Characterizing Effective Auction Mechanisms: Insights from the 2007 TAC Market Design Competition

Jinzhong Niu, Kai Cai
Simon Parsons
City University of New York
{jniu,kcai}@gc.cuny.edu
parsons@sci.brooklyn.cuny.edu

Enrico Gerding
School of Electronics and Computer Science
University of Southampton
Southampton SO17 1BJ, UK
eg@ecs.soton.ac.uk

Peter McBurney
Computer Science
University of Liverpool
Liverpool L69 7ZF UK
mcburney@liverpool.ac.uk

ABSTRACT

This paper analyzes the entrants to the 2007 TAC Market Design competition. It presents a classification of the entries to the competition, and uses this classification to compare these entries. The paper also attempts to relate market dynamics to the auction rules adopted by these entries and their adaptive strategies via a set of post-tournament experiments. Based on this analysis, the paper speculates about the design of effective auction mechanisms, both in the setting of this competition and in the more general case.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence: *Multiagent systems*

General Terms

Design, Economics, Experimentation, Measurement, Performance

Keywords

Double auction, mechanism design, trading agent competition

1. INTRODUCTION

This paper is concerned with the Trading Agent Competition (TAC) Market Design competition, a competition known as CAT. A CAT game consists of a set of agents. Each of these is either a buyer, a seller, or a specialist. Each specialist operates and sets the rules for a single exchange market, a double auction, and buyers and sellers — collectively called *traders* — trade in one of the available markets. Buyers and sellers make offers to trade, known as *shouts*, and specialists identify compatible traders, and then *clear* the markets. In the CAT competition, the traders are provided by the game organizers, and use standard trading strategies from the literature. While entrants know what strategies may be used, they are not told the precise makeup of the trader population. Specialists, and hence the rules of the markets, are designed by the entrants.[1] A typical CAT game consists of a CAT server and several CAT clients, which may be traders or specialists. CAT clients do not talk to each other directly; instead they connect to the CAT server through sockets and the server responds to messages from clients and forwards information if needed. A plain-text-based protocol called CATP, similar to HTTP, regulates the communication between the CAT server and the clients.

A CAT game lasts a certain number of *day*s, each day consists of *round*s, and each round lasts a certain number of *tick*s, or milliseconds. Trading is only permitted during rounds, and hence during a day. After a day closes, information on the profit made by each specialist and the number of traders registered with it are disclosed. This allows specialists to change their market rules, adapting these rules to improve their competitiveness. Between days traders may change the specialist that they trade with, and they migrate to specialists that allow more profitable trades [11].

Specialists make a profit by charging traders. They are allowed to charge traders a fee for registering to trade with them on a given day, for placing a shout, for obtaining information on the shouts made by other traders, for making a transaction, and they may charge a fraction of the bid/ask spread at a transaction (which we call the *profit fee*). In the first competition, in 2007, specialists were rated by a combination of the profit they made on specific *assessment days*, the market share they obtained on those days, and the success rate of transactions on those days.[2]

2. COMPONENTS OF SPECIALISTS

A specialist may adopt various auction rules. JCAT, the software platform that supports the CAT games [7], provides a reference implementation of a parameterizable specialist that can be easily configured and extended to use policies regulating different aspects of an auction. This section briefly describes a classification of those aspects that we have derived from the policies provided by JCAT and those used by specialists in the 2007 tournament. This classification is an extension of the parametric model of [14]. Section 3 relates these policies to the CAT 2007 finalists.

2.1 Matching policy

Matching policies define how a market matches shouts made by traders. *Equilibrium matching* (ME) is the most commonly used matching policy [9, 13]. The offers made by traders form the *reported demand and supply*, which is usually different from the *underlying demand and supply*, and are determined by traders' private values and unknown to the specialist, since traders are assumed to be profit-seeking and make offers deviating from their private values. ME clears the market at the *reported* equilibrium price and matches intra-marginal asks (offers to sell) with intra-marginal

[1] This is the reverse of the other TAC games, hence the name CAT.

[2] More information may be found at `www.marketbasedcontrol.com` and in [4].

Cite as: Title, Author(s), *Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, Padgham, Parkes, Müller and Parsons (eds.), May, 12-16., 2008, Estoril, Portugal, pp. XXX-XXX.

bids (offers to buy) — with an intersecting demand and supply, the shouts on the left of the intersection (the equilibrium point) and their traders are called *intra-marginal* since they can be matched and make profit, while those on the right are called *extra-marginal*. It is worth mentioning that a shout, or a trader, that appears to be intra-marginal or extra-marginal in the reported demand and supply may not be so in the underlying demand and supply. *Max-volume matching* (MV) aims to increase transaction volume based on the observation that a high intra-marginal bid can match with a lower extra-marginal ask, though with a profit loss for the buyer.

2.2 Quote policy

Quote policies determine the quotes issued by markets. Typical quotes are the ask and bid quotes, which respectively specify the upper bound for asks and the lower bound for bids that may be placed in a *quote-driven* market. *Two-sided quoting*[3] (QT) defines the ask quote as the minimum of the lowest tentatively matchable bid and lowest unmatchable ask, and defines the bid quote as the the maximum of the highest tentatively matchable ask and highest unmatchable bid. *One-sided quoting* (QO) is similar to QT, but considers only the standing shouts closest to the reported equilibrium price from the unmatchable side. When the market is cleared continuously (see below), QO is identical to QT, but otherwise forms a possibly looser restriction on placing shouts.

