Evolutionary Pricing Games

Bernhard von Stengel*

September 6, 2022
— draft, not for distribution —

Abstract

This text describes a *research idea* related to evolutionary tournaments in a dynamic pricing setting. It originated in an experimental tournament from 1991 conducted by Claudia Keser, then a PhD student in Bonn supervised by Reinhard Selten. I participated in that tournament with a very successful strategy, which was fun to devise. The challenge was that participants knew the game (which has aspects of competition and cooperation) but not the actions likely chosen by their opponents. Interestingly, the good performance of my strategy was less visible in an evolutionary setting, where more cooperative strategies were more prevalent.

The open research questions (which will need to be made more specific) concern how one should design (and evaluate) evolutionary successful strategies in this context, and how and if this success can be achieved with learning algorithms, which may be tailored to the model, or general.

1 Background

Keser (1992) conducted and analyzed a tournament competition between game theorists. Her work belongs to experimental economics. The goal was to find out how people behave in a relatively complex game-theoretic scenario. A certain dynamic duopoly game was played in two ways: among subjects who pairwise play against each other in a laboratory, and – which is the topic of this note – according to the "strategy method" (see Brandts and Charness, 2011). In the present setting, subjects submit an explicit strategy in the form of a flowchart. The

*Department of Mathematics, London School of Economics, London WC2A 2AE, United Kingdom. Email: b.von-stengel@lse.ac.uk

strategies were solicited from a forum of academic game theorists, who submitted their strategies. These strategies were played against each other in a tournament, and ranked according to their cumulative payoffs. The results were sent back to the participants who could then submit a new strategy, with a new second tournament.

A third part of the study is reported in Keser (1992, pages 97–110), called an "evolutionary tournament", based on the strategies of the second tournament. This part of the study had not been announced to the participants, and was not conducted further (nor published in the journal article Keser, 1993). I explain possibilities and issues of possibly continuing this part of the study.

2 The oligopoly model

The model is a classic oligopoly game analyzed by Selten (1965) with the concept of subgame perfect equilibrium (SPE). The game is played over a fixed number of rounds T or, with discounted payoffs, infinitely often (the infinite case is analyzed in part II of Selten's paper, not cited). There are n firms. Each firm i has a certain "market strength", called a *demand potential* D_i, that determines its number of $D_i - p_i$ of sold units of a good when setting a price p_i (firm i's decision in each period), with a profit of $p_i - c_i$ per unit for the firm's production cost c_i.

The myopic monopoly profit maximizes $(D_i - p_i)(p_i - c_i)$ when $p_i = (D_i + c_i)/2$. This is the optimal action of the firm in the final round when played over T rounds.

The model has the following feature of *demand inertia*: At the start of the T periods, each firm i has the same demand potential D_i. After each round, the demand potential is changed for the next round in proportion to the difference of the firm's price p_i and the average price of all firms, where firms with lower prices gain and firms with higher prices lose demand potential (the sum of demand potentials stays the same).

The SPE found by Selten, found by backward induction from the action in the final round parameterized by the demand potentials, has the following features:

- in each period, the optimal action (price p_i) of a firm i depends only on its *own* current demand potential D_i (where p_i increases with D_i) and its cost c_i (where p_i decreases with c_i), and the discount factor if used. That is, the other players are anonymous (this makes sense from the game rules).
- prices quickly converge to a stable value, with a short "end effect" in the last periods where prices rise.

- prices are relatively low and favor firms with lower costs, with substantially lower profits compared to myopic monopoly price-setting.

3 The duopoly model in the tournament

In the tournament used for the strategy method (as well as in the round-by-round price setting among game playing subjects), the game was between two players, as a duopoly. The low-cost firm had cost $c_1 = 57$ and the high-cost firm cost $c_2 = 71$, and a starting demand potential of $D_1 = D_2 = 200$. (Also a discount factor of 1 percent per round; this was apparently not of relevance for any partipants.) The number of rounds was 25. Players have perfect information about the past but not about the chosen price of the other player in the current round. The demand potential is adjusted so that firm i gains from its opponent one "customer" (increase of D_i) for each price unit that it is cheaper than the opponent in the current time period t. That is, with the demand potentials and prices now superscribed with the current time period t:

$$\begin{aligned} D_1^{t+1} &= D_1^t + \tfrac{1}{2}(p_2^t - p_1^t)\,, \\ D_2^{t+1} &= D_2^t + \tfrac{1}{2}(p_1^t - p_2^t)\,. \end{aligned} \tag{1}$$

(Selten showed that the analysis of the SPE is valid as long as the combined proportionality factor, here $1 = \frac{1}{2} + \frac{1}{2}$, of the "inertia adjustment" does not exceed 2.)

