Strategic Interactions in a Supply Chain Game

Michael P. Wellman Joshua Estelle Satinder Singh
Yevgeniy Vorobeychik Christopher Kiekintveld
Vishal Soni
University of Michigan
Artificial Intelligence Laboratory
Ann Arbor, MI 48109-2110 USA
{ wellman, jestelle, baveja, yvorobey, ckiekint, soniv }@umich.edu

November 10, 2004

To appear in *Computational Intelligence*

Abstract

The TAC 2003 supply-chain game presented automated trading agents with a challenging strategic problem. Embedded within a high-dimensional stochastic environment was a pivotal strategic decision about initial procurement of components. Early evidence suggested that the entrant field was headed toward a self-destructive, mutually unprofitable equilibrium. Our agent, Deep Maize, introduced a preemptive strategy designed to neutralize aggressive procurement, perturbing the field to a more profitable equilibrium. It worked. Not only did preemption improve Deep Maize's profitability, it improved profitability for the whole field. Whereas it is perhaps counterintuitive that action designed to prevent others from achieving their goals actually helps them, strategic analysis employing an empirical game-theoretic methodology verifies and provides insight about this outcome.

1 Introduction

Like classic computer games, multiagent research competitions [Stone, 2002] present well-defined problems for testing and comparing AI techniques and systems. The annual Trading Agent Competition (TAC) series provides a forum for research on strategic market behavior, and has led to several promising concepts and methods for implementing strategies in such domains [Wellman et al., 2003].

The TAC Supply Chain Management (TAC/SCM) scenario [Arunachalam and Sadeh, to appear, Sadeh et al., 2003] defines a complex six-player game with severely incomplete and imperfect information, and high-dimensional strategy spaces. Like the real supply-chain environments it is intended to model, the TAC/SCM game presents participants with challenging decision problems in a context of great strategic uncertainty.

This paper is a case study of a strategic issue that arose in the first TAC/SCM tournament. We present our reasoning about the issue, and our effort to perturb the environment from an "equilibrium" we considered undesirable, to another more profitable domain of operation. We recount the experience as it played out in the competition, and analyze the outcome of this naturalistic experiment. We then perform a more controlled experimental analysis of the issue, applying empirical game-theoretic methods to produce compelling results, narrow in scope but arguably accounting well for strategic interactions.

The direct result of this study is validation of the insight behind our particular policy for strategic procurement in the TAC/SCM game. Our experimental analysis verifies that the predominant patterns we observed among agents in the tournament reflect strategically rational behavior for this issue. It also confirms the surprising phenomenon in which a tactic designed to preempt the actions of others actually leads to global welfare improvements. More broadly, we view this exercise as illustrating a general approach by which agent designers can reason through pivotal strategic issues in a principled way, despite computational and analytical intractability of their environments.

2 Strategic Analysis for Complex Domains

Given the significant strategic interactions in the supply-chain game, we as agent designers would naturally like to apply game-theoretic tools to predict the behavior of other agents, individually and in aggregate. Unfortunately, direct solution of the full game in the sense of finding equilibria (of any variety) seems out of the question on tractability grounds. Therefore, we seek other ways to exploit the concepts and methods of game theory, short of comprehensive solutions.

The economic literature on game-theoretic applications is full of highly abstracted, or *stylized* models,[1] designed to highlight some particular strategic issue of interest, suppressing or holding constant—in extremely summarized form—the many other factors that would actually be present in any real instance of the modeled scenario. In general, these unmodeled factors may well be relevant to strategic decisions of the players. In suppressing them, the modeler is judging that the effect of these factors would not interact so much as to invalidate the key strategic implications of the modeled factors.

We aim to employ a similarly reasoned use of abstraction, but in a situation where a completely detailed description of the game is actually available. Whereas this availability does not obviate the need for abstraction or other simplifying approximations, it does present opportunities for calibration and validation that may not exist when the modeler is abstracting the real world directly. For instance, we can often summarize the unmodeled features analytically or through statistical simulation, and can achieve some validation through sensitivity analysis or comparison of alternative models based on different simplifying assumptions.

TAC/SCM is an example of a very detailed model intended to capture aspects of a real situation, offering the prospect that interesting phenomena may emerge from

[1]For example, peruse any text on game theory [Fudenberg and Tirole, 1991, Gintis, 2000] for a wealth of such enlightening examples.

interactions of the details, even if unanticipated by designers. Given a game in this form, analysts may form hypotheses about identified features of interest based on reasoning (about the game specification) and observation (of actual game instances), and then construct a more abstracted model focusing on these features. After using the detailed model to quantify the abstracted model, one can solve the abstracted model using standard techniques.

In this study, we construct such a model by estimation applied to data from simulations. Its basis in a sampled experience renders this model an *empirical game*. Since the empirical game considers only a limited repertoire of strategies, it also constitutes what Walsh et al. [2003] term a *heuristic strategy payoff matrix*. The approach we take to construction and analysis of this model builds on our previous application of empirical game-solving methodology [Reeves et al., to appear]. It shares much in motivation and operation with the constrained-equilibrium approach advocated by Armantier et al. [2000], as well as the approach to various strategic environments explored by Kephart and colleagues [Kephart and Greenwald, 2002, Kephart et al., 1998, Walsh et al., 2002].

The present study contributes several elements to this emerging empirical game-theoretic approach. First, we employ standard variance-reduction techniques, in particular the method of control variates, to reduce the amount of simulation required to obtain statistically meaningful results. Second, from the empirical game we derive both symmetric (mixed strategy) and non-symmetric (pure strategy) Nash equilibria. We employ statistical hypothesis testing as well as analysis of the benefits from deviation as means to assess the stability of candidate strategy profiles.

Our investigation of strategic procurement in TAC/SCM illustrates more generally how the combination of gaming, simulation, experimental manipulation, and game-theoretic analysis can provide insight into a complex strategic environment. We believe that these methods comprise a promising methodology for problems characterized by significant dynamics and uncertainty, with fine-grained interaction among agents. Real-world commerce environments (e.g., supply chains) very often exhibit these features. For such problems, comprehensive direct analysis of detailed models is intractable, but ex ante stylization risks abstracting away pivotal features. Our methodology does not completely avoid these risks, but increases confidence through validation based on maintaining an explicit bridge between the detailed and stylized models.

3 TAC/SCM Game

In the TAC/SCM scenario,[2] six agents representing PC (personal computer) assemblers operate in a common market environment, over a simulated production horizon. The environment constitutes a *supply chain*, in that agents trade simultaneously in markets for supplies (PC components) and the market for finished PCs. Agents may assemble for sale 16 different models of PCs, defined by the compatible combinations of the four

[2]For complete details of the game rules, see the specification document [Arunachalam et al., 2003]. This is available at `http://www.sics.se/tac`, as is much additional information about TAC/SCM and TAC in general. Arunachalam and Sadeh [to appear] discuss the challenges posed by the game, and present an account of the 2003 tournament.

component types: CPU, motherboard, memory, and hard disk.

Figure 1 diagrams the basic configuration of the supply chain. The six agents (arrayed vertically in the middle of the figure) procure components from the eight suppliers on the left, and sell PCs to the entity representing customers, on the right. Trades at both levels are negotiated through a *request-for-quote* (RFQ) mechanism, which proceeds in three steps:

1. Buyer issues RFQs to one or more sellers.
2. Sellers respond to RFQs with *offers*.
3. Buyers accept or reject offers. An accepted offer becomes an *order*.

The suppliers and customer implement fixed negotiation policies, defined in the game specification, and discussed in detail below where applicable.

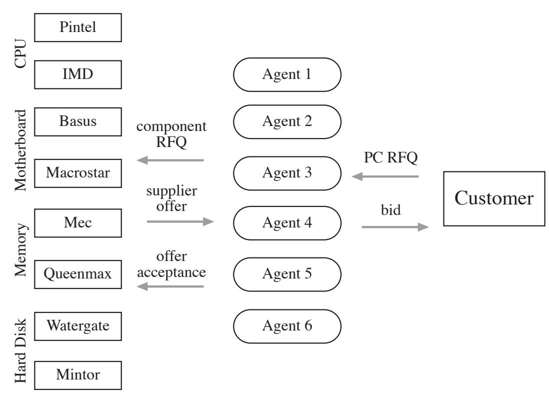

Figure 1: TAC/SCM supply chain.

The game runs for 220 simulated days. On each day, the agent may receive offers and component delivery notices from suppliers, and RFQs and offer acceptance notifications from customers. It then must make several decisions:

1. What RFQs to issue to component suppliers.
2. Given offers from suppliers (based on the previous day's RFQs), which to accept.
3. Given component inventory and factory capacity, what PCs to manufacture.
4. Given inventory of finished PCs, which customer orders to ship.
5. Given RFQs from customers, to which to respond and with what offers.

In the simulation, the agent has 15 seconds to compute and communicate its daily decisions to the game server. At the end of the game, agents are evaluated by total profit, with any outstanding component or PC inventory valued at zero.

As we describe below, a key stochastic feature of the game environment is level of demand for PCs. The underlying demand level is defined by an integer parameter Q

(called RFQ_{avg} in the specification document [Arunachalam et al., 2003, Section 6]). Each day, the customer issues a set of $\hat{Q}$ RFQs, where $\hat{Q}$ is drawn from a Poisson distribution with mean value defined by the parameter Q for that day. Since the order quantity, PC model, and reserve price are set independently for each customer RFQ, the number of RFQs serves as a sufficient statistic for the overall demand, which in turn is a major determinant of the potential profits available to the agents.