2.3 Shout accepting policy

Shout accepting policies judge whether a shout made by a trader should be permitted in the market. *Always accepting* (AA) accepts any shout. *Quote-beating accepting* (AQ) allows only those shouts more competitive than the corresponding market quote. This has been commonly used in both experimental settings and real stock markets, and is sometimes called the "New York Stock Exchange rule" since that market adopts it. *Equilibrium-beating accepting* (AE) learns an estimate of the equilibrium price based on the past transaction prices in a sliding window, and requires bids to be higher than the estimate and asks to be lower. This policy was suggested in [10] and found to be effective in reducing transaction price fluctuation and increasing allocative efficiency in markets populated with ZI-C traders [6]. *Self-beating accepting* (AS) accepts all first-time shouts but only allows a trader to modify its standing shout with a more competitive price. *Transaction-based accepting* (AT) tracks the most recently matched asks and bids, and uses the lowest matched bid and the highest matched ask to restrict the shouts to be accepted. In a clearing house (CH) [3], the two bounds are expected to be close to the estimate of equilibrium price in AE, while in a continuous double auction (CDA), AT may produce much looser restriction since extra-marginal shouts may steal a deal. *History-based accepting* (AH) is derived from the GD trading strategy [5]. GD computes how likely a given offer is to be matched, based on the history of previous shouts, and AH uses this to accept only shouts that will be matched with probability no lower than a specified threshold.

2.4 Clearing condition

Clearing conditions define when to clear the market and execute transactions between matched asks and bids. *Continuous clearing* (CC) attempts to clear the market whenever a new shout is placed. *Round clearing* (CR) clears the market after all traders have submitted their shouts.

2.5 Pricing policy

A pricing policy is responsible for determining transaction prices for matched ask-bid pairs. The decision making may involve only the prices of the matched ask and bid, or more information including market quotes. *Discriminatory k-pricing* (PD) sets the transaction price of a matched ask-bid pair at some point in the interval between their prices. The parameter $k \in [0, 1]$ controls which point is used and usually takes value 0.5 to avoid a bias in favor of buyers or sellers. *Uniform k-pricing* (PU) is similar to PD, but sets the transaction prices for all matched ask-bid pairs at same point between the ask quote and the bid quote. PU cannot be used with MV because the price intervals of some matched ask-bid pairs do not cover the spread between the ask quote and the bid quote. *n-pricing* (PN) was introduced in [10], and sets the transaction price as the average of the latest n pairs of matched asks and bids. If the average falls out of the price interval between the ask and bid to be matched, the nearest end of the interval is used. This policy can help reduce transaction price fluctuation and has little impact on allocative efficiency. *Side-biased pricing* (PB) is basically PD with k set to split the profit in favor of the side on which fewer shouts exist. Thus the more that asks outnumber bids in the current market, the closer k is set to 0.

2.6 Charging policy

Charging policies determine how charges are imposed by a specialist. *Fixed charging* (GF) sets charges at a specified fixed level. *Bait-and-switch charging* (GB) makes a specialist cut its charges until it captures a certain market share, and then slowly increases charges to increase profit. It will adjust its charges downward again if its market share drops below a certain level. *Charge-cutting charging* (GC) sets the charges by scaling down the lowest charges of markets imposed on the previous day. This is based on the observation that traders all prefer markets with lower charges. *Learn-or-lure-fast charging* (GL) adapts charges towards some target following the scheme used by the ZIP trading strategy [1].

3. SPECIALISTS IN THE 2007 FINALS

The first CAT competition was held in conjunction with AAAI in July 2007. Table 1 lists the finalists in descending order of their final rankings[4] and identifies the auction rules we inferred from the programs of the CAT 2007 competition final (held in the TAC repository) against the policies we described in Section 2. All specialists for which we have data fit into the generic double auction mechanism framework introduced above and Table 1.

We found that most specialists use ME to clear markets at the equilibrium price. `IAMwildCAT` and `Mertacor` are the only two attempting to match competitive intra-marginal shouts with extra-marginal shouts close to the equilibrium point in order to obtain high transaction success rates. QT, familiar from classic CDAs and CHs, is a popular quote policy, but its effectiveness is bound to the matching policy that is used with it since different matching algorithms, such as ME and MV, can generate significantly varying quotes. Furthermore, quote policies only affect the performance of the specialists when AQ is used as an accepting policy.

Specialists use a wide range of shout accepting policies, which reflects the importance of this aspect in performing well in CAT games. In contrast, only `CrocodileAgent` and `Mertacor` use a clearing condition that isn't one of the standard policies provided in JCAT. Since JCAT ensures that specialists impose uniform charges

[3]The name follows [9] since either quote depends on information on both the ask side and the bid side.

[4]Due to technical problems, two teams, `TacTex` and `MANX`, were not able to participate in all the games. Some teams were banned from parts of some games — `PSUCAT` and `Havana` for exceeding reconnection limits, and `CrocodileAgent`, `Havana`, `MANX`, `PSUCAT`, `TacTex`, and `jackaroo` for invalid fees.

