Competition à la Cournot with bounded rationality

Màxim Sandiumenge i Boy

Universitat Pompeu Fabra

Advisor: Francesco Cerigioni

Code: EME08

June 13, 2019

Abstract

Rules of thumb, deviations from the standard objective functions, or explicit cognitive bounds have been introduced in Industrial Organization to tackle irrationality. At the same time, a considerable literature has been extending level-k since Nagel (1995), but little has been done on its application outside experimental settings. To address this we define a model of boundedly rationality and we apply it to a Cournot competition game, finding that the accuracy of anticipation and the distribution of level 0 players' action play a critical role in explaining which firms become larger in dynamic settings.

Contents

1 Introduction

Cournot and Bertrand explained the consequences of competition under some simplifying assumptions, each one considering a different choice variable of the firm. Hotelling (1929) introduced the idea that for a given commodity, "some customers buy from one seller, some from another seller, in spite of moderate differences of price". These authors obtained different stable outcomes for different assumptions by considering rational players and common knowledge of rationality through finding points where all optimal conditions are satisfied at the same time. They found what later on was modelled by Nash (1951) in a formal and generalised manner. All these contributions are the bases of what is known about firm behaviour and specially in firm competition, but they rely on strong assumptions. Rational assumptions seem to be suitable when referring to big firms, or to equilibrium situations as it might be the case of small (maybe non-rational) firms that have been competing with each other for a long time and reach an equilibrium, but standard models have little to say about how competition becomes effective when a market starts from zero, and this analysis is not based on rational assumptions. They also have little to say about the facts explaining how some firms may become larger than others starting from a situation where all competitors have the same fundamentals.

In this paper we depart from Nash and we study how firms behave in a first period without assuming fully rational thinking but boundedly rational agents. We go through the level-k literature and propose a model as an alternative to the ones suggested by Nagel, Camerer and Chong. Then, we apply this model to a Cournot competition example in three different settings. We will first consider duopoly competition between two retailers. Secondly, we will assume an upstream monopoly and a downstream duopoly, the former charging uniform pricing to the laters. Finally, we will analyse the previous case but allowing the monopoly to price discriminate.

1.1 Level-k literature

Level k (LK) models (Nagel, 1995; Stahl and Wilson, 1994; Costa-Gomes et al., 2001; Costa-Gomes and Crawford, 2006), Cognitive hierarchy (CH) models (Camerer et al., 2004), and more recently the Generalized cognitive hierarchy (GCH) model (Chong et al., 2016) have been the most successful models in explaining out of equilibrium behaviours.

They start from the model proposed by Nagel. This defines an individual whose action is a random number (a_0) among the available closed interval $[\underline{a}, \bar{a}]$ as a level 0, an individual that best responds to his/her belief of all other agents being level 0 as level 1, an individual that thinks that all other agents are level 1 and best responds to that, as level 2 and so on, being able to go to each extra cognitive stage is defined as "depths of reasoning".

Nevertheless, they differ in the way in which individuals' mental process incorporates the cognitive process of other players when facing a game. While in level k models subjects of level k best respond uniquely to a population of level $k-1$, in cognitive hierarchy models, individuals of level k best respond to a boundedly rational heterogeneous population of levels between 0 and $k-1$. At the same time, GCH represents a generalization over CH since the former allows two parameters (α and β) that for certain values GCH is identical to CH. GCH represents an improvement over CH at two different levels. First, the parameter α permits to adjust the belief of the relative proportion of lower levels agents to which a level k would best respond, which implies a difference in the perception of the population to which best respond.[1] Second, they introduce a generalized behaviour for level 0 players. On the basis of the random behaviour rule used so far, they propose a rule for which level 0 players are β times more likely to choose strategies inside a set of preferred actions. This set is composed of actions that never yield minimum payoff for any possible action of the opponents. This makes level 0 agents less likely

---

[1] In the CH model, level k agent best responds to a population of lower levels with distribution equal to the normalized true proportion, i.e. level k's belief of the relative proportion of level h is equal to $\frac{f(h)}{\sum_{h'=0}^{k-1} f(h')}, \forall h < k$, while in GCH this is equal to $\frac{f(h)^{\alpha}}{\sum_{h'=0}^{k-1} f(h')^{\alpha}}, \forall h < k$.

to play dominated strategies relative to the fully random behaviour.