The demand parameter Q evolves according to a given stochastic process. In each game instance, an initial value, Q_0, is drawn uniformly from [80,320]. If Q_d is the value of Q on day d, then its value on the next day is given by [Arunachalam et al., 2003, Section 6]:

$$Q_{d+1} = \min(320, \max(80, \tau_d Q_d)), \tag{1}$$

where τ is a trend parameter that also evolves stochastically. The initial trend is neutral, $\tau_0 = 1$, with subsequent trends updated by a perturbation $\epsilon \sim U[-0.01, 0.01]$:

$$\tau_{d+1} = \max(0.95, \min(1/0.95, \tau_d + \epsilon)). \tag{2}$$

In a given game, the demand may stay at predominantly high or low levels, or oscillate back and forth. The probabilistic behavior of Q figures importantly in our analysis, as presented in Section 7 below.

Although it necessarily simplifies the PC market and manufacturing process to a great extent, the TAC/SCM game does introduce several realistic elements not typically incorporated in trading games. Specifically, it embeds trading in a concrete production context, and incorporates stochastic (and partially observable) exogenous effects. Like actors on real supply chains, TAC/SCM agents make decisions under uncertainty over time, dealing with both suppliers and customers, in a competitive market environment. Negotiation concerns several facets of a deal (price, quantity, delivery time, penalty), and takes place simultaneously at multiple tiers.

4 Deep Maize

The University of Michigan's entry in TAC-03/SCM is an agent called Deep Maize [Kiekintveld et al., 2004a,b]. The agent is organized in modular functional units controlling procurement, manufacturing, and sales. Its behavior is based on distributed feedback control, in that it acts to maintain a reference zone of profitable operation. To coordinate the distributed modules, Deep Maize employs aggregate price signals, derived from a market equilibrium analysis and continual Bayesian demand projection. The design of Deep Maize optimizes for performance in the steady state, with little explicit attention to transient or end-game behaviors.

In the present study we focus on one pivotal feature of Deep Maize's strategy, described in full detail below. We thus defer specifics of the rest of our agent's strategy to our other reports (which in turn do not address the strategic analysis presented here).

5 Day-0 Procurement Strategies

A close examination of the game rules suggests that procurement of components at the very beginning of the game (*day-0 procurement*) may be a pivotal strategic issue. This was indeed borne out by the behavior observed in preliminary rounds of the tournament, as discussed below. In this section, we explain the reason for expecting day-0 procurement to be so significant, and its ramifications for Deep Maize and other agents.

5.1 Supplier Pricing

In the TAC/SCM market, suppliers set prices for components based on an analysis of their available capacity. Conceptually, there exist separate prices for each type of component, from each supplier. Moreover, these prices vary over time: both the time that the deal is struck, and time that the component is promised for delivery.

The TAC/SCM component catalog [Arunachalam et al., 2003, Figure 3] associates every component c with a *base price*, b_c. The correspondence between price and quantity for component supplies is defined by the suppliers' pricing formula [Arunachalam et al., 2003, Section 5.5]. The price offered by a supplier at day d for an order to be delivered on day $d + i$ is

$$p_c(d + i) = b_c - 0.5 b_c \frac{\kappa_c(d + i)}{500i}, \tag{3}$$

where for any j, $\kappa_c(j)$ denotes the cumulative capacity for c the supplier projects to have available from the current day through day j. The denominator, $500i$, represents the *nominal capacity* controlled by the supplier over i days, not accounting for any capacity committed to existing orders.

Supplier prices according to Eq. (3) are date-specific, depending on the particular pattern of capacity commitments in place at the time the supplier evaluates the given RFQ. A key observation is that component prices are never lower than at the start of the game ($d = 0$), when $\kappa_c(i) = 500i$ and therefore $p_c(i) = 0.5b_c$, for all c and i.[3] As the supplier approaches fully committed capacity ($\kappa_c(d + i) \to 0$), $p_c(d + i)$ approaches b_c.

In general, one would expect that procuring components at half their base price would be profitable, up to the limits of production capacity. Customer reserve prices range between 0.75 and 1.25 the base price of PCs, defined as the sum of base prices of components. Therefore, unless there is a significant oversupply, prices for PCs should easily exceed the component cost, based on day-0 prices.

An agent's procurement strategy must also take into account the specific TAC/SCM RFQ process. Each day, agents may submit up to 10 RFQs, ordered by priority, to each supplier. The suppliers then repeatedly execute the following, until all RFQs are

[3] As discussed below, this creates a powerful incentive for early procurement, with significant consequences for game balance. In retrospect, the supplier pricing rule was generally considered a design flaw in the game, and has been substantially revised for the 2004 TAC/SCM tournament.

exhausted: (1) randomly choose an agent,[4] (2) take the highest-priority RFQ remaining on its list, (3) generate a corresponding offer, if possible. In responding to an RFQ, if the supplier has sufficient available capacity to meet the requested quantity and due date, it offers to do so according to its pricing function. If it does not, the supplier instead offers a partial quantity at the requested date and/or the full quantity at a later date, to the best of its ability given its existing commitments. In all cases, the supplier quotes prices based on Eq. (3), and reserves sufficient capacity to meet the quantity and date offered.

5.2 Implications of Aggressive Day-0 Procurement

From the discussion above, it would appear advantageous to any agent that it attempt to procure a large number of components on day 0. We call this strategy *aggressive day-0 procurement*, or simply *aggressive*. From each agent's perspective, the main effect of being aggressive is on its own component procurement profile. If every agent is aggressive, however, it can significantly change the character of the game environment.

An aggressive day-0 procurement commits to large component orders before overall demand over the game horizon is known. This leaves agents with little flexibility to respond to cases of low demand, except by lowering PC prices to customers. Since component costs are sunk at the beginning, there is little to keep prices from dropping below (ex ante) profitable levels. Ketter et al. [2004] confirm the importance of day-0 procurement in determining overall performance, finding a strong positive correlation between components obtained on day 0 and profitability for high-demand games, and a negative correlation for games characterized by low demand.

As more agents procure aggressively, several factors make aggressiveness even more compelling. The aggressive agents reserve significant fractions of supplier capacity, thus reducing subsequent availability and raising prices, according to their pricing function (3). A natural response might induce a "race" dynamic, where agents issue day-0 RFQs in increasingly large chunks, ultimately requesting all components they expect to be able to use over the entire game horizon. Not only does this exacerbate the risk of locking in aggregate oversupply, it also produces a less interleaved and more unbalanced distribution of components, especially at the beginning of the game. This in turn can prevent many agents from being able to acquire key components needed for particular PC models until relatively far into the production year.

For all these reasons, the aggressive strategy is appealing to individual agents, yet potentially quite damaging for the agent pool overall. We considered this situation particularly bad for our agent, given that it was designed for high performance in the steady state [Kiekintveld et al., 2004a]. Deep Maize devotes a considerable effort toward developing accurate demand projections, and thus is quite responsive to actual demand conditions. If most of the game's component procurement is up front, we never reach a steady state, and the ability to respond to demand conditions is much less relevant.

[4] At the start of the game, suppliers select among agents with equal probability. Thereafter, suppliers employ a reputation mechanism whereby the probability of agent choice depends on its record of accepting previous offers. We discuss the operation and effectiveness of this mechanism in Section 6.2.

Agent	Affiliation	Average Profit ($M)		
		Qualifying	Seeding 1	Seeding 2
TacTex	U Texas	33.65	32.66	32.97
RedAgent	McGill U	15.09	24.57	29.52
Botticelli	Brown U	13.88	17.29	28.03
Jackaroo	U Western Sydney	14.89	35.55	19.23
WhiteBear	Cornell U	–3.17	13.57	16.50
PSUTAC	Pennsylvania State U	–120.0	15.52	15.25
HarTAC	Harvard U	12.41	4.19	10.72
UMBCTAC	U Maryland Baltimore Cty	–13.94	30.16	10.23
Sirish		–109.4	–0.17	8.27
Deep Maize	U Michigan	1.85	0.45	7.49
TAC-o-matic	Uppsala U	0.22	1.79	7.07
RonaX	Xonar GmbH	–0.92	9.24	4.29
MinneTAC	U Minnesota	10.88	6.56	–0.32
Mertacor	Aristotle U Thessaloniki	9.29	–0.38	–3.53
zepp	Poli Bucharest	–24.83	–7.80	–5.46
PackaTAC	N Carolina State U	–5.11	–25.67	–5.71
Socrates	U Essex	–48.94	–3.31	–6.84
Argos	Bogazici U	3.65	–4.24	–8.43
DummerAgent		–8.08	–20.56	—
DAI_hard	U Tulsa	–11.36	–39.05	—

Table 1: TAC-03/SCM tournament participants, and their performance in preliminary rounds. Results from the qualifying rounds are weighted, seeding rounds are unweighted.

The Deep Maize development team therefore decided not to employ aggressive day-0 procurement in the preliminary rounds, instead treating it just like any other day. We did not really expect that others would miss the opportunity, but did not want to encourage or accelerate it.

6 TAC-03 Tournament

The twenty agents who participated in the TAC-03/SCM tournament are listed in Table 1. The table presents average scores from each of three preliminary rounds, measured in millions of dollars of profit. Results from the semifinal and final rounds are presented in Section 6.3 below.