Table 1: Comparison between the CAT 2007 finalists.

specialists	matching	quote	accepting	clearing	pricing	charging
IAMwildCAT	ME+MV*	$QT + QO + \Delta$	AQ+AE+AS+Δ	CR	PB*	Δ
PSUCAT	ME	(QT)	AE*	CC	PB*	Δ
CrocodileAgent	ME	$(QT + QO^*)$	AE	CR*	PN*+PB*	GL*
jackaroo	ME	QT*	AQ	CC	PN	GC*+Δ
Havana	ME*	QT	AQ	CC	PD	Δ
PersianCat	ME*	(QT)	AT*+Δ	CC	PD	GF*+Δ
PhantAgent						
Mertacor	MV*	(QT)	AE*	CR*	PB*	Δ
TacTex	ME	(QT)	AA	CR	PD	GB*+GC*
MANX	ME	QT	AQ	CR	PD	GC*+GL*

XX^* denotes a policy that can be viewed as a modified or improved XX; Δ stands for some mechanism that cannot be related to any policy in Section 2; (XX) represents a quote policy that is defined by the specialist but has no effect on its behavior due to its adoption of some non-AQ accepting policy; and $XX + YY$ means some combination of XX and YY. Blanks are left due to lack of information — PhantAgent is not in the TAC repository.

on all traders registered with it on a trading day, it is not possible to attract specific traders by levying differential charges. However, about half the entrants managed to bias their pricing policy to promote the quality of their trader population.

Entrants seem to have contributed more effort to charging policies than to any other aspect of auction mechanisms. Table 2 in particular compares:

1. How charges are updated over time.

   Some specialists *adapt* their charges while others *directly calculate* the charges that they expect to bring a certain payoff without explicitly considering how they charge currently. A third choice is to combine the two approaches by setting charges that move *gradually* from the current level to the target level.

2. Whether different types of charges are treated differently.

   About half of the specialists impose only or mainly registration fees and charges on profits. TacTex charges only shout fees. All the three specialists without a bias towards a certain kind of fee — CrocodileAgent, Havana and MANX — adapt charges without using any heuristic knowledge of the fee types.

3. Whether traders are identified and treated differentially.

   Only IAMwildCAT tracks individual traders and records information on them.

4. How much profit a trader and/or a specialist can make on average.

   IAMwildCAT and jackaroo are the only two specialists that lay down a road map for achieving some desired or target profit. IAMwildCAT is the only one that tracks the absolute value of the daily overall profit of specialists, which, when small, can be exploited by the specialist to obtain a fairly high share of the profit without imposing massive fees.

5. Whether a specialist learns from the history of charges and performances of its own and/or the other specialists.

   It is a common practice among the specialists for fees to be set based on information about their competitors' charges and performances, though the lengths of history monitored vary from only the previous day, to a sliding multi-day window, to the full game history.

6. Whether a specialist tries to lure traders by charging less in the early stage of a game (*start effect*) and/or imposes higher charges when the game is about to end (*deadline effect*).

   Most specialists feature start and deadline effects, taking advantage of a definitive game duration and traders exploring widely at the beginning of a CAT game.

The characterization of the specialists in Table 2 may help establish the relationship between the features of auction rules and their performances, and guide appropriate modification of an auction mechanism to achieve desirable behavior.

4. EXPERIMENTS WITH THE FINALISTS

To further examine the strategies of the specialists that participated in the CAT 2007 competition, we ran a series of games with the same setup as in the 2007 finals.

4.1 Experimental setup

Every game in our experiment ran for 500 trading days with 10 rounds per day and 1 second per round. The trader population comprised 180 ZIP traders [1], 180 RE traders [2], 20 ZI-C traders [6], and 20 GD traders [5]. Buyers and sellers were evenly split in each trader sub-population. The private values of all the traders were independently drawn from a uniform distribution between 50 and 150, and each trader was allowed to buy or sell up to 3 commodities per day. The specialists in our games included all 8 specialists released on the TAC web site's agent repository. The same scoring criteria were used as in the tournament [4] but, unlike the tournament, all the game days were assessed. The results and plots shown in the following sections were averaged over a total of 10 games.

[5] PSUCAT however does identify traders to adjust parameters in its pricing policy.

Table 2: Comparison between the charging policies of the CAT 2007 finalists.

specialist	**fee update**	**fee type bias**	**trader id**	**profitability**		**fee history**		**score history**		**start effect**	**deadline effect**
				traders	**specialists**	**self**	**others**	**self**	**others**		
`IAMwildCAT`	⇒∘	✓	✓	✓	✓	❙	❙	◇	◇	✓	✓
`PSUCAT`	∘⇒	✓	✗[5]	✗	✗	❙	❙	❙	❙	✓	✓
`CrocodileAgent`	∘⇒∘	✗	✗	✗	✗	◇	◇	◇	◇	✓	✗
`jackaroo`	⇒∘	✓	✗	✓	✗	❙	❙	❙	❙	✓	✓
`Havana`	⇒∘	✗	✗	✗	✗	◇	◇	◇	◇	✓	✗
`PersianCat`	∘⇒	✓	✗	✗	✗	❙	✗	❙	✗	✓	✓
`PhantAgent`											
`Mertacor`	⇒∘	✓	✗	✗	✗	❙	◀	❙	◀	✓	✓
`TacTex`	∘⇒	✓	✗	✗	✗	❙	❙	◀	◀	✓	✗
`MANX`	⇒∘	✗	✗	✗	✗	❙	❙	◇	◇	✗	✗

✓ has this feature ✗ does not have this feature ◇ sliding window ❙ single day
◀ full history ∘⇒ adapting ⇒∘ direct calculation ∘⇒∘ gradual learning

specialist	**score**	**std. dev.**
`IAMwildCAT`	240.22	2.82
`PSUCAT`	209.26	12.01
`CrocodileAgent`	179.64	17.53
`jackaroo`	182.80	24.30
`PersianCat`	128.82	5.57
`Mertacor`	100.11	8.57
`TacTex`	166.66	8.99
`MANX`	140.09	31.03

Table 3: The scores of specialists in our experiments. `Havana` relies on the CPLEX library, and since at the time of writing we do not have a licence for CPLEX, we were unable to include it in our experiments.