1.1.1 Degree-0 behaviour

The model does not use the rule for degree 0 players proposed in GCH since in the games that we focus on individuals are not likely to detect strategies that never yield minimum payoff.[2] As specified in Chong et al. (2016), there are two prevailing behaviour models for degree zero players. One of them assumes that level zero randomises over the set of possible actions, and another claims that degree zero chooses a salient action among this same set. Even though the latter is criticised due to its unpredictability, we claim that this rule is more likely to represent the cognitive process of a degree zero player.

In section 2 we define a model for boundedly rational individuals that reproduce the *cognitive path* that agents follow in these situations. Our model differs from the previous at two main stages. First, each individual starts its own cognitive process from his/her belief about degree 0 players, so we break up with the common knowledge of degree 0 players' rule. Second, we also introduce a *corrective term* that agents use to readjust their action once they finish their cognitive process. This is due to their awareness of their inability to perfectly forecast others' behaviour. Furthermore, the model does not use the rule for degree 0 players proposed in GCH since in the games that we focus on individuals are not likely to detect strategies that never yield minimum payoff. This structure given in GCH is plausible to be applied in less cognitively costly games. Contrary, we state that level 0 players are likely to start from focal points (probably anchors due to framing in the game). In section 3 and its corresponding subsections, we apply a simplification of the model previously defined to a quantity-setting game and we provide the main insights of the analysis.

---

[2] Here we use the terminology used by John Maynard Keynes in "The General Theory of Employment, Interest, and Money". Then we will use the terms level k and degree k indistinguishably.

2 Formal Definitions and Terminology

In this section we will introduce the main definitions and terminology of the model about bounded rationality decision making that this paper is proposing. We will give some basic and standard definitions that will be complemented by some newer concepts.

We shall start defining the space of games that we have in mind. These will be simultaneous games with N players, where each individual $i \in N$ chooses an action $a_i \in A_i \in \mathbb{R}$, and to each player there is an associated payoff function, p_i that maps the set of all n-tuples into the real numbers. We refer to this n-tuple as a sequence of N elements, each element corresponding to one player, in this case each element being the action submitted by each individual. Then, the game will be defined as $\{N, (A_i)_{i \in N}, (p_i)_{i \in N}\}$.

Being $\tilde{b}_{i,j}$ the belief of player i of player $j's$ degree thinking, let $\tilde{b}_i$ be the *belief tuple* of payer i. A n-tuple that has as elements the beliefs of player i about the degree thinking of player j, $\forall j \in N$, being $j = i$, his/her own awareness of his/her thinking level, the first element of the sequence. Then, $\tilde{b}_1 = (\tilde{b}_{1,1}, \tilde{b}_{1,2}, ..., \tilde{b}_{1,n-1}, \tilde{b}_{1,n})$

Let $\tilde{a}_{i,j}$ be the belief of player i about the belief that player j has of degree 0 players' action; and $\tilde{a}_i$ be the n-tuple that aggregates the beliefs of individual i, $\forall j \in N$, being $i = j$ the first element. Then $\tilde{a}_i$ has as elements the belief of player i about the beliefs of all players about degree 0 actions, i.e. $\tilde{a}_1 = (\tilde{a}_{1,1}, \tilde{a}_{1,2}, ..., \tilde{a}_{1,n-1}, \tilde{a}_{1,n})$.

At the same time, let $\dot{a} \in \dot{A}$ represent a n-tuple, in the space of all possible tuples, with the actual actions submitted by the n players.

The *payoff function*, p_i, associates to each element of the domain exactly one element of the codomain. Being the domain composed by $\dot{A}$ and the codomain by $\mathbb{R}$. Being $p_i(\dot{a}^{th}) \equiv p_i(a_1, a_2, ..., a_n)$ for some $\dot{a}^{th} \in \dot{A}$.

Let $\tilde{p}_i$ be a subjective payoff function (or expected payoff function) that has as inputs the personal perception of the cognitive (degree) distribution of the opponents (i.e. $w_i * \tilde{b}_i$), and $\tilde{a}_i$. Being w_i some undetermined, for the moment, subjective weighting function of $\tilde{b}_i$.