The aggressive SPE prescribes low prices, near about 95 in the middle stretch of rounds, with cumulative profits of 137k and 61k for the low- and high-cost firm, and a demand split for them of about 222 and 178, respectively. The corresponding cooperative myopic monopoly profits are 156k and 109k, a demand split of 207 and 193, and a much higher stabilizing price of 132.

Tournament participants submitted a separate strategy (in the form of a flowchart for computing the current price) for both the low-cost and the high-cost player. Different costs were chosen because this makes it harder to identify a focal point of "cooperative" behavior which would be given, by symmetry, if costs were the same (namely an equal split of the demand potential).

The tournament had 45 participants in the first round and 34 participants in the second round. Keser (1992) analyzes and classifies the different types of strategies, according to their cooperativeness and vulnerability, with various statistical analyses.

4 My strategy

I was one of the participants and had the best-scoring high-cost strategy (as participant number 10) in the second round (among 34 participants), and a near-best low-cost strategy. The challenge was that there was only guesswork about how any opponent would behave. My strategy

- aimed for a cooperative split of the demand potentials of 207 and 193 as in the myopic outcome;
- would choose a moderately lower price (minus 7) compared to myopic monopoly pricing as this would lose little in a repeated setting compared to "full cooperation", but would always gain extra "customers" for the next round;
- used, uniquely among all strategies, a *predicted* price of the opponent for the next period with a very simple linear model based on the opponent's demand potential (in effect, mimicking my own strategy, so that it would quickly stabilize when playing against myself);
- based on this prediction immediately "fought back" to claw back my demand potential if that fell below 207 resp. 193;
- as a feature of the second tournament after I observed better strategies in the first tournament: exploit "suckers" that do not claw back their demand potential by keeping a low price even when having an own "undeserved" high demand potential, because a low price and many customers would still give high profits; this was achieved by the very simple device of never setting a price above 125 (which is 7 less than the "cooperative" price of 132). See Keser (1993, page 81).

It was clearly of help understanding the game mechanics well.

5 More details on game mechanics

The following has not yet been implemented as a strategy, but may lead to models of opponents.

Suppose we start with the prices that are stable in the myopic monopoly setting: Recall that the per-unit cost of the Low-cost firm is 57, of the High-cost firm 71. The "myopic" demand potential split is $D_1 = 207$, $D_2 = 193$ with an optimal price of 132 for both, with a profit of $(207 - 132)(132 - 57) = 75^2 = 5625$ for the low-cost firm and a profit of $(193 - 132)(132 - 71) = 61^2 = 3721$ for the high-cost firm per time period.

Now suppose Low-cost reduces this price by x. Its current profits reduce from 75^2 to $(75+x)(75-x) = 75^2 - x^2$, a quadratic loss that is very small when x is small. The *gain* is strong: in the next period, the firm sells $x/2$ additional units, because D_1 has increased by $x/2$.

For example, if $x = 10$ then its profit per sold unit is 65, which in the next period (with the same price) means its profits increase by $5 \times 65 = 325$, compared to the loss of 100 in the current period. That's why I chose $x = 7$ (a current loss of 49) against a gain of $\frac{7}{2} \times (75-7) = \frac{7}{2} \times 68 = 238$ in the next period. Of course, this is cancelled if the opponent does the same, so this is mildly (but only mildly) competitive.

Another consideration is how to beat an *imitation* (or "tit-for-tat") strategy that *copies* the opponent's price from the previous period (and starts out being "nice"). My intuition is that the firm facing this tit-for-tat strategy should steal as many customers as possible from in the first round with a low price (rather than doing so gradually, to be proved) and then keep the same price so that the demand potential is constant from then on.