Two seeding rounds were held during the periods 7–11 and 14–18 July,[5] with each agent playing 60 and 66 games, respectively. Two agents were eliminated based on scores and/or inactivity after Seeding Round 1. The remaining 18 agents advanced to

[5] An earlier “qualifying” round spanned 16–27 June, but this was mainly for debugging and no agents were eliminated.

the semifinals, with assignment to heats based on standing in Seeding Round 2. The semifinals and finals were held live at IJCAI-03, 11–13 August in Acapulco, Mexico, each round consisting of nine games in one day. Semifinal 1 heat 1 (S1H1) comprised agents seeded 1–6 and 16–18, and the 7–15 seeds played in S1H2. The top six teams from each S1 heat (9 games) proceeded to the second semifinal round. S2H1 comprised teams ranked 1–3 in S1H1, and those ranked 4–6 in S1H2. The top three in S1H2 played, along with the second three in S1H1, in S2H2. The top three from each of S2H1 and S2H2 then entered the finals on 13 August. Further details about the TAC-03 tournament are available at `http://www.sics.se/tac`.

6.1 Evolution of Day-0 Policies in Preliminary Rounds

As we expected, competition entrants noticed the individual advantages of aggressive day-0 procurement. Early in the qualifying rounds we noticed Jackaroo's distinct saw-toothed profit timeline, indicating a steady increase in wealth with large periodic drops corresponding to bulk deliveries of supplies. This pattern was the result of large supply orders placed early in the game (over the first seven days, not just day 0) for delivery at regular intervals [Zhang et al., 2004].

Based on our subsequent analysis of early game logs,[6] we can identify TacTex [Pardoe and Stone, 2004] as the first to employ an aggressive day-0 strategy in competition. In their very first qualifying round game, TacTex requested 8000 of each component from each supplier. Although we have found many agents performed mild day-0 procurement during the qualifying rounds, TacTex was more aggressive, earlier—likely a factor in their supremacy this first round.

Throughout the first seeding round, more agents began using increasingly aggressive day-0 procurement strategies. In particular we noticed the successful agents TacTex, Botticelli, RedAgent, UMBCTAC, and Jackaroo ordering very large quantities on day 0 and very little later in the game. Interestingly, there was no discussion of the issue on the TAC/SCM message boards, possibly because entrants recognized its strategic sensitivity. By the second seeding round it was obvious that the majority of agents were using aggressive strategies. In particular, we verified that all the agents that placed higher than Deep Maize in the second seeding round (see Table 1) employed aggressive day-0 procurement.

While observing the increase in aggressiveness, we compiled detailed dossiers describing the day-0 strategies of other agents. We hoped to use this data to understand how widespread the use of day-0 procurement had become, and to understand how it was affecting the dynamics of the game.

6.2 Deep Maize Preemptive Strategy

After much deliberation, we decided that the only way to prevent the disastrous rush toward all-aggressive equilibrium was to *preempt* the other agents' day-0 RFQs. By requesting an extremely large quantity of a particular component, we would prevent

[6] The TAC/SCM game server records all agent actions (e.g., RFQs, manufacturing, bids) along with supplier and customer behavior, and releases the log files after each game instance is complete.

the supplier from making reasonable offers to subsequent agents, at least in response to their requests on that day. Our premise was that it would be sufficient to preempt only day-0 RFQs, since after day 0 prices are not so especially attractive.

The Deep Maize preemptive strategy operates by submitting a large RFQ to each supplier for each component produced. The preemptive RFQ requests 85000 units—representing 170 days' worth of supplier capacity—to be delivered by day 30. See Figure 2. It is of course impossible for the supplier to actually fulfill this request. Instead, the supplier will offer us both a partial delivery on day 30 of the components they can offer by that date (if any), and an earliest-complete offer fulfilling the entire quantity (unless the supplier has already committed 50 days of capacity). With these offers, the supplier reserves necessary capacity. This has the effect of preempting subsequent RFQs, since we can be sure that the supplier will have committed capacity at least through day 172. (The extra two days account for negotiation and shipment time.) We will accept the partial-delivery offer, if any (and thereby reject the earliest-complete), giving us at most 14000 component units to be delivered on day 30, a large but feasible number of components to use up by the end of the game.

Figure 2: Deep Maize's preemptive RFQ.

In the situation that our preemptive RFQ gets considered after the supplier has committed 50 days of production to other agents, we will not receive an offer, and our preemption is unsuccessful. For this reason, we also submit backup RFQs of 35000 to be delivered on day 50, and 15000 to be delivered on day 70.

The effect of a preemptive RFQ clearly depends on the random order by which its target supplier selects agents to consider. On each selection, Deep Maize will be picked with probability 1/6, which means that with probability $1 - \frac{5}{6}^k$ it is selected within the first k RFQs. For example, the preemptive RFQ appears among the first four 51.8% of the time. Given these orderings are also generated independently for each supplier-component combination, with high probability Deep Maize is expected to successfully preempt a significant number of the other agents' day-0 RFQs.

The TAC/SCM designers anticipated the possibility of preemptive RFQ generation, (there was much discussion about it in the original design correspondence), and took steps to inhibit it. The designers instated a *reputation mechanism*, in which refusing offers from suppliers reduces the priority of an agent's RFQs being considered in the future. This is accomplished by adjusting agent i's selection probability π_i as follows [Arunachalam et al., 2003, Section 5.1]:

$$\begin{aligned} \text{weight}_i &= \max\left(0.5, \frac{\text{QuantityPurchased}_i}{\text{QuantityRequested}_i}\right), \\ \pi_i &= \frac{\text{weight}_i}{\sum_x \text{weight}_x}. \end{aligned}$$

Agent	Average Profit ($M)		
	Semifinal 1	Semifinal 2	Final
RedAgent	12.75 (H1)	25.09 (H1)	11.61
Deep Maize	10.51 (H2)	15.28 (H1)	9.47
TacTex	1.85 (H1)	–15.54 (H2)	5.02
Botticelli	5.69 (H1)	–4.83 (H2)	3.33
PackaTAC	18.31 (H1)	8.70 (H1)	–1.68
WhiteBear	5.26 (H1)	–9.58 (H2)	–3.45
PSUTAC	17.81 (H1)	–1.56 (H1)	—
TAC-o-matic	–1.24 (H2)	–13.50 (H1)	—
Sirish	15.86 (H2)	–20.21 (H2)	—
MinneTAC	13.92 (H2)	–24.98 (H2)	—
UMBCTAC	10.78 (H2)	–29.91 (H2)	—
HarTAC[8]	2.59 (H2)	–32.95 (H1)	—

Table 2: Results for twelve agents participating in the second semifinal and final rounds.

Even with this deterrent, we felt our preemptive strategy would be worthwhile. Since most agents were focusing strongly on day 0, priority for RFQ selection on subsequent days might not turn out to be crucial.

6.3 Tournament Story

Having developed the preemptive strategy, we still faced the question of when to deploy it. Based on our performance in preliminaries, we were reasonably confident that we could make the top six out of nine in S1H2 without resorting to preemption, and instead chose to implement a moderate form of aggressive day-0 procurement. As expected, other agents actually scaled up their day-0 procurement, and consequently, Deep Maize did not put on a very strong showing in this round. Fortunately, fourth place was sufficient to advance to the next round.

Table 2 presents results for the top twelve agents after Semifinal 1. Network problems at the competition venue caused difficulties for agents running locally—Jackaroo and HarTAC, in particular.[7]

After the first semifinal closed, the next few hours were filled with a great deal of hustle as the team activated the preemptive strategy that would be played the next day. These hours were also filled with anxiety. We had intuition about the effect of preemptive strategy on Deep Maize and other agents, but had never had a chance to test it against other competitors. On the other hand, we could hardly wait to see the "unexpected" dramatic change in Deep Maize behavior in the arena with presumably

[7]The problems did not affect the majority of agents communicating over the Internet from entrants' home institutions to the servers in Sweden.

[8]The score of HarTAC in Semifinal 2 was adversely affected by one game in which it experienced connectivity problems and lost $364M. Omitting this game would boost their average profit to $8.46M.

	S1H1	S1H2	S2H1	S2H2	Finals
(DM?, P?, N)	–,–,9	DM,–,9	DM,P,8	–,–,9	DM,P,16
components	59390	46989	27377	70744	27172
avg profits	2.97	–3.05	7.02	–17.51	4.05

Table 3: Effect of preemption on day 1 component orders and average profits.

the three best agents (since we did not place very highly in the first round, we would play the top three placing agents from the other heat).

In the morning of 12 August, the Deep Maize team stood waiting by the computer screen as the second round of semifinals began. As day 29 rolled around, everyone held their breath, releasing it when the first large delivery of components dropped in. Once we saw distinct manifestations of the preemptive strategy, we began to wonder how other agents would react. Our suspense did not last long: soon after the game's midpoint, a comment emerged in the TAC game chatroom: "why we can't get hard disks? How server handle purchase RFQs? is the administrator around!!!?" Apparently, one agent at least was taking for granted that its day-0 requests would be fulfilled.

At the end of S2H1, Deep Maize came in second behind the eventual tournament winner, Red Agent [Keller et al., 2004], followed closely by PackaTAC [Dahlgren, 2003]. These agents, it turned out, were relatively resilient to the preemptive strategy, as they did not excessively rely on day-0 procurement, but adaptively purchased components throughout the game.

Although none had anticipated it explicitly, it turned out that most agents playing in the finals were individually flexible enough to recover from day-0 preemption. By preempting, it seemed that Deep Maize had leveled the playing field, but RedAgent's apparent adaptivity in procurement and sales [Keller et al., 2004] earned it the top spot in the competition rankings.