The subjective best response function of individual $i \in N$, degree $k \in \mathbb{W}$, is the correspondence that maps the set $\tilde{a}$ to the set of feasible actions, and it is defined as:

$$\pi_{i,k}(w_i * \tilde{b}_i, \tilde{a}_i) = \underset{a\prime_i \in A_i}{\operatorname{argmax}} \tilde{p}_i(w_i * \tilde{b}_i, \tilde{a}_i)$$

For us, the *cognitive path* will be the intellectually costly mental process that each individual $i \in N$ does silently with the final goal of completing the task. It is composed by his/her belief about what all players think degree-0 would play (i.e. $\tilde{a}_i$), his/her belief about the cognitive distribution of the population, and a *corrective term* in order to readjust the calculus upwards or downwards. It is the best response function $\pi_{i,k}$ plus the corrective term $\epsilon_{i,k}$. This new term permits to incorporate final adjustments given their own awareness of being incapable to perfectly forecast the outcome. This corrective term is understood as a randomisation among a set $\epsilon \subset \mathbb{R}$, that contains the zero, whose cardinality may depend on all the previous parameters. Therefore, the behavioural prediction for player i is $\Pi_{i,k} \equiv \pi_{i,k}(w_i * \tilde{b}_i, \tilde{a}_i) + \epsilon_{i,k}$.

Then, the aggregate behavioural prediction of the static game will be composed by the result of the N cognitive paths. This is, the aggregate behavioural prediction of the game defined by $\{N, (A_i)_{i \in N}, (p_i)_{i \in N}\}$ will be composed by $\Pi = (\Pi_{1,k}, ..., \Pi_{N,k})$.

3 Competition à la Cournot with bounded rationality

In this section we will analyse duopoly quantity-competition in different situations and with different underlying assumptions than when it was first treated by A. Cournot in 1838. We will first consider duopoly competition between two retailers (D_1 and D_2) with bounded rationality and zero costs. Second we will assume an upstream rational monopoly (U) and a downstream boundedly rational duopoly (D_1 and D_2), the former charging uniform pricing to the laters. Finally, we will treat the previous case but allowing the monopoly to price discriminate.

The following few assumptions will be common all through out the exercise. Retailers will face a market with the following linear demand: $P(\dot{Q}) = 100 - \dot{Q}$ being $\dot{Q}$ the sum of the actual quantities produced by the retailers. In our analysis we will assume that $\tilde{a}_{D1,D2} = \tilde{a}_{D2,D1}$ (from now on we will do a slight change in notation for the sake of clearness, using q as a has been used so far), basically ensuring that the action that both retailers think that a level 0 player would choose is the same. In addition, D_1 is level 1, D_2 is level 2 and thinks that D_1 is level 1 (using our previous notation: $\tilde{b}_1 = (\tilde{b}_{1,2} = 0)$ and $\tilde{b}_2 = (\tilde{b}_{2,1} = 1)$), and U is perfectly rational and able to forecast which level each retailer is since it observes the quantities they demand. During these examples we do not consider the corrective term previously specified.

3.1 Duopoly with bounded rationality

Following the usual analysis, in a quantity-setting game of this kind, the firms will observe the demand they face and will try to maximize their objective function with respect to their choice variable. The only difference here will be when stating the best response functions. Being D_1 a level 1, it will react to its belief of D_2 acting as a level 0. Similarly, D_2 will best respond to a level 1 firm. Then, the corresponding subjective best response functions will be the following

(notice they obviously intersect in the Nash quantity):

$$\pi_{D_1,1} = \frac{100 - \tilde{q}_{D1,D2}}{2} = \frac{100 - \tilde{q}}{2} \tag{1}$$

$$\pi_{D_2,2} = \frac{100 + \tilde{q}_{D2,D1}}{4} = \frac{100 + \tilde{q}}{4} \tag{2}$$

The best response function for firm D_1 is, as usual, decreasing in its belief of quantity introduced in the market by D_2. Then, the lower D_1 thinks the competition will be, the less it will produce. Let us call this effect *expected competition effect*, which in this case is negative. Contrary, the best response function of D_2 is obviously increasing in $\tilde{q}$, since the larger $\tilde{q}$ is, the lower the production of D_1 will be, which will keeping prices high. Therefore, D_2 will have incentives to expand a little bit its quantity to still apply an increase of price to and even larger amount of units; D_2 will face lower competition the larger $\tilde{q}$. This makes the actual total quantity produced to be decreasing in $\tilde{q}$, and therefore the price to be increasing in this variable.