To obtain a clearer view, plots were smoothed out with each datum being the average of a 10-day sliding window around it.

4.2 Winners

The results of our experiments broadly agree with the rankings in the tournament [12]. The 2007 CAT champion, `IAMwildCAT`, still wins in our experiments and `PSUCAT`, which placed second in the competition, comes second as shown in Table 3. The only changes in ranking are due to `TacTex` and `MANX` increasing their scores since they could participate in every game. Figure 1 shows the daily components of the scores and Figure 2 shows some of the daily charges made by the specialists.

4.3 Trader migration

The competition among specialists is reflected directly by the migration of intra-marginal traders and extra-marginal traders. Traders migrate based on estimates of expected profits, where the estimate for a given specialist is based on past experience with that specialist. Generally speaking, the more intra-marginal traders and the fewer extra-marginal traders in a market, the more potential profit there is, and the easier it is to make transactions and achieve a high transaction success rate. To measure the balance of intra-marginal and extra-marginal demand and supply, we introduce the *marginal coefficient*, β. For demand,

$$\beta_D = \frac{D_i}{D_i + D_e} \tag{1}$$

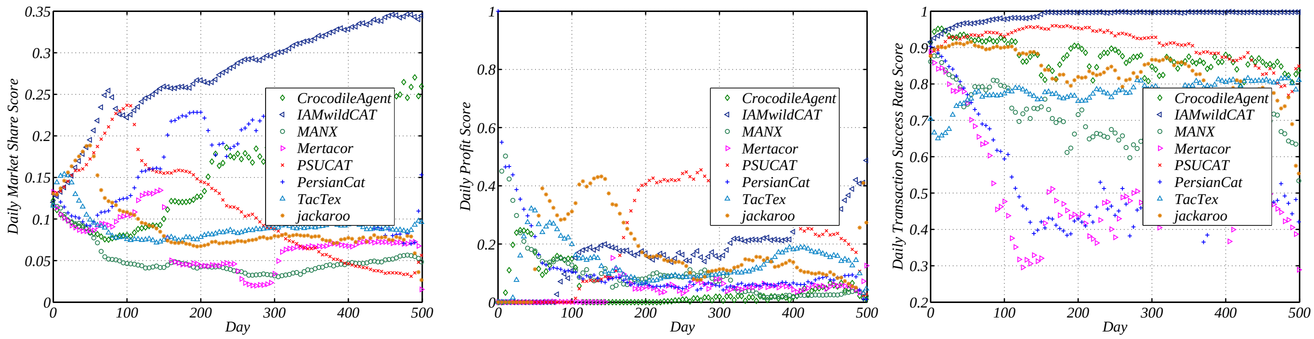
where D_i is the intra-marginal demand — the equilibrium — and D_e is the extra-marginal demand. The marginal coefficient for supply, β_S, can be defined similarly. β_D varies between 0 and 1. A value of 0 indicates that all the buyers in the market are extra-marginal while 1 indicates all the buyers are intra-marginal. Figure 3(a) shows the daily value of β_D in the individual markets managed by the specialists. Since β_D provides no information on the absolute equilibrium quantity or profit, Figure 3(b) gives the daily equilibrium profits in the markets.

As Figure 3(a) shows, $\beta_D \approx 0.5$ in all the markets when the game starts. Then the β_D of `IAMwildCAT`, `TacTex`, and `PSUCAT` increases while that of `CrocodileAgent`, `PersianCat`, and `Mertacor` decreases. Since a falling β indicates losing intra-marginal traders and/or gaining extra-marginal traders, these changes indicate that intra-marginal traders and extra-marginal traders have different preferences over the different markets.

Intra-marginal traders seem to be sensitive to matching policies and charges, especially charges on profit. However, they seem to be relatively insensitive to other charges so long as they can still profit from trades. Figure 3(a) shows that the β_D of `Mertacor`, `PersianCat`, and `CrocodileAgent` decreases significantly at the beginning of the game and remains low all the way through the game. However these decreases occur for different reasons.