6.4 Analysis

Did Deep Maize's preemption strategy work? We can first examine whether it had its intended direct effect, namely, to reduce the number of components ordered at the very beginning of the game. Table 3 presents, for each tournament round, the number of components ordered on day 1 (based on day-0 RFQs). Each value represents a total over delivery dates and agents, averaged over the 16 supplier-component pairs. Above the component numbers we indicate whether Deep Maize played in that round (DM), whether it employed preemption (P), and the number of games. Note that this data includes one game in S2H1 and two in the finals in which Deep Maize failed to preempt due to network problems. It does exclude one anomalous S2H1 game, in which HarTac experienced connectivity problems, to wildly distorting effect.

From the table, it is clear that the preemptive day-0 strategy had a large effect. The difference is most dramatic in Semifinal 2, where the heat with Deep Maize preempting saw an average of 27377 components committed on day 1, as compared to 70744 in the heat without Deep Maize.

The tournament results also indicate that preemption was successful. The fact that Deep Maize performed well overall is suggestive, though of course there are many other elements of Deep Maize contributing to its behavior. Evidence that the preemptive strategy in particular was helpful can be found in the results from Semifinal 1, where Deep Maize did not preempt and ended up in fourth place. This was sufficient for advancing in the tournament, but clearly not as creditable as its second place showing in the finals, among the (presumably) top agents in the field.

We can conclude, then, that preemption helped Deep Maize. How did it affect the rest of the field? Table 3 also suggests a positive relation between preemption and profits averaged over *all* agents. Again, the contrast is greatest between S2H1 and S2H2. In the heat without Deep Maize, it appears that competition among aggressive agents led to an average *loss* of $17.51M. With Deep Maize preempting in S2H1, profits are a healthy $7.02M per agent. Preemption was also operative in the finals, and profits there were also positive. That it is preemption and not Deep Maize per se is supported by examination of Semifinal 1, in which the heat without our agent appears to be substantially more profitable on average.

Pooling all of these semifinal and final games, we compared average profits for games with and without preemption. Games with preemption averaged $3.97M in profits, compared to a loss of $4.02M in games without preemption. Given the small dataset and large variance, this difference is only marginally statistically significant ($p = .09$).

Drawing inferences from tournament results is complicated by the presence of many varying and interacting factors. These include details of participating agents, and random features of environment, in particular the level of demand. To test the influence of demand, we measured the overall demand level for each game, $\bar{Q}$, defined as the average number of customer RFQs per day. Figure 3 presents a scatterplot of the tournament games, showing $\bar{Q}$ and per-agent profits for each. We distinguish the games with and without preemption, and for each class, fit a line to the points. The linear fit was quite good for the games with preemption ($R^2 = 0.84$), capturing somewhat less of the variance for the games without ($R^2 = 0.66$).

As seen in the figure, with or without preemption, demand clearly exhibits a significant ($p < 10^{-6}$) relation to profits. The relation is attenuated by preemption, and indeed the revealed trend indicates that preemption is beneficial when demand is low, and detrimental in the highest-demand games. This is what we would expect, given that the primary effect of preemption is to inhibit early commitment to large supplies. Given the apparently important influence of demand, we developed a more elaborate mechanism (described in Section 7) to control for demand in our analysis of tournament games as well as our post-competition experiments.

6.5 ICEC Exhibition Tournament

From 1–3 October 2003, organizers of the Fifth International Conference on Electronic Commerce (ICEC) conducted a TAC/SCM exhibition event. Twelve agents participated, including all six finalists from the TAC-03 tournament. Deep Maize ran unchanged since August, and we suspect that most others (with one exception noted below) were also the same as their competition versions. The exhibition tournament

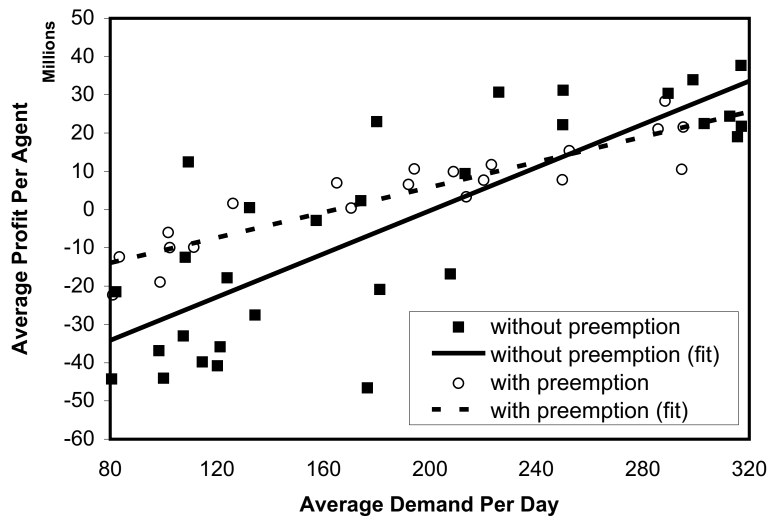

Figure 3: Profits versus $\bar{Q}$ in TAC-03 tournament games. The lines represent best fits to data from games with and without preemption.

was organized as a series of randomly-matched agent profiles, with each participant playing 26 games.

It is important not to draw firm conclusions from the results of any exhibition event, and indeed the ICEC results are particularly noisy due to the high absentee rate (15%) of participating agents. We examined the data fairly closely, in part to understand why Deep Maize had the highest average profits (despite being "absent" from five games), and of course to examine the effect of preemption. At first the results seemed quite anomalous to us, until we discovered that another agent—Botticelli—had been modified to play a preemptive strategy as well!

We plot average profit versus demand for the ICEC games in Figure 4. We partition the 52 games into three classes, based on whether there were zero, one, or two agents playing a preemptive day-0 strategy. The fitted lines for these cases suggest that preemption ameliorated the effect of demand here as well. However the data is quite noisy, and none of the comparisons of average profits are even remotely statistically significant.

7 Demand Adjustment

Given a sufficient number of random instances, the problem of variance due to stochastic demand would subside, as the sample means for outcomes of interest would converge to their true expectations. However, for TAC/SCM, sample data is quite expensive, as each game instance takes approximately one hour. (55 minutes of game simulation time, plus a few minutes for pre- and post-game processing) Therefore, datasets from tournaments and even offline experiments will necessarily reflect only

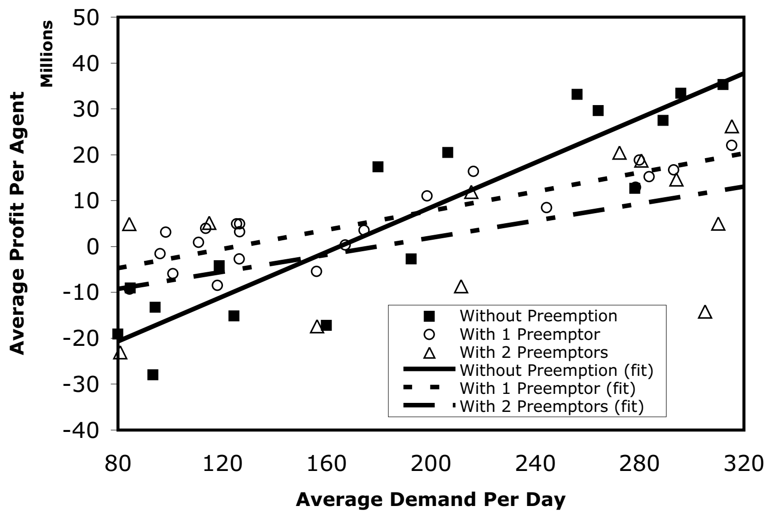

Figure 4: Profits versus $\bar{Q}$ in the ICEC-03 exhibition tournament.

limited sampling from the distribution of demand environments.

7.1 Demand-Adjusted Profit

To address this issue, we can calibrate a given sample with respect to the known underlying distribution of demand ($\bar{Q}$). Our approach is an instance of the standard method of variance reduction by *control variates* [L'Ecuyer, 1994, Ross, 2002]. Given a specification for the expectation of some game statistic y as a function of $\bar{Q}$, its overall expectation accounting for demand is given by

$$E[y] = \int_{\bar{Q}} E[y|\bar{Q}] \Pr(\bar{Q}) d\bar{Q}. \tag{4}$$

Although we do not have a closed-form characterization of the density function $\Pr(\bar{Q})$, we do have a specification of the underlying stochastic demand process. From this, we can generate Monte-Carlo samples of demand trajectories over a simulated game.[9] We then employ a kernel-based density estimation method using Parzen windows [Duda et al., 2000] to approximate the probability density function for $\bar{Q}$. This distribution is shown in Figure 5. Its mean is 196, with a standard deviation of 77.4. Note that much of the probability is massed at the extremes of demand, with a skew toward the low end. The tendency toward the extremes comes from the combination of trend (τ) momentum and bounding of Q. The skew toward the low end comes from the fact that the trend is multiplicative, so the process tends to transition more rapidly while at the higher levels of demand.

[9] We could also use historical game data, but simulating Eqs. (1) and (2) is much faster. The 200,000 data points we generated for our density estimate would take 22.8 years of game simulation time to produce.

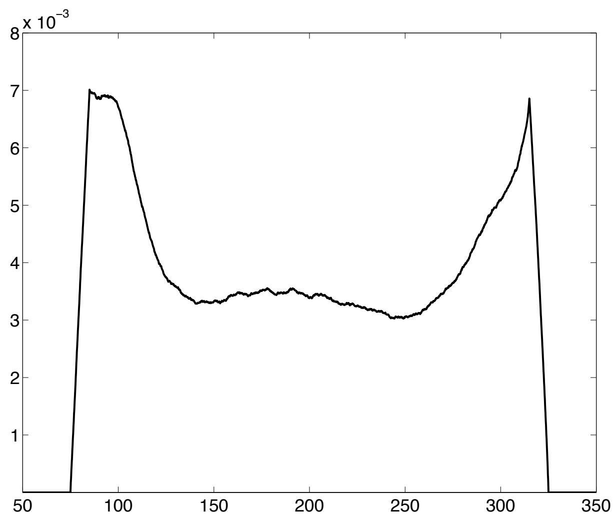

Figure 5: Probability density for average RFQs per day ($\bar{Q}$).