$$\dot{Q} = \frac{300 - \tilde{q}}{4} \tag{3}$$

$$P(\dot{Q}) = \frac{100 + \tilde{q}}{4} \tag{4}$$

$$\Pi_{D_1} = \frac{1}{8}(100 + \tilde{q})(100 - \tilde{q}) \tag{5}$$

$$\Pi_{D_2} = \frac{1}{16}(100 + \tilde{q})^2 \tag{6}$$

This makes the profits for D_1 to be concave in $\tilde{q}$ with a maximum in $\tilde{q} = 0$, and the profits of D_2 to be convex in this variable with a maximum in $\tilde{q}$ being equal to the maximum number we may define as feasible.[3] This two profit functions are obviously equal in $\tilde{q} = q^{Nash}$. In particular, for $\tilde{q} > q^{Nash}$, $\Pi_{D_1} < \Pi_{D_2}$, level 1 players shrink their production, since expected

---

[3] This definition will be case-specific. In the current case we may think that even a degree 0 player would never produce more than 100 units.

competition for level 1 players is relatively high, and level 2 players take advantage of this by expanding quantities. Hence, the belief of level 1 players about large competition becomes true only because of their inability to either commit to produce larger quantities, or to go one step further in rational reasoning about others' actions. Contrary, for $\tilde{q} < q^{Nash}$, $\Pi_{D_1} > \Pi_{D_2}$, D_2 is harmed by the expectation of level 1 players about low competition, which makes D_1 to expand production and therefore, D_2 to reduce its.

This suggests that should we consider a dynamic game rather than a static one where players form new expectations by somehow using the previous actions of players, ceteris paribus, chances are that for any starting feasible $\tilde{q}$, in the medium run Nash quantities will be reached.[4] Once this happens, firms will do Nash profits indefinitely, everything else kept constant.

$$CS = \frac{1}{2}\left(\frac{300 - \tilde{q}}{4}\right)^2 \tag{7}$$

$$W = -\frac{1}{32}(500 + \tilde{q})(\tilde{q} - 300) \tag{8}$$

While consumer surplus is convex in the expectation of level 0 quantity, total welfare is concave. Both with a maximum in $\tilde{q} = 0$, where total quantity is maximum and price minimum. As it has been implicitly explained, this scenario mimics the Nash or Cournot case for $\tilde{q} = \frac{100}{3}$. Therefore, for this value this exercise predicts the same consumer surplus and total welfare that in the standard result. Nevertheless, both curves are larger than this result for $\tilde{q} < q^{Nash}$, since this triggers the total quantity to be larger.

For consumer surplus and total welfare it also happens that, in a dynamic game with this expectations generating rule, they tend to their Nash values.

---

[4] In fact, to ensure this, say that the rule of thumb for generating expectations is the following. In period $t = 1$ the static game is played, for $t > 1$, firm D_1 infers from $\dot{Q}$ (showed at the very end of the previous period) the actual production of firm D_2 in $t - 1$, and best responds to that; firm D_2 knows this and best responds.

Fig. 1: Duopoly with bounded rationality
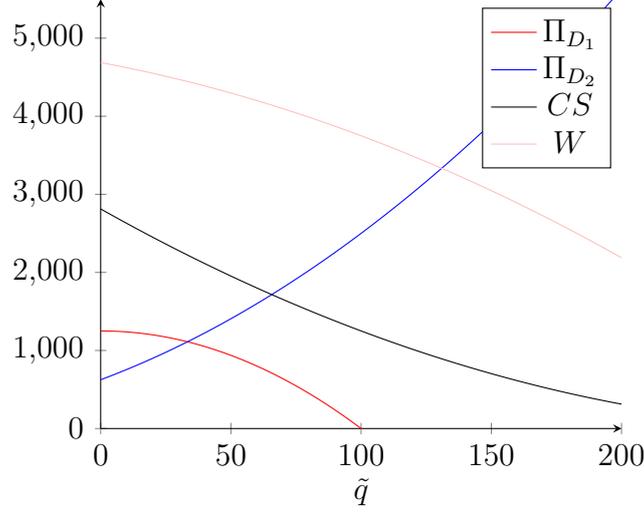
We have seen that in a dynamic setting, starting from our setting we would reach the Nash prediction in the medium run, ceteris paribus. This means that in the end, only the fundamentals matter on the variables represented in the previous figure. This suggests the importance it has for firms to invest in their cost structure in the very first periods so they can head the equilibrium in a more favourable position. Nevertheless, it is not the "most rational" firm the one that would be able to perform this investment. In fact, the firm with more availability of resources in the first period will depend on how accurately they forecast other players but also on the actual system of level-0 actions. Then, it is possible that for some $\tilde{q}$, the firm that makes more profits is not the one that perfectly forecasts the others, as it is the case in our example.