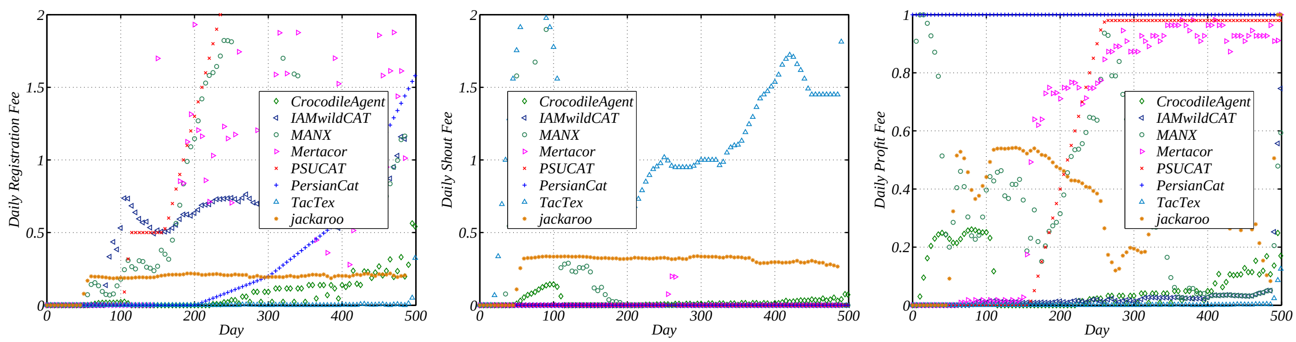
The low allocative efficiency of `Mertacor` means a great portion of the potential social welfare is not achieved, suggesting an inefficient matching policy[6]. A close examination of `Mertacor`'s mechanism found that its `MV`-like matching policy strategically executes extra-marginal trades so as to increase its transaction success rate, but this leads to much lower profit for the intra-marginal

[6] During the CAT 2007 competition, `TacTex` and some others announced invalid fees on some trading days, causing them to be banned from the games for a certain period. This is equivalent to the use of a very inefficient matching policy. Our experiments have rounded their fees into the valid ranges and avoided banning the specialists.



(a) Market share (b) Profit (c) Transaction success rate

Figure 1: Scores of specialists in our experiments



(a) Registration fee. (b) Shout fee. (c) Charge on profit.

Figure 2: Daily fees charged by specialists in our experiments.

traders involved in those trades. In addition, `Mertacor` mistakenly disregards the unmatched shouts every time after the market is cleared. This will then make the traders that made these shouts unable to either improve their standing shouts or place new ones since the game server believes they still have active shouts. Some of these traders may possibly be intra-marginal traders, therefore causing unrealized intra-marginal trades. These two issues provide sufficient 'excuse' for intra-marginal traders to flee.

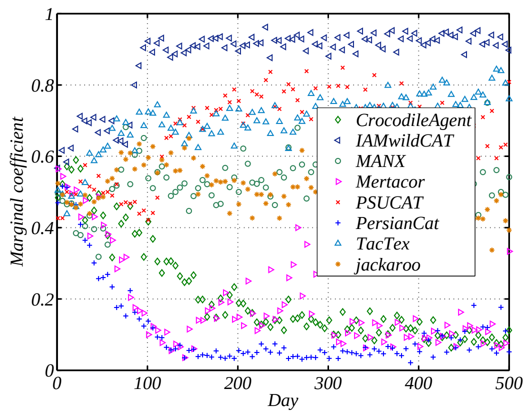
`PersianCat` and `CrocodileAgent` lose traders due to imposing high profit charges. `PersianCat` charges 100% on profit for the whole game and this drives β_D down very quickly. `CrocodileAgent` levies a lower fee than `PersianCat` and therefore has a modestly decreasing β_D. The decrease of β_D in `PSUCAT` and `jackaroo` starting from days 250–300 follows an aggressive increase in the profit fee.

The rest of the specialists have much higher β_D despite their use of similar policies. `IAMwildCAT`, for instance, though adopting a version of MV, refrains from using it in the early rounds of a day, which usually are sufficient to realize most intra-marginal trades. `MANX`, though levying a high, yet volatile, profit fee, also levies other fees without bias considerations, which together scare away both extra-marginal traders and intra-marginal traders at approximately the same pace. Its β_D therefore zigzags around 0.5. The three specialists that obtain a β_D higher than 0.6 during the most time of the game, `IAMwildCAT`, `PSUCAT`, and `TacTex`, all produce allocative efficiency higher than 85%, again suggesting the importance of matching policies in keeping a high-quality trader population.

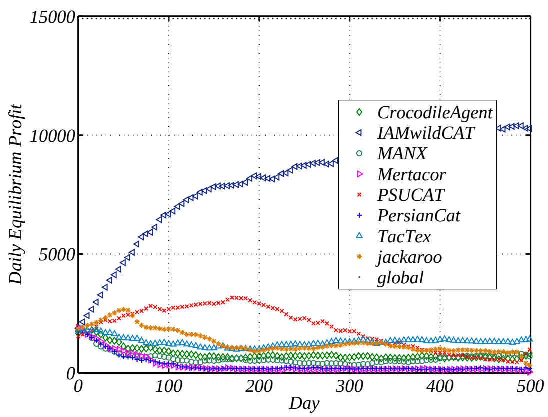
Registration fees appear to help filter out extra-marginal traders, and information fees have the same effect on GD and ZIP traders (which require such information). `IAMwildCAT` and `jackaroo` consistently impose one or both of these fees. As a result, the number of extra-marginal traders in those markets falls the most.

Shout fees also affect extra-marginal traders, but the degree of the effect depends on the shout accepting policy used. If the accepting policy is a strong filter and extra-marginal traders have little chance to place shouts, they can avoid losing money due to charges and thus are indifferent to shout charges. Their staying with a specialist therefore does not harm the market's transaction success rate, and on the contrary, only adds to its market share. `TacTex`, uniquely among the specialists, charges only shout fees and consistently does so all the way through the game, as shown in Figure 2(b). This policy together with its AA accepting policy — the weakest one possible — causes the extra-marginal traders to leave quickly.