Given this distribution, we define *demand-adjusted profit* (DAP) as the expected profit, adjusted for demand. We calculate this by substituting the per-agent profit for y in Eq. (4). Using this formula requires an estimate for profits as a function of $\bar{Q}$, which we obtain by linear regression from the sample data. The two lines in Figure 3 thus represent our estimates for profits given $\bar{Q}$ for the two sets of TAC-03 tournament games. Although the actual relationship is not linear, the fitted line provides an estimate of the mean equivalent to that of the control variates method [L'Ecuyer, 1994]. As we see below, for limited samples, adjusting for $\bar{Q}$ in this manner indeed produces a substantial reduction in variance, introducing only a small additional bias.

7.2 Variance Reduction

To evaluate the effectiveness of demand adjustment as a means to increase the statistical power of our limited sample, we performed a simple experiment comparing the accuracy of DAP estimates to that of unadjusted mean profit. For one (arbitrarily) selected profile, we collected a particularly large number of samples, D (in our experiment, $|D| = 439$). For each value $1 \leq m \leq |D|$, we then measured the bias and variance of $N = 1000$ independent subsamples of size m of D, employing both raw and demand-adjusted profits for one of the agents. To define a gold standard, we treat the sample mean of raw profits over D as the true mean, z. We can then define the bias, variance, and mean squared error (MSE) of a particular estimate v as follows:

$$\begin{aligned} \text{bias}^2 &\equiv (\bar{v} - z)^2 \\ \text{variance} &\equiv \frac{1}{N}\sum_{i=1}^{N}(v_i - \bar{v})^2 \\ \text{MSE} &\equiv \text{variance} + \text{bias}^2 \\ &= \frac{1}{N}\sum_{i=1}^{N}(v_i - z)^2 \end{aligned}$$

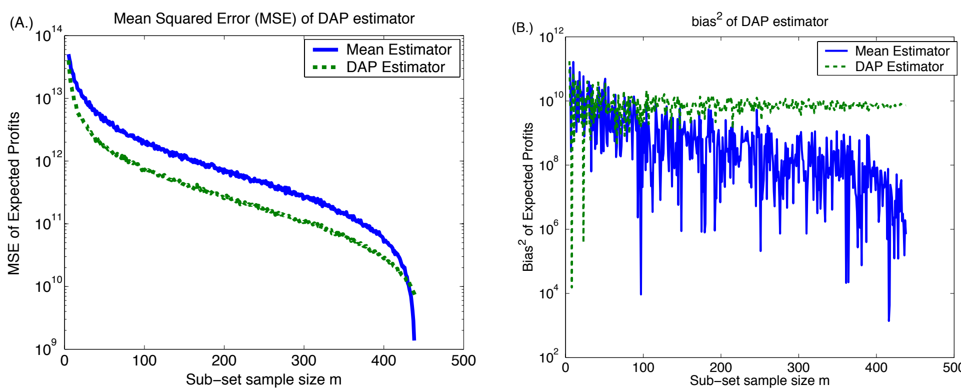

Figure 6: (A.) Mean squard error (MSE) of DAP estimator. (B.) $Bias^2$ of DAP estimator.

Figure 6 compares the MSE and bias of estimates using DAP and unadjusted means. Note that MSE is dominated by variance, as bias^2 is several orders of magnitude smaller, for either estimator. Only as m approaches $|D|$, where variance approaches zero by construction,[10] do we see bias become a significant component of MSE. We also observe in Figure 6(B) that the $bias$ of the DAP estimator does not change as we increase m. It becomes less noisy (due to the reduced variance), but the bias itself remains consistent.

Another way to evaluate the benefit of demand adjustment is to determine the number of samples required for the DAP estimate to achieve the same MSE as the mean estimator with m samples. Figure 7 shows this relationship in our experimental data, demonstrating that we always need strictly fewer samples using DAP to achieve a given level of MSE. As shown here, DAP can reduce the number of required samples by up to 50%.

7.3 DAP Analysis of Preemption in TAC-03

From the linear model of profits given $\bar{Q}$, we can obtain a summary comparison of overall profits with and without preemption. For the TAC-03 games without preemption,

[10]This is an artifact of our assumption that the sample mean from our fixed dataset D is the true mean.

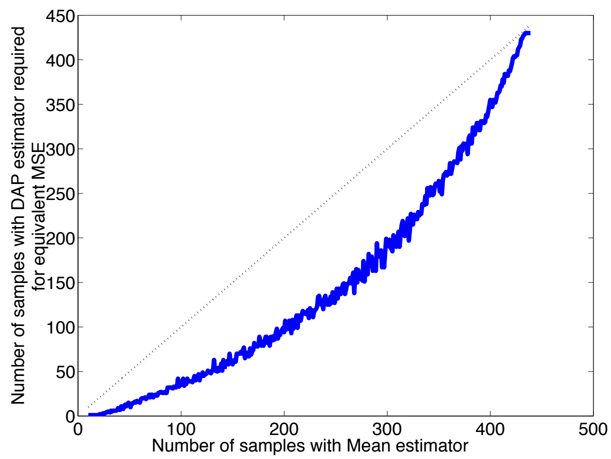

Figure 7: Number of samples required by DAP estimator to acheive the same MSE as a given number of samples with the mean estimator.

DAP was –\$1.41M. Preemption increased DAP to \$5.20M. Thus, we find that on average, Deep Maize's preemptive strategy improved not only its own profits, but those of the other agents as well. These results are corroborated by controlled experiments described below.

8 Game-Theoretic Model

Although the tournament results presented above are illuminating, it is difficult to support general conclusions due to the many contributing factors and differences among agents. To isolate the effect of preemption on the key strategic variable (aggressiveness of day-0 procurement), we developed a stylized game-theoretic model, then calibrated it using simulation experiments. Game-theoretic analysis supports conclusions about *equilibrium* behavior of *rational* agents, providing further evidence that the phenomena observed are not merely transient outcomes produced by a particular set of irrational agents.

We proceed by defining a normal-form game, restricting attention to characteristic strategies, and estimating payoffs through simulation experiments. We then identify a series of equilibria, pure and mixed, with respect to various restrictions of the game. Examination of these equilibria and their payoffs confirms that the strategic dynamic observed in the 2003 TAC/SCM tournament obtains in our more controlled setting as well.

8.1 Normal-Form Model Structure

As noted at the outset, TAC/SCM defines a six-player game of incomplete and imperfect information, with an enormous space of available strategies. The game is *symmet-*

ric [Cheng et al., 2004], in that agents have identical action possibilities, and face the same environmental conditions. In our stylized model, we restrict the agents to three strategies, differing only in their approach to day-0 procurement. Each strategy is implemented as a variant of Deep Maize. By basing the strategies on a particular agent, we clearly cannot capture the diversity of approaches to all aspects of TAC/SCM. Fixing much of the behavior, however, enables our focus on the particular issue of strategic procurement.

In strategy A (aggressive), the agent requests large quantities of components from every supplier on day 0. The specific day-0 RFQs issued correspond to aggressive day-0 policies we observed for actual TAC-03/SCM participants. We encoded four of these as RFQ quantity lists:

1. (4250,5000,5000,2500,1250), based on TacTex [Pardoe and Stone, 2004].
2. (3000,3000,3000,3000,3000), based on UMBCTAC.
3. (4000,3000,8000), based on HarTac [Dong et al., 2004].
4. (1672,1672,1672,1672,1672), and double this for CPU components, based on Botticelli [Benisch et al., 2004].

Strategy A randomly selects among these at the beginning of each game instance.

In strategy B (baseline), the agent treats day 0 just like any other day, issuing requests according to its standard policy of serving anticipated demand and maintaining a buffer inventory [Kiekintveld et al., 2004a]. Strategy P (preemptive) is actually Deep Maize as we ran it in the tournament, with preemptive day-0 procurement as described above. Each of these strategies follows the standard Deep Maize procurement policy after day 0.

We consider three versions of this game in our analysis. The first is an *unpreempted* six-player game, where agents are restricted to playing A or B. The second is a five-player game, with the sixth place taken up by a fixed agent playing strategy P. We refer to this as the *single-preemptor* game. The third is the full six-player game where agents are allowed to play any of the three strategies A, B, or P.

Since the three strategies incorporate specified policies for conditioning on private information, we represent the game in normal form. By symmetry there are only seven distinct profiles for the unpreempted game, corresponding to the number j of agents playing A, $0 \leq j \leq 6$. There are six distinct profiles for the single-preemptor game, and a total of twenty-eight for the full game (including the thirteen from the more restricted games). Payoffs for each profile represent the expected profits for playing A, B, or P, respectively, given the other agents, with expectation taken over all stochastic elements of the game environment.

8.2 Simulation Results

To estimate our game's expected payoff function, we sampled an average of around 30 game instances for each strategy profile—834 in total. For each sample, we collected the average profits for the As, Bs, and Ps, as well as the demand level, $\bar{Q}$. We then used the demand-adjustment method described above to derive DAP for each strategy, which we take as its payoff in that profile.