3.2 Upstream monopoly, downstream duopoly, and uniform pricing

After maximizing the objective function with respect to the choice variable, and assuming the previous beliefs tuples of D_1 and D_2 (i.e. $\tilde{b}_1 = (\tilde{b}_{1,2} = 0)$ and $\tilde{b}_2 = (\tilde{b}_{2,1} = 1)$), we know that the subjective best response functions are the following:

$$\pi_{D_1,1} = \frac{100 - p_u - \tilde{q}_{D_1,D_2}}{2} = \frac{100 - p_u - \tilde{q}}{2} \tag{9}$$

$$\pi_{D_2,2} = \frac{100 - p_u + \tilde{q}_{D_2,D_1}}{4} = \frac{100 - p_u + \tilde{q}}{4} \qquad (10)$$

$$\dot{Q} = \frac{300 - 3p_u - \tilde{q}}{4} \qquad (11)$$

One more time, the best response function of D_1 is the usual one. We observe that the effect of $\tilde{q}$ in the best response function of D_2 is the same than before, but even though the marginal cost still has a negative effect on its production, this direct effect is smaller than the usual one. This is due to the fact that there is an indirect effect which consists on the firm anticipating that for an increase in the marginal cost, the other firm will reduce its production, being the total effect less negative than a pure increase in marginal costs.

Since the direct expected competition effect of $\tilde{q}$ in both companies is obviously larger than the indirect one over D_2, the actual quantity produced is decreasing in this variable. Furthermore, the actual quantity produced is decreasing in the price set by the upstream monopoly.

The upstream monopoly is able to forecast this quantity demanded, as previously assumed, (one may either think that retailers demand the products or inputs just before selling to the customers). Then this firm maximizes its profits with respect to the uniform price charged.

$$\Pi_U = \left(\frac{300 - 3p_u - \tilde{q}}{4} \right) p_u \qquad (12)$$

$$p_u^{BR} = \frac{300 - \tilde{q}}{6} \qquad (13)$$

The optimal price strategy for the monopoly is represented by (13). This optimal pricing is decreasing in $\tilde{q}$ due to the negative final effect that this variable has on total quantity.

$$\pi_{D_1,1} = \frac{300 - 5\tilde{q}}{12} \qquad (14)$$

$$\pi_{D_2,2} = \frac{300 + 7\tilde{q}}{24} \tag{15}$$

$$\dot{Q} = \frac{300 - \tilde{q}}{8} \tag{16}$$

Here we rewrite (9), (10), and (11) with the optimal p_u and we see that now, each increase in $\tilde{q}$ is less harmful for the downstream firms since the monopolist will also take it into account and will reduce its choice variable. The effect on $\pi_{D_1,1}$ is still negative because the positive externality that generates the behaviour of firm U does not compensate the direct negative effect of the expected competition. Contrary, for D_2 both effects are positive, since the effect of $\tilde{q}$ per se (direct plus indirect) was positive, and now the behaviour of firm U also incentivises to expand quantity.

$$P(\dot{Q}) = \frac{500 + \tilde{q}}{8} \tag{17}$$

$$\Pi_{D_1} = \frac{5}{288}(300 + 7\tilde{q})(60 - \tilde{q}) \tag{18}$$

$$\Pi_{D_2} = \frac{1}{576}(300 + 7\tilde{q})^2 \tag{19}$$

The profit function of firm D_1 is concave with a maximum in $\tilde{q} = 8.571$. Then, some expected competition from a level 0 player is beneficial since this would allow this firm to expand production and face a lower marginal cost. However, larger expected competition will make this firm production to shrink too much. At the same time, the profit function of D_2 is convex with a maximum in the highest largest belief $\tilde{q}$ we might think as feasible. If we compare with the profits that the firms would obtain under Nash, we see that for low values of $\tilde{q}$, D_1 is better off with bounded rationality but D_2 is worst off, for intermediate values both are better off, and for larger values D_2 prefers bounded rationality but D_1 would like to be in a Nash scenario.

$$\Pi_U = \frac{1}{48}(300 - \tilde{q})^2 \tag{20}$$

11

$$CS = \frac{1}{128}(300 - \tilde{q})^2 \qquad (21)$$

$$W = -\frac{1}{128}(1300 + \tilde{q})(\tilde{q} - 300) \qquad (22)$$

In this setting, the profits of the upstream monopoly are decreasing in the expectation of initial competition (for all feasible values of $\tilde{q}$). This was expected since larger initial competition harms firm U both via price and quantity. The lower the initial expectation of competition is, the larger the consumer surplus will be since this leads to the highest possible quantity and lowest marginal cost. Relatedly, since in this case is where the charged marginal cost is the lowest and the closest to the real marginal cost of production in the economy, total welfare is also maximum at $\tilde{q} = 0$.

