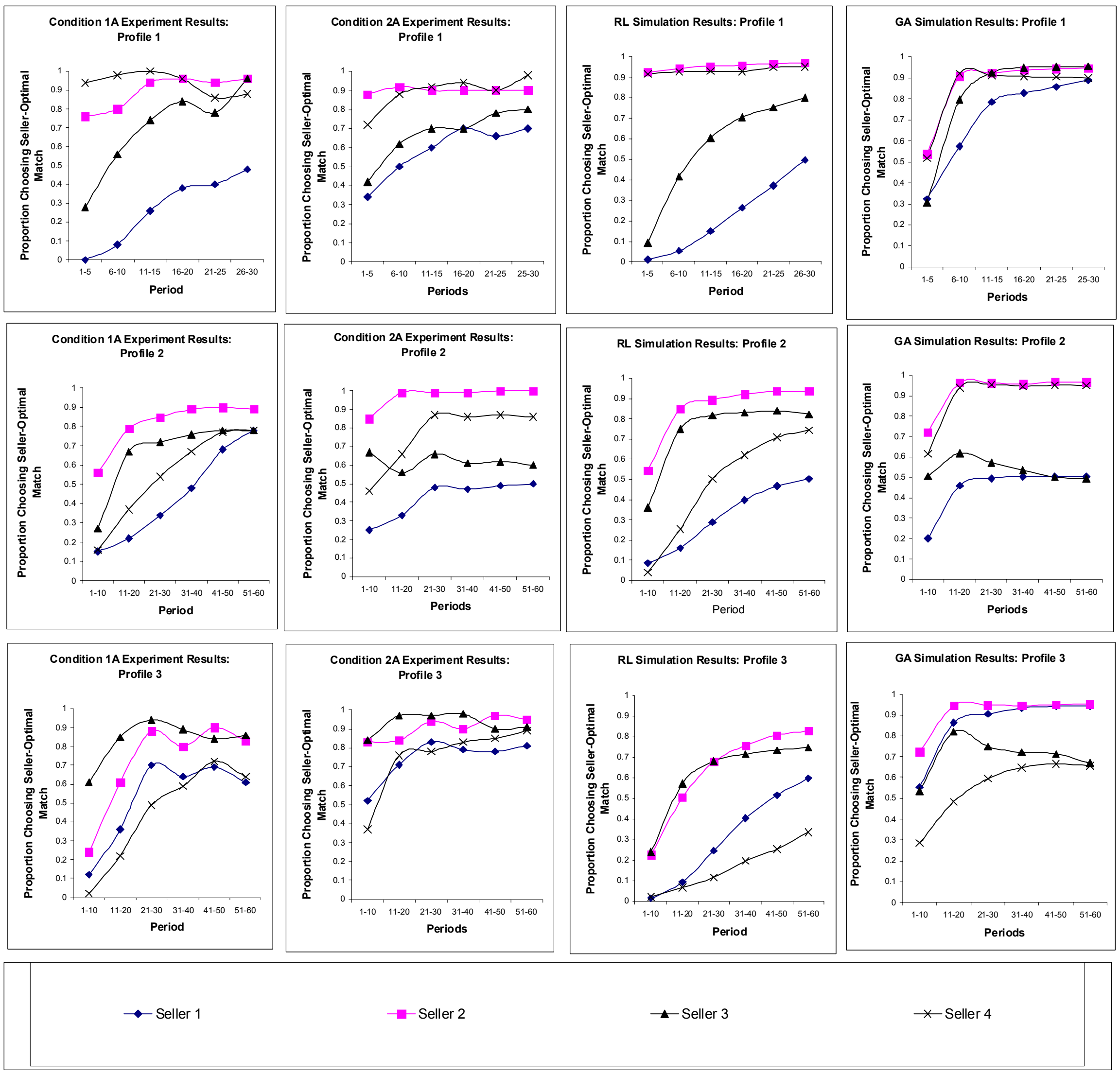

Figure 4: **Conditions with Automated Buyers:** Propostion of Sellers Proposing to Their Seller-Optimal Stable Matching Partner