`Mertacor` managed to draw a large number of extra-marginal traders during the first 200 days, due to its free-market policy. Its policy change, starting to charge heavily on registration as in Figure 2(a), explains why it loses almost all its extra-marginal traders shortly afterwards and its β increases significantly around day 200. Actually, higher registration fees in `PSUCAT` after day 150 and `PersianCat` after day 200, are both accompanied with a loss of market share in extra-marginal traders. `CrocodileAgent` increases its registration fee as well around day 200 but the modestly increased



(a) Daily marginal coefficient of demand, β_D



(b) Daily equilibrium profit

Figure 3: Properties of daily equilibria in individual markets.

fee is still lower than those charged by most of other specialists, therefore it is still popular among extra-marginal traders.

In conclusion, extra-marginal traders, as expected, flee from those markets with high registration fees and information fees (also high shout fees as in `TacTex`) to other markets, while intra-marginal traders migrate from markets with high profit fees and inefficient matching policies to those that do not have high charges and realize the most potential social welfare.

4.4 Learning and adaptation in specialists

The numbers of traders registered daily with specialists, profit made in markets, and their daily charges are all accessible to specialists via CATP. This makes it possible for specialists to learn and adapt their own policies. The transaction success rates however are unavailable unless a specialist is willing to obtain shout and transaction information direct from other specialists, paying any necessary fees. Specialists' payments for this purpose are not observable during the games.

Though specialists may adapt various types of auction policies, changes in charging policy are more obvious than other aspects from the data collected. `MANX` copies the charges of the leading markets in terms of profit share and market share combined, producing the most scattered charges among the specialists through the games. Looking at its charges gives us an approximate pattern of adaption of the other markets.

1. At the opening stage, `PersianCat` charges the most (though only profit fees) while most of the others are free markets.

2. `TacTex` then starts to impose shout fees, but its payoff and winning position is not sustainable. Its market share declines significantly as seen in Figure 1(a) around day 20.

3. Around day 50, `jackaroo` begins to impose all types of fees heavily, and similarly to `TacTex`, `jackaroo`'s market share decreases. Figure 1(a) shows that before day 50, the free `jackaroo` market attracts more and more traders, but after that, traders flee quickly first and then slowly. Figures 3(a), 3(b), and 1(a) further indicate that intra-marginal traders are more sensitive and flee faster than extra-marginal traders immediately after day 50, causing a plunge in market share immediately after day 50 and an increasing β between days 50 and 100. Around day 100, β starts to drop again, suggesting extra-marginal traders leave at a slower and slower pace and intra-marginal traders still flee gradually if not even faster.

4. From around day 85, `IAMwildCAT` disregards its free-market policy and turns to charge registration fees, as shown in Figure 2(a), which scares away extra-marginal traders, and Figure 3(a) shows a significantly faster increase of β. `PSUCAT` afterwards does the same thing and causes an increasing β between days 100 and 150.[7]

5. `IAMwildCAT` and `jackaroo`, are designed to take advantage of the known length of games. One after another, they increase their charges to much higher levels and make huge profits during the last days of the games, though JCAT has taken measures to avoid traders going bankrupt in this situation and disregards any due charge beyond the capabilities of traders. The huge daily profits obtained, however, did not greatly increase their final scores since the scoring mechanism adopted by CAT normalizes profits before scoring.

The comparison between the charges of `MANX`, which copies charges, and those of the specialists mentioned above show clearly which have adapted their policies and become the daily front-runners at each point.

`IAMwildCAT` exhibits stable performance according to almost all criteria and is worth further investigation. Profit share is the most sensitive metric since fee changes may immediately and dramatically cause the relative profit shares to go up or down. In Figure 1(b), `TacTex`, `jackaroo`, and `PSUCAT`, one after another, increase their charges and claim big profit shares. However every subsequent increase leads to an apparent profit share drop for the previous front-runner, including what `IAMwildCAT` does to `PSUCAT` by increasing its profit charge gradually as shown in Figure 2(c).[8] Despite this common theme, `IAMwildCAT` is to a great extent immune to the changes of other specialists' charges in terms of its profit share. This should be attributed to its target-oriented charging policy and the direct calculation of fees to achieve a certain target profit. `Mertacor` takes a similar approach, but its sub-optimal calculation method and other problematic auction rules prevent the approach from working well.

[7]The y axis in Figure 2(a) has an upper bound of 2, and does not show the constant registration charges of 10 by `PSUCAT` in the second half of the game. This aims to obtain a better general view, avoiding the curves of other specialists (usually below 2) being squeezed together and approaching the x axis. The even higher charges by the specialists near the end of the game, due to the same reason, are not shown in Figures 2(a)-2(c).

[8]The increase of shout fees in `TacTex` around day 300 may also play a role in lowering `PSUCAT`'s profit share.

5. DISCUSSION

Here we extract some general guidance for market design from the analysis above.

5.1 ME vs. MV

If a high transaction success rate is desirable, then specialists have to explicitly take this into account, for example by matching intra-marginal and extra-marginal shouts, just as MV does. However, caution should be exercised when using an MV-like policy. MV may cause intra-marginal traders to lose profits and in a competitive situation may lead them to prefer non-MV markets. This is exactly what happened to `Mertacor`. In addition, the extra-marginal trades may lower market efficiency. `IAMwildCAT`'s matching policy is a mixture of ME and a MV-like policy. It uses the former in the first rounds of a day and the latter in the rest of the day. Figure 1(c) shows that `IAMwildCAT` obtains high transaction success rates, very close or equal to 100%, after day 150 when the specialist starts to use the MV-like policy for more rounds in a day. As a consequence, `IAMwildCAT`'s efficiency has a striking 5% drop. Unlike `Mertacor`, `IAMwildCAT` did not show a loss of intra-marginal traders when it did this. This is because most of the intra-marginal traders traded in the early rounds of each day — when the MV-like policy was used, most of the traders still shouting were extra-marginal traders, few shouts made by these traders can pass the specialist's shout accepting policy, and these limited extra-marginal shouts did no great harm to the remaining intra-marginal traders.