Here is an analysis against an imitation strategy: Suppose the firm's current demand potential is $D = 200$, its cost is c, and its opponent chooses a price of q in the first period. The firm chooses a single price p throughout, and ignores its payoff in the first round. Its demand potential from then on will be $D^* = D + (q-p)/2$, and its future profits per period will be $(D^* - p)(p - c)$, which is equal to $(D + q/2 - p/2 - p)(p-c)$ or $\frac{3}{2}(\frac{2}{3}D + \frac{1}{3}q - p)(p-c)$, and maximized by $p = (\frac{2}{3}D + \frac{1}{3}q + c)/2$ (the midpoint between the zeros of the previous parabola in p). For example, if $c = 57$ and $D = D_1 = 200$ and $q = 131$ (a bit below the myopic price of $(200+71)/2$ for the high-cost opponent) then

$$p = (\frac{2 \times 200 + 131}{3} + 57)/2 = 117\,. \tag{2}$$

The firm's demand potential goes up to 207 and its future profits per period are $(207-117)(117-57) = 6300$.

I wonder if this is stable and the firm can still keep stealing slowly (and if so should maybe do so from the beginning, which above I conjectured *not* to be the right way to start).

However, the high-cost opponent may anticipate this action, that is, not be "nice" in the first period, and choose $q = 114.2$, which is the solution to $q = p$ in (2) above; in general, it is $q = \frac{3}{5}D + \frac{2}{5}c$. Then the low-cost firm's demand potential does not change at all ($D = D^*$) and its profits will be a mere $(200-q)(q-57) \approx 4908$ and those of the high-cost player about 3707 per period (which might be lower with an improved strategy of the low-cost player, to be investigated).

All these are observations about how to optimize a firm's unilateral payoffs and not fully game-theoretic. They may be useful examples of modeling other agents.

6 The evolutionary tournament: a surprise

In the (unannounced) evolutionary tournament, Keser (1992) put the mutual results among the 34 strategies (for both firms) of the second tournament into a 34×34 bimatrix game. Starting from equal populations shares of 1/34, this was then put into a model of discrete replicator dynamics, where the population share of each firm increased or decreased according to its performance in the current population mix. In addition, new "mutant" entrants where added at the rate of about 10^{-7} per round so that no strategy ever died out, even if it performed poorly. This dynamics was then run over 100,000 rounds and the resulting proportions recorded, which oscillated around a small number of strategies that survived with significant fractions.

The main fractions where two cooperative strategies with fractions of about 0.6 and 0.3. My own "best" high-cost strategy against a uniform distribution of 34 opponents had a fraction of 0.05. A mixed Nash equilibrium with support size 4 on both sides was found that included these strategies that had high fractions in the populations, and mixed strategy probabilities near the average fractions in the dynamic round (which were not stable but had an apparent orbit).

A possible explanation why my own "clever" strategy had the low population share was possibly my exploitation of "suckers" that led to a different treatment of my opponents. In a mixed Nash equilibrium, all strategies in the support need to have equal payoffs. This constraint would not tolerate my strategy with a high mixed-strategy probability, which is determined by the opponent payoffs, because it treated opponents so differently.

If there had been a third round aimed at doing well in the evolutionary tournament, I would have re-designed my strategies with this goal in mind.

7 Issues of conducting a third "evolutionary" tournament

Keser (1992, page 110) writes: "We might suspect that the subjects would develop their strategies further if they played more rounds. Some subjects submitted for the second tournament strategies with a structure completely different from their strategies of the first tournament. In the evolutionary tournament, however, we restrict the strategy set to the strategies participating in the second tournament. New strategies cannot evolve during the evolutionary process. Ignoring this problem, we might consider our result as a hint where the trend might go to if we repeated our tournament strategies more often: cautiously cooperative to moderately aggressive strategies."

There are a number of questions here:

a. Methodological – how would one conduct further rounds? Ask participants to submit flowcharts each time? How often?

b. What is the aim of the study?
   - (presumably) find out how subjects approach this problem;
   - (questionable) model a market of pairwise interaction; it is not clear what a model of random pairwise interactions represents: e.g., in each village/neighbourhood/street, at most two stores can open, and how should they compete? A more adequate model could be that of an oligopoly with $n = 34$ participants, but then the game is played very differently, with a fight for a "global" market share. But this is an altogether different game. Maybe accuracy of the model should not be an issue for a beginning theoretical investigation.

c. I think it is interesting to even address the problem of how to do well in an evolutionary setting (see end of Section 6 above). That is, I would like to extend my own strategy in this respect. But what are my opponents, and how should I test my strategy?

d. In the quote by Keser at the beginning of this section, *evolve* is a crucial word. How could one parameterize a strategy to run the 25-round game (in the pairwise interaction setting) to *learn* how to play it? Maybe with neural net?

e. Maybe it is worth abstracting from the duopoly setting altogether, and just call it a "base game" that has certain constraints of acting cooperatively or noncooperatively, with a set of payoff pairs for the two players (these constraints would have to be chosen interestingly enough, and not just model a Prisoner's dilemma interaction).