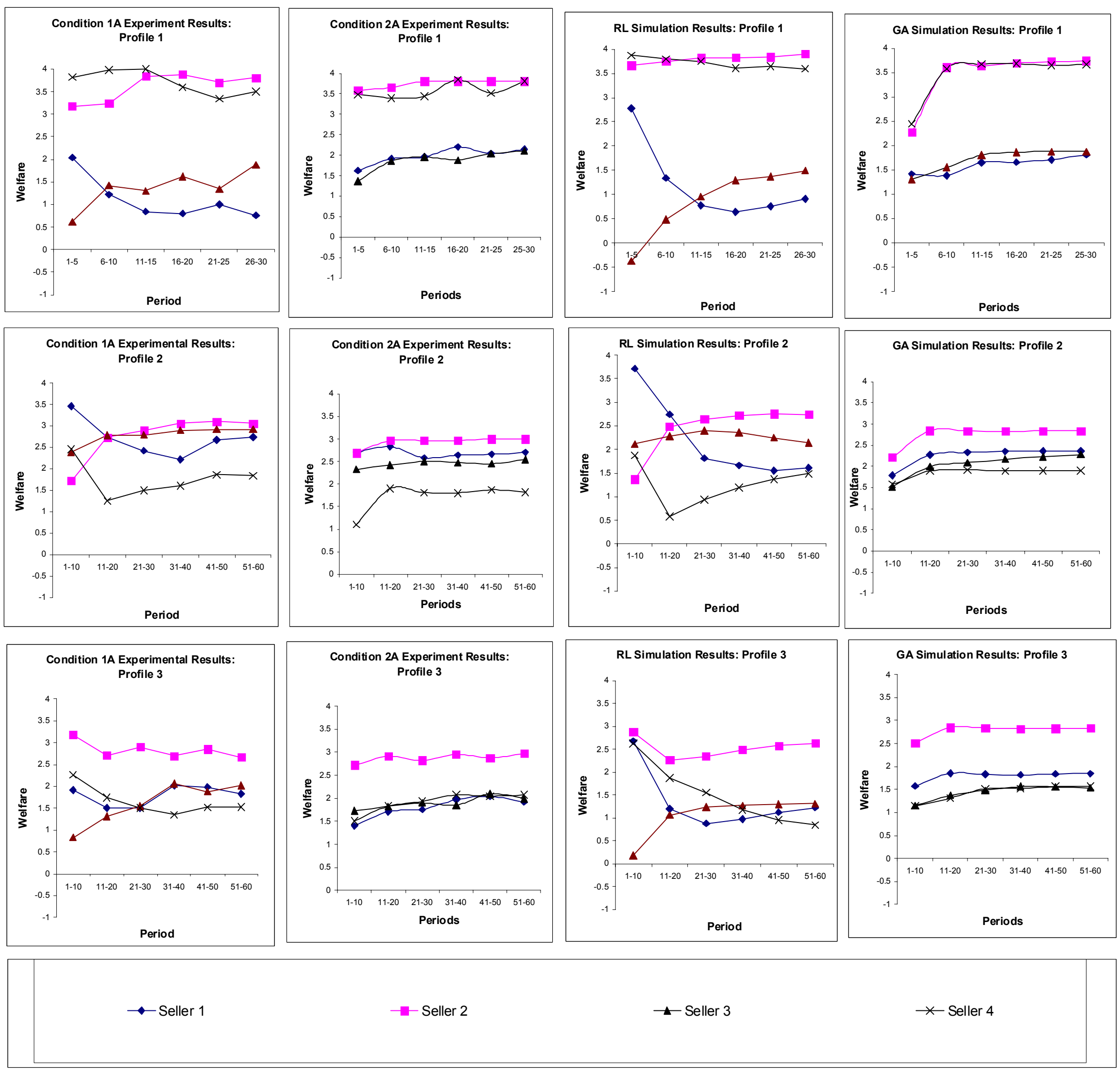

Figure 5: **Conditions with Automated Buyers:** Welfare of Sellers