Since traders are profit-seeking, MV-like policies can actually increase market allocative efficiency in some cases. For instance, a greedy intra-marginal trader may make an extra-marginal shout, which, when ME is used, will not be matched and therefore add to the number of unrealized intra-marginal trades. When MV is used, this extra-marginal shout can be matched by an intra-marginal trader, and the efficiency loss can thus be reduced or avoided.

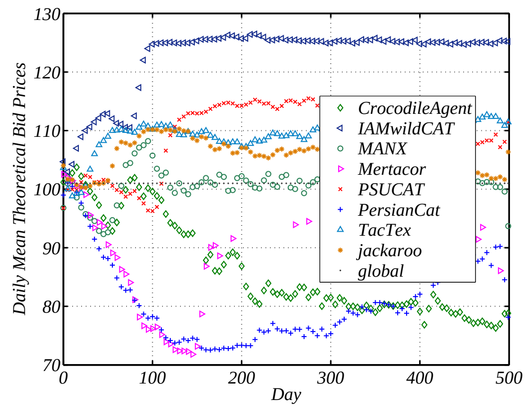
5.2 Open vs. closed shout accepting

Shout accepting policies have a direct impact on the effectiveness of other auction rules. An open shout accepting policy places a heavy burden on the matching policy. When the matching policy is also ineffective, intra-marginal traders fail to profit and tend to leave. In contrast, if the shout accepting policy filters out most extra-marginal shouts, a simple matching policy can work well.

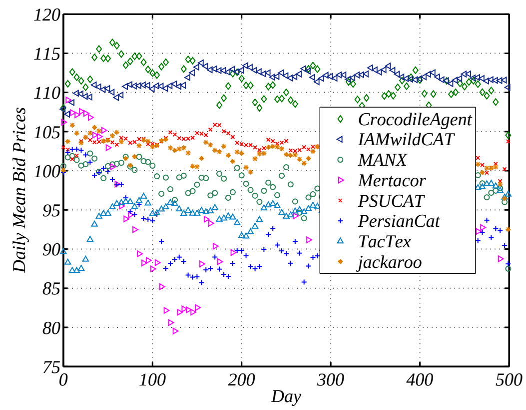
For example, `CrocodileAgent` and `PersianCat` have similar trader populations in terms of their competitiveness, as shown in Figures 4(a) and 3(a), and both use the ME matching policy. However, they produce significantly different shout sets as shown in Figure 4(b) and transaction success rates as illustrated in Figure 1(c). This is due to the AE accepting policy in `CrocodileAgent`, which is much more effective than the policy in `PersianCat`.

In addition, AQ, the common accepting policy, may leave the door wide-open at the start of a day. In CAT games, shouts automatically expire at the end of a day. This resets the market quotes in AQ and loses valuable information from the previous day on the underlying demand and supply schedules, which do not usually change dramatically over days. This may explain why `jackaroo` and `MANX`, the two AQ markets, with higher mean theoretical demand prices in Figure 4(a) than those in `CrocodileAgent`, produce bid sets with lower mean prices as in Figure 4(b) and lower transaction success rates.

We believe a good accepting policy in the current CAT game setting should be able to reflect the collective properties of traders and carry this knowledge from day to day, as the history-based policy AH does. We expect that most specialists would be better off when using AH. `PSUCAT`'s customized AE is another potential policy. The



(a) Underlying demand, as calculated from trader private values.



(b) Reported demand, as calculated from shouts.

Figure 4: Daily mean demand prices.

mean theoretical demand price in the `PSUCAT` market jumps around day 100 in Figure 4(a), and β_D follows in Figure 3(a), but this did not cause the mean bid price to climb as well, indicating the effectiveness of its shout accepting policy, which successfully prevented extra-marginal traders placing shouts.

5.3 Market share vs. profits

In CAT games it is common for specialists to find that increasing fees will boost profits but gradually lead to loss of market share. If market share falls too low, the profit return cannot be sustained. In contrast, low charges help gain market shares but harm profits. However, if a charging policy is properly designed, it may keep both measures at suitably high levels. Imposing small, flat, fees after a game has been running for a while, may not have much negative effect on market shares if the good reputation of a specialist has been established and the traders continue to make profit that is much higher than the fees. In this way, on the basis of a big market share, small fees may still bring a considerable amount of profit. `IAMwildCAT` demonstrates this.

Bias towards different types of fees in charging policies can also benefit specialists. For example, `IAMwildCAT` and `PSUCAT` use registration fees to drive extra-marginal traders away, make it easy for the trader population to find partners, and obtain high transaction success rates. However as discussed in Section 4.3, a powerful shout accepting policy may make this unnecessary or even harmful, since it may filter out most extra-marginal shouts and avoid their negative effect on transaction success rates. With a strong accepting policy and without charges on registration and information, a market actually becomes a free place for extra-marginal traders to stay. If other markets impose these charges, these traders are sure to be willing to stay with this free market and boost market share.