From this data, we verify that increasing the prevalence of aggressiveness degrades aggregate profits. We show that inserting a single preemptive agent neutralizes the effect of aggressiveness, diminishing the incentive to implement an aggressive strategy, and also ameliorating its negative effects. Moreover, the presence of a preemptor tends to improve performance *for all agents* in profiles containing a preponderance of agents playing A. We then study the equilibrium behavior of each of the three versions of the game. From the empirical game models, we derive asymmetric pure-strategy equilibria, as well as symmetric mixed-strategy equilibria, for each of the games.[11] Comparison of the features of equilibrium behavior in the respective games confirms our findings about the effects of strategic preemption.

To test our hypothesis that aggressive strategy has a negative effect on total profits, we regressed total DAP for each profile on the number of aggressive agents in the profile. For profiles without preemption, the linear relationship was quite strong ($p = 0.0018$, $R^2 = 0.88$), with each A in the profile subtracting \$20.9M from total profits, on average.

In the single-preemptor game, the effect of number of aggressive agents on average total profits was statistically insignificant, explaining little variance ($p = 0.54$, $R^2 = 0.10$). For unpreempted profiles with four or more aggressive players, agents playing either strategy would benefit substantially (at least \$6.5M in average profits) from one of the others (either type) switching to play P. Thus, preemption appears to eliminate the detrimental effect that aggressive agents exert on total profits, and for individual profits as well compared to profiles with a predominance of strategy A.

We also confirmed that preemption levels the playing field, as the difference in average profits between aggressive and baseline agents was on the order of \$10M for the unpreempted profiles, as compared to \$1M for the single-preemptor case. Examining the variance across agents in each particular game, we observe that average variance for unpreempted profiles was an order of magnitude larger than that for profiles with preemption. The variances are tabulated in Table 4.

8.3 Pure Strategy Equilibria

A pure-strategy Nash equilibrium is a strategy profile such that no agent can improve its payoff by changing strategies, assuming all other agents play according to the profile. We identify pure strategy Nash equilibria for both of the two-strategy games, as well as the full three-strategy game.

8.3.1 Two-Strategy Games

In a two-strategy ($\{$A,B$\}$) symmetric game, a profile is defined by the number of As. Profile $0 \leq i \leq N$ is a Nash equilibrium if and only if:

[11] It can be shown that for any N-player two-strategy symmetric game, there must exist at least one equilibrium in pure strategies, and there also must exist at least one symmetric equilibrium (pure or mixed) [Cheng et al., 2004].

Profile	Variance	Baseline DAP ($M)	Aggressive DAP ($M)	Preemptive DAP ($M)	Average DAP ($M)
AAAAAA	8.37E+15	n/a	–12.26	n/a	–12.26
AAAAAB	8.29E+15	–9.03	–13.58	n/a	–12.83
AAAABB	1.14E+16	–10.78	–1.47	n/a	–4.57
AAABBB	9.70E+15	–10.36	9.73	n/a	–0.31
AABBBB	6.79E+15	–2.47	19.16	n/a	4.74
ABBBBB	2.61E+15	–1.83	13.28	n/a	0.69
BBBBBB	1.84E+13	8.17	n/a	n/a	8.17
PAAAAA	6.38E+14	n/a	6.86	9.98	7.38
PAAAAB	6.36E+14	7.23	8.84	10.29	8.82
PAAABB	6.50E+14	3.62	5.15	9.04	5.29
PAABBB	7.53E+14	6.06	7.34	10.96	7.30
PABBBB	1.18E+15	4.19	5.75	11.41	5.65
PBBBBB	1.03E+15	6.06	n/a	13.64	7.32
PPAAAA	5.90E+14	n/a	3.67	5.45	4.26
PPAAAB	4.24E+14	5.11	4.71	6.69	5.44
PPAABB	4.28E+14	4.55	4.70	6.71	5.32
PPABBB	4.95E+14	1.74	2.57	4.46	2.79
PPBBBB	9.03E+14	4.78	n/a	7.31	5.63
PPPAAA	2.49E+14	n/a	7.41	7.30	7.35
PPPAAB	1.75E+14	5.76	5.84	6.32	6.07
PPPABB	1.99E+14	10.10	10.14	10.08	10.10
PPPBBB	3.51E+14	3.76	n/a	4.30	4.03
PPPPAA	2.33E+14	n/a	2.26	1.50	1.75
PPPPAB	2.13E+14	6.98	7.24	6.16	6.48
PPPPBB	2.87E+14	5.77	n/a	5.69	5.72
PPPPPA	1.43E+14	n/a	7.74	6.64	6.82
PPPPPB	2.04E+14	5.46	n/a	4.39	4.56
PPPPPA	1.19E+14	n/a	n/a	4.14	4.14

Table 4: Payoffs by strategy profile. "Variance" refers to the mean variance across agents for games with the corresponding profile.

1. the payoff to A in i exceeds the payoff to B in $i-1$ (or $i=0$), *and*
2. the payoff to B in i exceeds that to A in $i+1$ (or $i=N$).

We consider the games defined by DAP payoffs, as well as raw average profits. The full set of DAP payoffs are provided in Table 4. As we ran our simulations, we observed that DAP results anticipated those we would obtain from raw averages after collecting more samples. This is consistent with the experiment described in Section 7.2, in which we found that DAP estimates exhibit lower mean-squared-error compared to sample means, for a range of subsample sizes going well beyond what we could collect for each profile. Given our relatively small datasets, therefore, we have greatest confidence in the DAP results. An advantage of the raw averages is that we have associated variance measures, enabling statistical hypothesis testing.

Let iA denote the profile with no preemption, and i agents playing A (the rest playing B). Whether we define payoffs by DAP or raw averages, the unique pure-strategy Nash equilibrium is 4A. That this is an equilibrium for DAP payoffs can be seen by comparing adjacent columns in the bar chart of Figure 8. Arrows indicate for each column, whether an agent in that profile would prefer to stay with that strategy (arrow head), or switch (arrow tail). Solid black arrows denote statistically significant comparisons, as discussed below. Profile 4A is the only one with only in-pointing arrows.

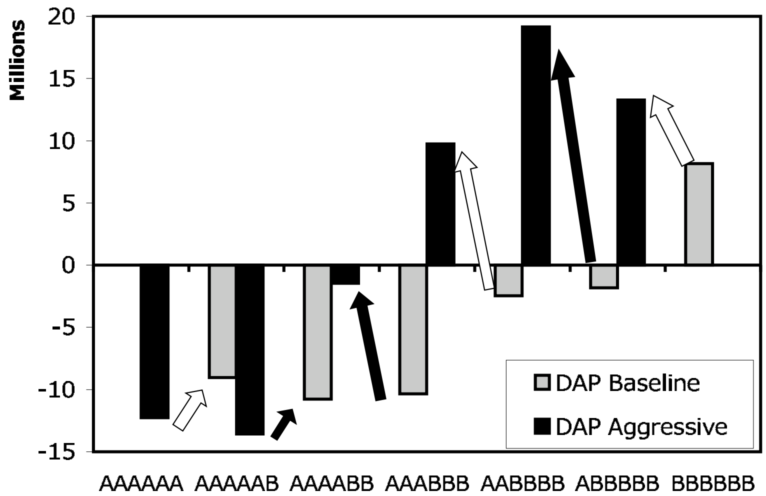

Figure 8: DAP payoffs for unpreempted strategy profiles.

Let PiA denote the profile with a preemptive agent, and i As. In the game with preemption, we find several pure-strategy Nash equilibria. P4A and P2A are equilibria under either payoff measure, and P0A is an equilibrium in the game defined by DAP, but not raw averages. The DAP comparisons are illustrated by Figure 9.

To assess the robustness of these equilibria, we conducted statistical tests. For each relevant comparison we performed two-sample t-tests, using average profits, assuming unequal variance. The p-values are presented in Table 5. Whereas some of the comparisons in the unpreempted game indicate significant differences, for the preemptive profiles none of the comparisons are particularly significant. Thus, the equilibria we found should be considered suspect, or weak equilibria at best. Since payoffs in the preemptive games have much lower variance, if anything we would expect significant differences to show up earlier. This is consistent with our finding above that the preemptive agent neutralizes the difference between strategies A and B. In that respect, identifying an equilibrium is less important in this context.

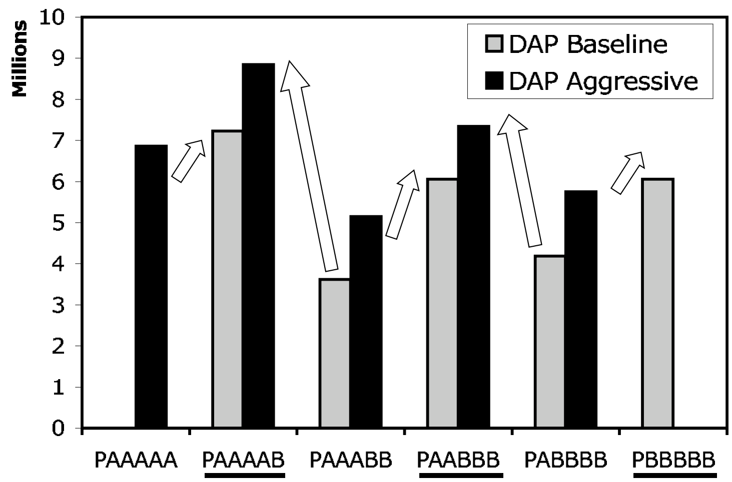

Figure 9: DAP payoffs, with preemption.

Regardless of which equilibrium is played, both A and B agents are clearly better off in the single-preemptor game. In all its equilibria, all agents earn over $6M profit. In the unpreempted game equilibrium (4A), in contrast, all profits are negative, with the B agents losing over $10M each.