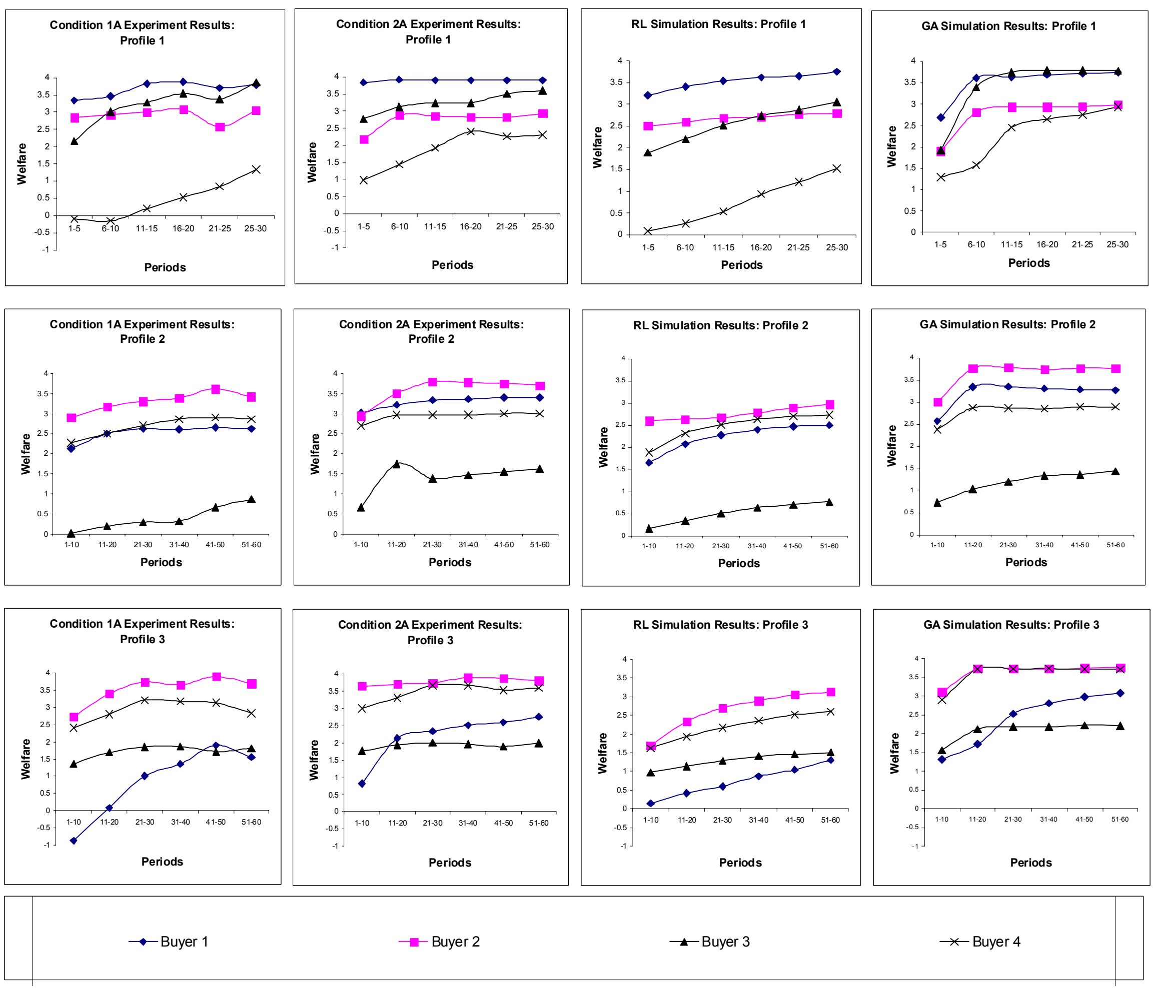

Figure 6: **Conditions with Automated Buyers:** Welfare of Buyers