5.4 Targeted vs. non-targeted charges

Specialists in the tournament adapt their daily charges differently (Table 2). Some do this by setting specific performance targets, determining these targets from estimates of the expected actions of other specialists, while others increase or decrease their current charges without setting targets or modeling the effect of the changes. `IAMwildCAT`, for instance, determines a reasonable portion of the profit it desires to make via registration fees, and calculates its registration fee and profit fee by taking into consideration the average profit a trader has been able to make in its market. In contrast, parameter values and charge levels of most other specialists are decided rather arbitrarily. As a result, `IAMwildCAT` has a stable performance in the face of changes by other specialists.

Several specialists are reactive, copying the fees that other, well-performing, specialists charge. `MANX` in particular does this. This approach is problematic for two reasons. First, it is usually based on a short-term assessment and may not optimize the long-term outcome. Second, copying a winning specialist may not be a winning strategy. The effect of fees is closely related to other auction rules of specialists and the properties of their trader population at that moment [11]. `MANX`'s follow-the-leader approach demonstrates impressive performance during the early part of a game when the trader populations in all individual markets are quite similar. However it fails to lead to a similar outcome after traders have diverged to prefer different markets.

6. SUMMARY

This paper provides an analysis of the entrants in the 2007 TAC Market Design competition. We believe that it makes three main contributions to the literature of electronic markets.

First, this paper provides a more extensive assessment of the performance of the entrants to the 2007 competition than was possible in the competition itself. Each game runs for around 8 hours, and given the technical problems experienced by both organizers and competitors, this meant that it was only possible to run two games during the three days of the competition, and not all of these involved all the competitors. Running more games and including all the competitors gives more definitive results, and confirms the dominance of `IAMwildCAT`.

Second, this paper provides the first classification of the strategies used by 2007 Market Design competition entrants, and the first comparison of the effects of these strategies in a rigorous, systematic experiment. While there are many more experiments to be run before we fully understand the comparative strengths of the strategies, we believe that these aspects of the paper will be of help to future entrants in the competition.

Third, the paper provides a discussion of the implications of the design of the various components of double auction mechanisms, in particular the interaction between the component policies, and their effect on auction performance. We hope that this part of the paper will help to guide future research on the design of double auctions, not least in suggesting new market designs that involve new combinations of component policies.

Of course, there are limits on what this analysis tells us. Since the results are likely, as are all market games of this complexity, to depend heavily on the population of participants, the conclusions we draw here are only valid in the context of the specialist and trader populations we experimented with. To obtain more robust results, we need to carry out the kind of empirical game-theoretic analysis presented in [8], and we are currently working on this.

Acknowledgments

The authors are grateful for financial support received from the National Science Foundation under grant NSF IIS-0329037 and from the EPSRC under grant GR/T10657/01. Our grateful thanks go to the entrants to the 2007 TAC CAT tournament for releasing the binaries of their specialists, to Haizheng Zhang for information on `PSUCAT`, and to Elizabeth Sklar for advice on the design of JCAT.

7. REFERENCES

[1] D. Cliff and J. Bruten. Minimal-intelligence agents for bargaining behaviours in market-based environments. Technical report, Hewlett-Packard Research Laboratories, Bristol, England, 1997.

[2] I. Erev and A. E. Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review*, 88(4):848–881, September 1998.

[3] D. Friedman and J. Rust, editors. *The Double Auction Market: Institutions, Theories and Evidence*. Santa Fe Institute Studies in the Sciences of Complexity. Perseus Publishing, 1993.

[4] E. Gerding, P. McBurney, J. Niu, S. Parsons, and S. Phelps. Overview of CAT: A market design competition. Technical Report ULCS-07-006, Department of Computer Science, University of Liverpool, Liverpool, UK, 2007. Version 1.1.

[5] S. Gjerstad and J. Dickhaut. Price formation in double auctions. *Games and Economic Behavior*, 22:1–29, 1998.

[6] D. K. Gode and S. Sunder. Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy*, 101(1):119–137, 1993.

[7] `http://jcat.sourceforge.net/`.

[8] P. R. Jordan, C. Kiekintveld, and M. P. Wellman. Empirical game-theoretic analysis of the TAC supply chain game. In *Proceedings of the Sixth International Joint Conference on Autonomous Agents and Multiagent Systems*, Honolulu, Hawaii, 2007.

[9] K. A. McCabe, S. J. Rassenti, and V. L. Smith. Designing a uniform price double auction. In Friedman and Rust [3], chapter 11, pages 307–332.

[10] J. Niu, K. Cai, S. Parsons, and E. Sklar. Reducing price fluctuation in continuous double auctions through pricing policy and shout improvement rule. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*, Hakodate, Japan, 2006.

[11] J. Niu, K. Cai, S. Parsons, and E. Sklar. Some preliminary results on competition between markets for automated traders. In *Workshop on Trading Agent Design and Analysis*, Vancouver, Canada, 2007.

[12] `http://www.marketbasedcontrol.com/blog/`. Entry for July 26th 2007.

[13] P. R. Wurman, W. E. Walsh, and M. P. Wellman. Flexible double auctions for electronic commerce: Theory and implementation. *Decision Support Systems*, 1998.

[14] P. R. Wurman, M. P. Wellman, and W. E. Walsh. A parametrization of the auction design space. *Games and Economic Behavior*, 35:304–338, 2001.