8.3.2 Full Three-Strategy Game

Our analysis of the two-strategy games confirms our hypothesis that introducing a single preemptive agent neutralizes the effect of aggressiveness and moves equilibrium play toward a more profitable space. The success of preemption, however, raises the question about whether an incentive to preempt will create a similar mutually destructive competition among preemptors. To check this, we can perform the same kind of equilibrium analysis in the three-player game, where agents are allowed to choose strategy P. The twenty-eight profiles are arrayed in Figure 10, with arrows indicating the transitions between profiles induced by agents switching strategies.

The four pure-strategy Nash equilibria of this game are indicated in bold: PAAAAB, PPBBBB, PPPAAA, and PPPABB. Although the average scores vary across equilibria, in every case the A and B players earn substantial profit, unlike the unpreempted case. Indeed, there exists only one unpreempted profile (2A) from which an A would not deviate, and no unpreempted profiles where playing B is stable.

We also note that as more agents adopt a preemptive strategy, the difference in performance among strategies diminishes. Almost all the comparisons between profiles with preemptive agents are statistically insignificant, as can be seen by the thin arrows in Figure 10. One way to quantify the indifference between strategies given preemption is to consider the ϵ-Nash equilibria. A profile is ϵ-Nash if no agent can improve its payoff by more than ϵ by deviating from its assigned strategy. In Figure 10, we display for each profile the minimum ϵ that would render it an ϵ-Nash equilibrium. For

Comparison	P-value
AAAAAA $\rightarrow$ AAAAAB	0.6595
AAAAAB $\rightarrow$ AAAABB	0.0020
AAAABB $\leftarrow$ AAABBB	0.0246
AAABBB $\leftarrow$ AABBBB	0.5123
AABBBB $\leftarrow$ ABBBBB	0.0001
ABBBBB $\leftarrow$ BBBBBB	0.2879
PAAAAA $\rightarrow$ PAAAAB	0.7678
PAAAAB $\leftarrow$ PAAABB	0.2294
PAAABB $\rightarrow$ PAABBB	0.4413
PAABBB $\leftarrow$ PABBBB	0.3436
PABBBB $\leftarrow$ PBBBBB	0.3845

Table 5: Statistical significance of profile comparisons.

example, although PPPPPA is not an equilibrium, agents can gain at most \$0.34M by deviating from their assigned strategies. Among the 21 preemptive profiles, 17 of them are ϵ-Nash equilibria at an ϵ of \$5.38M or less. In contrast, none of the unpreempted profiles are ϵ-equilibria at that level.

8.4 Mixed Strategy Equilibria

Although the pure-strategy equilibria are interesting, we might consider *symmetric equilibria* more natural, given the symmetry of the game and its lack of identifying roles [Kreps, 1990]. In order to identify a symmetric equilibrium, we need in general to consider mixed strategies.

8.4.1 Two-Strategy Games

Let N be the total number of strategies in the profile (in our context, $N = 6$ without preemption, and $N = 5$ when we include a single fixed preemptive agent). Define DAP(X, j) as the DAP of strategy X (A or B) when j agents out of N play strategy A. If k agents each independently choose whether to play A with probability α (henceforth, "play α"), then the probability that exactly i will choose A is given by

$$\Pr(\alpha, i, k) = \binom{k}{i} \alpha^i (1-\alpha)^{k-i}.$$

Let $V(A, \alpha)$ denote the DAP of an agent playing A when the remaining agents play α:

$$V(A, \alpha) = \sum_{i=0}^{N-1} \Pr(\alpha, i, N-1)\text{DAP}(A, i+1).$$

Similarly, we define DAP values for playing B or α, respectively, in the setting where

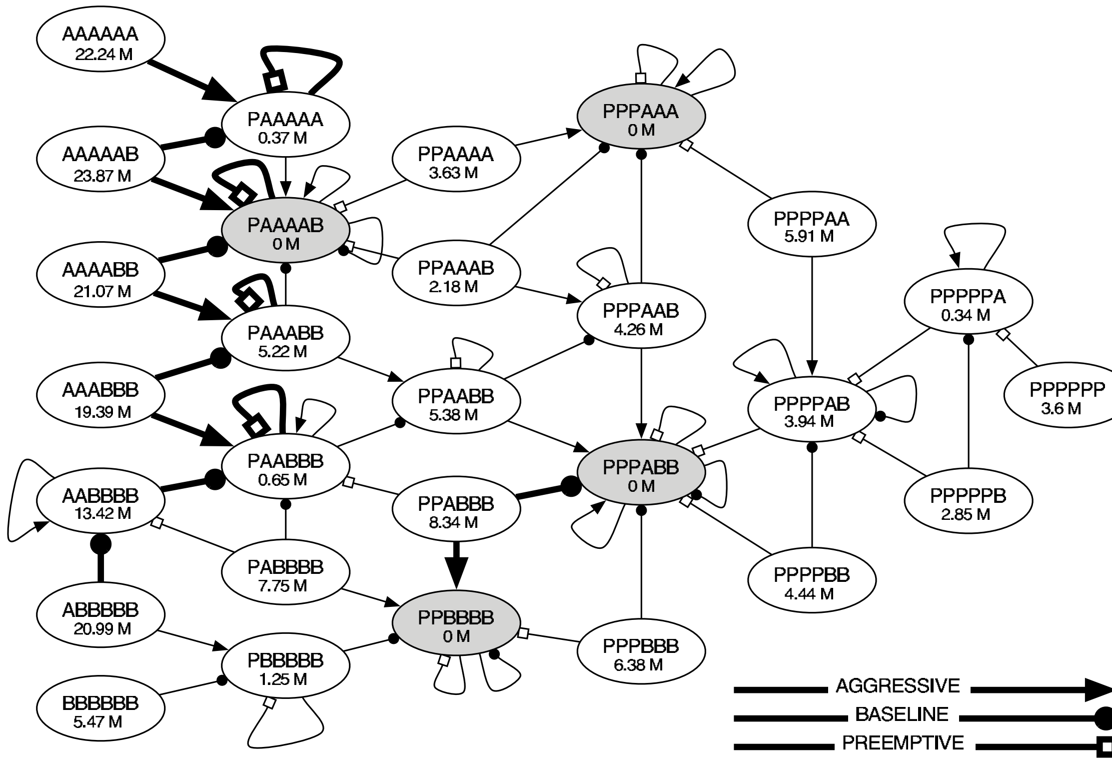

Figure 10: Profiles for the full three-strategy game, with arrows indicating a desire by the associated agent to change its strategy. Statistically significant comparisons are indicated by bold arrows. Values specified for each profile represent the minimum ϵ such that the profile constitutes an ϵ-Nash equilibrium.

others play α:

$$\begin{aligned} V(B,\alpha) &= \sum_{i=0}^{N-1} \Pr(\alpha,i,N-1)\text{DAP}(B,i), \\ V(\alpha,\alpha) &= \alpha V(A,\alpha) + (1-\alpha)V(B,\alpha). \end{aligned}$$

We plot these values of playing A, B, or α in response to α, for the two games, in Figure 11. A necessary and sufficient condition for a symmetric mixed-strategy equilibrium is

$$V(A,\alpha) = V(B,\alpha).$$

Therefore, we can identify such equilibria by the points in these figures where the curves intersect. For the game without preemption, we have a single symmetric mixed-strategy equilibrium, at $\alpha = 0.82$. When the preemptive agent is present, we find two symmetric mixed-strategy equilibria: $\alpha = 0.03$ and $\alpha = 0.99$.

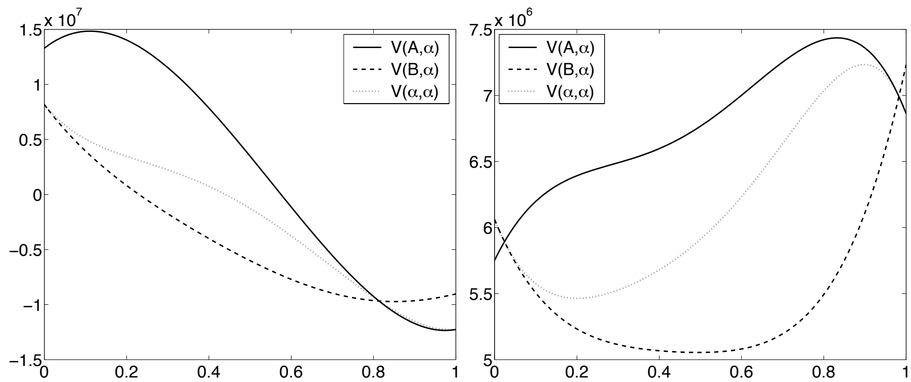

Figure 11: Response to mixed strategy α, for the unpreempted (left) and single-preemptor (right) games. Note that the payoff scale is an order of magnitude wider in the left graph.

The expected payoff for the equilibrium strategy (equal for A and B, by definition) of the game without preemption is a *loss* of \$9.59M. With a single preemptor, the two equilibria have expected payoffs of \$5.92M and \$7.01M, respectively. The preemptive agent itself also does well, earning profits of \$13.3M and \$9.99M in the respective equilibria.

Although we have no direct way to perform a statistical hypothesis test using demand-adjusted values, a conservative option is to compare the mean DAP scores using the variance of the raw averages. In this instance, DAP for the two preemptive equilibria exceed that of the non-preemptive equilibrium at p-values less than 0.0001.