8 What is optimal behavior?

Continuing from d. in the previous section, suppose we want to "learn" how a firm plays *well* in this game, say with a neural net that computes current prices with some inputs such as current demand potential and (possibly the whole) history.

The question is *how to measure this*. In the first two Keser tournaments, it was the performance against the uniform distribution of opponents.

The third evolutionary tournament used a replicator dynamics (with small-probability random entrants from the existing pool) and looked at the surviving strategies. This probably depends on the starting distribution; e.g., playing the SPE strategies against each other should be a pure and probably strict equilibrium and thus an ESS (evolutionary stable strategy).

My original question (as a participant in a third round) was how to *design* (rather than learn) a strategy that does well in the evolutionary setting, i.e., survive with a high frequency.

An analytical derivation (such as "treat your opponents equally") is probably difficult, but one could *compare* how a good strategy against a uniform distribution does in an evolutionary setting, or some other distribution, maybe with that distribution chosen as a starting point. (For a general game, I am sure initially successful strategies can quickly die out.)

9 Simplifying / specifying the base game

The 25-period duopoly game is kind of complicated, but could be simplified as follows:

- (Suggested by Simina Brânzei.) Drop the end effect and run the game forever, with the discount factor $\delta < 1$ as the probability that the game continues to the next round, and expected running time $1/(1-\delta)$ (and long runs exponentially unlikely). Simulations would need to be done multiple times to get good averages.
- Identify the characteristics of the game via a parameter that specifies a range of "cooperative" versus "competitive" behavior. The result is probably a kind of continuous Prisoner's Dilemma, let's call it a "one-parameter PD". It may be useful to distill this as the essence of the duopoly game. Or maybe we need two parameters? Probably this makes quite a difference.

Related is what I find attractive about the tournament so far: A participant *designs* their own strategy, without knowledge of what opponent to expect. So while the rules of the dynamic "base game" (the 25-round duopoly game) are clear, the available strategies are so vast that one doesn't just pick some row of a big bimatrix game, but designs and explores the possible strategies.

In this setting, rather than just studying some evolutionary dynamics, the player has *agency* in playing the game by picking a strategy. This is a bit vague but I think combining agency (rational analysis) and evolutionary adaption (no agency, just reproductive fitness) is an interesting, possibly new topic. The catchy title of such a study could be "Agency and evolutionary fitness". Suppose we have a one-parameter PD game where the agency consists in picking that parameter. Is this still interesting?

10 Related work

This is not a full study yet and related work is surely vast.

Of interest is Calvano, Calzolari, Denicolo, and Pastorello (2020) (and its summary on a blog at https://www.law.ox.ac.uk/business-law-blog/blog/2019/02/artificial-intelligence-algorithmic-pricing-and-collusion). They use a Bertrand model of competition (which is also price-setting, although its relationship to Selten's model of demand inertia will need to be examined). Their findings show that agents using Q-learning learn how to become semi-collusive based on their own learning, without communication.

Our proposed approach in Section 8 based on Selten's model has the "design" aspect (actively designing a good strategy) and its comparison with Keser (1992) as an additional feature, but may confirm these findings.

References

Brandts, J. and G. Charness (2011). The strategy versus the direct-response method: a first survey of experimental comparisons. *Experimental Economics* 14(3), 375–398.

Calvano, E., G. Calzolari, V. Denicolo, and S. Pastorello (2020). Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review* 110(10), 3267–3297.

Keser, C. (1992). *Experimental Duopoly Markets with Demand Inertia: Game-Playing Experiments and the Strategy Method*, volume 391 of *Lecture Notes in Economics and Mathematical Systems*. Springer Verlag, Berlin.

Keser, C. (1993). Some results of experimental duopoly markets with demand inertia. *The Journal of Industrial Economics* 41(2), 133–151.

Selten, R. (1965). Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit: Teil I: Bestimmung des dynamischen Preisgleichgewichts. *Zeitschrift für die gesamte Staatswissenschaft/Journal of Institutional and Theoretical Economics* 121(2), 301–324.