Inspection of Figure 11 confirms our prior finding that preemption reduces the difference between A and B strategies. One way to quantify this is to identify an ϵ^* for each game such that *any* mixed strategy is a symmetric ϵ-Nash equilibrium at $\epsilon = \epsilon^*$. In our context, ϵ^* is therefore the maximum payoff difference between playing the best-response strategy, and playing α:

$$\epsilon^* = \max_{\alpha} \left(\max_{X \in \{A,B\}} V(X, \alpha) - V(\alpha, \alpha) \right).$$

For games without preemption, ϵ^* is \$10.6M. With preemption, ϵ^* is only \$0.97M. This provides a bound on how much it can matter to make the right choice about aggressiveness, given a symmetric set of other agents.

8.4.2 Full Three-Strategy Game

We were also able to derive a symmetric mixed-strategy equilibrium for the full three-strategy game, using replicator dynamics [Schuster and Sigmund, 1983]. In equilibrium, agents play A with probability 0.23, B with probability 0.19, and P with 0.58. The expected payoff for this mixed strategy is \$5.78M. This is not quite as good as the environment allowing only a single preemptor, but of course much better than the unpreempted situation.

9 Procurement Strategy in TAC-04 and Beyond

Based on the 2003 experience, the TAC/SCM game designers revised the game rules for TAC-04, with the primary intent of reducing the attraction of aggressive day-0 procurement. The modified supplier pricing rule imposed premiums for large orders, and the new storage fee added a proportional cost for holding component inventory. Although these may have had some deterrent effect, they were apparently not sufficient to prevent aggressive day-0 procurement in the 2004 tournament. They did, however, effectively rule out preemptive remedies of the sort that Deep Maize provided in TAC-03. In consequence, the tournament games exhibited major procurement imbalances, and surprisingly volatile profits given the much steadier demand process introduced for 2004. Further analysis will be required to characterize agents' procurement strategies in TAC-04, and explain the full strategic implication of the rule changes.

In order to remove day-0 procurement as an overriding issue once and for all, the designers for TAC-05 [Collins et al., 2004] are completely overhauling the method by which suppliers allocate offers. We look forward to investigating the other strategic trading issues that will inevitably come to the fore in the 2005 TAC/SCM game.

10 Conclusion

The TAC supply-chain game presented automated trading agents (and their designers) with a challenging strategic problem. Embedded within a highly-dimensional stochastic environment was a pivotal strategic decision about initial procurement of components. Our reading of the game rules and observation of the preliminary rounds suggested to us that the entrant field was headed toward a self-destructive, mutually unprofitable equilibrium of chronic oversupply. Our agent, Deep Maize, introduced a preemptive strategy designed to neutralize aggressive procurement. It worked. Not only did preemption improve Deep Maize's profitability, it improved profitability for the whole field. Whereas it is perhaps counterintuitive that actions designed to prevent others from achieving their goals actually helps them, strategic analysis explains how that can be the case.

Investigating strategic behavior in the context of a research competition has several distinct advantages. First, the game is designed by someone other than the investigator, avoiding the kinds of bias that often doom research projects to success. Second, the entry pool is uncontrolled, and so we may encounter unanticipated behavior of individual agents and aggregates. Third, the games are complex, avoiding many of the biases following from the need to preserve analytical or computational tractability. Fourth, the environment model is precisely specified and repeatable, thus subject to controlled experimentation. We have exploited all of these features in our study, in the process developing a repertoire of methods for empirical game-theoretic analysis, which we expect to prove useful for a range of problems.

There is no doubt that this form of study also has several limitations, for example in justifying generalizations beyond the particular environment studied. Nevertheless, we believe that the methods developed here provide a useful complement to the kinds of (a priori) stylized modeling most often pursued in game-theoretic analysis, and to the non-strategic analyses typically applied to simulation environments.

Acknowledgments

Thanks to the TAC/SCM organizers and participants, and to anonymous reviewers for constructive comments. At the University of Michigan, Deep Maize was designed and implemented with the additional help of Shih-Fen Cheng, Thede Loder, Kevin O'Malley, and Matthew Rudary. Daniel Reeves assisted our equilibrium analyses. This research was supported in part by NSF grant IIS-0205435.

References

Olivier Armantier, Jean-Pierre Florens, and Jean-Francois Richard. Empirical game-theoretic models: Constrained equilibrium and simulations. Technical report, State University of New York at Stonybrook, 2000.

Raghu Arunachalam, Joakim Eriksson, Niclas Finne, Sverker Janson, and Norman Sadeh. The TAC supply chain management game. Technical report, Swedish Institute of Computer Science, 2003. Draft Version 0.62.

Raghu Arunachalam and Norman M. Sadeh. The supply chain trading agent competition. *Electronic Commerce Research and Applications*, to appear.

Michael Benisch, Amy Greenwald, Ioanna Grypari, Roger Lederman, Victor Naroditskiy, and Michael Tschantz. Botticelli: A supply chain management agent designed to optimize under uncertainty. *SIGecom Exchanges*, 4(3):29–37, 2004.

Shih-Fen Cheng, Daniel M. Reeves, Yevgeniy Vorobeychik, and Michael P. Wellman. Notes on equilibria in symmetric games. In *AAMAS-04 Workshop on Game-Theoretic and Decision-Theoretic Agents*, New York, 2004.

John Collins, Raghu Arunachalam, Norman Sadeh, Joakim Eriksson, Niclas Finne, and Sverker Janson. The supply chain management game for the 2005 trading agent competition. Technical report, Carnegie Mellon University, 2004.

Erik Dahlgren. PackaTAC: A conservative trading agent. Master's thesis, Lund University, 2003.

Rui Dong, Terry Tai, Wilfred Yeung, and David C. Parkes. HarTAC: The Harvard TAC SCM '03 agent. In *AAMAS-04 Workshop on Trading Agent Design and Analysis*, New York, 2004.

Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification*. Wiley-Interscience, second edition, 2000.

Drew Fudenberg and Jean Tirole. *Game Theory*. MIT Press, 1991.

Herbert Gintis. *Game Theory Evolving*. Princeton University Press, 2000.

Philipp W. Keller, Félix-Olivier Duguay, and Doina Precup. RedAgent: Winner of TAC SCM 2003. *SIGecom Exchanges*, 4(3):1–8, 2004.

Jeffrey O. Kephart and Amy R. Greenwald. Shopbot economics. *Autonomous Agents and Multiagent Systems*, 5:255–287, 2002.

Jeffrey O. Kephart, James E. Hanson, and Jakka Sairamesh. Price and niche wars in a free-market economy of software agents. *Artificial Life*, 4:1–23, 1998.

Wolfgang Ketter, Elena Kryzhnyaya, Steven Damer, Colin McMillen, Amrudin Agovic, John Collins, and Maria Gini. Design and analysis of the MinneTAC-03 supply-chain trading agent. Technical Report 04-016, University of Minnesota, Dept of Computer Science and Engineering, 2004.

Christopher Kiekintveld, Michael P. Wellman, Satinder Singh, Joshua Estelle, Yevgeniy Vorobeychik, Vishal Soni, and Matthew Rudary. Distributed feedback control for decision making on supply chains. In *Fourteenth International Conference on Automated Planning and Scheduling*, pages 384–392, Whistler, BC, 2004a.

Christopher Kiekintveld, Michael P. Wellman, Satinder Singh, and Vishal Soni. Value-driven procurement in the TAC supply chain game. *SIGecom Exchanges*, 4(3):9–18, 2004b.

David M. Kreps. *Game Theory and Economic Modelling*. Oxford University Press, 1990.

Pierre L'Ecuyer. Efficiency improvement and variance reduction. In *Twenty-Sixth Winter Simulation Conference*, pages 122–132, Orlando, FL, 1994.

David Pardoe and Peter Stone. TacTex-03: A supply chain management agent. *SIGecom Exchanges*, 4(3):19–28, 2004.

Daniel M. Reeves, Michael P. Wellman, Jeffrey K. MacKie-Mason, and Anna Osepayshvili. Exploring bidding strategies for market-based scheduling. *Decision Support Systems*, to appear.

Sheldon M. Ross. *Simulation.* Academic Press, third edition, 2002.

Norman Sadeh, Raghu Arunachalam, Joakim Eriksson, Niclas Finne, and Sverker Janson. TAC-03: A supply-chain trading competition. *AI Magazine*, 24(1):92–94, 2003.

P. Schuster and K. Sigmund. Replicator dynamics. *Journal of Theoretical Biology*, 100:533–538, 1983.

Peter Stone. Multiagent competitions and research: Lessons from RoboCup and TAC. In *Sixth RoboCup International Symposium*, Fukuoka, Japan, 2002.

William E. Walsh, Rajarshi Das, Gerald Tesauro, and Jeffrey O. Kephart. Analyzing complex strategic interactions in multi-agent systems. In *AAAI-02 Workshop on Game-Theoretic and Decision-Theoretic Agents*, Edmonton, 2002.

William E. Walsh, David Parkes, and Rajarshi Das. Choosing samples to compute heuristic-strategy Nash equilibrium. In *AAMAS-03 Workshop on Agent-Mediated Electronic Commerce*, Melbourne, 2003.

Michael P. Wellman, Shih-Fen Cheng, Daniel M. Reeves, and Kevin M. Lochner. Trading agents competing: Performance, progress, and market effectiveness. *IEEE Intelligent Systems*, 18(6):48–53, 2003.

Dongmo Zhang, Kanghua Zhao, Chia-Ming Liang, Gonelur Begum Huq, and Tze-Haw Huang. Strategic trading agents via market modelling. *SIGecom Exchanges*, 4(3): 46–55, 2004.