Fig. 2: Upstream monopoly, downstream duopoly, and uniform pricing
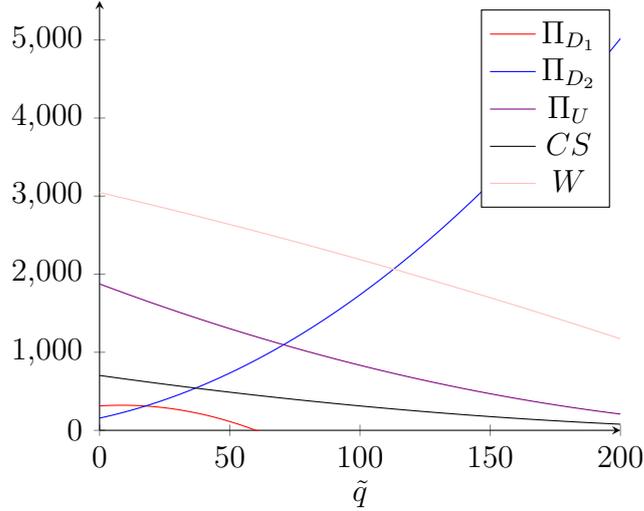
In contrast with the example without upstream firm, in the current case the dynamic transition towards Nash equilibrium is not as direct. Here we need the distorting behaviour of firm U and the actual quantities produced by the downstream firms to tend to the Nash stable outcome. It turns out that, for a favourable generating expectations rule, the downstream firms become higher degree overtime.[5] This ensures that the optimal pricing behaviour of the monopoly

---

[5] One may think in the generating process specified before. At $t = 1$ the static game just defined is played. At $t > 1$ downstream firms reformulate their subjective best response function. From the standard best response function $q_i^{BR} = \frac{100 - p_u - q_j}{2}$, firm D_1 take as q_{D_2} the quantity it can infer from the information available in the market. Firm D_2 best responds to this action.

tends to charge Nash price for any value of $\tilde{q}$ which, in turn, make the subjective best response function tend to the Nash quantity.

3.2.1 Dynamic transition towards Nash outcome

In this section we briefly describe how the static outcome tends to the Nash stable outcome overtime under the not very strong assumption that the process of generating expectations lead downstream firms to behave as if they where higher degree payers.[6]

First notice that the subjective best response function for this setting for a firm i level l that best responds to a firm level $l-1$ is the following.

$$\pi_{i,l} = \frac{100 - p_u}{2} - \{\mathbb{1}(l \neq 0, 1)\}\frac{100 - p_u}{2} - \sum_{i=1}^{l-1}\left(-\frac{1}{(-2)^i}\right) + \frac{\tilde{q}}{(-2)^l} \tag{23}$$

The next to be observed is that for larger values of l, the term that contains $\tilde{q}$ quickly becomes negligible and then the rest of the equation will be $\frac{100-p_u}{2}$ times something that will not affect the optimal behaviour of the upstream monopolist. Thus achieving the result expected. The monopolist will end up charging $p_u = 50$ for all values of $\tilde{q}$, and this will make the downstream firms to choose quantities that tend to the Nash quantities.

3.3 Upstream monopoly, downstream duopoly, and price discrimination

In this scenario, the new subjective best response functions follow the same logic that we have seen so far. The only difference will be that the direct and indirect effect that the upstream

---

[6] This idea was already mentioned in Nagel (1995). See Roth and Erev (1995) for an extension on learning models explaining such phenomena.

firm's price represents on the retailers will be separately identified on p_{D_1} and p_{D_2}.

$$\pi_{D_1,1} = \frac{100 - p_{D_1} - \tilde{q}}{2} \tag{24}$$

$$\pi_{D_2,2} = \frac{100 + p_{D_1} - 2p_{D_2} + \tilde{q}}{4} \tag{25}$$

The direct effect of the expected initial competition on firm D_1 is obviously larger than the indirect one on D_2. Thus the total actual production will be decreasing in this variable. At the same time, given that p_{D_1} supposes a positive externality for firm D_2, this variable reduces less total production than the price charged to D_2.

$$p_{D_1} = 60 - \frac{7}{15}\tilde{q} \tag{26}$$

$$p_{D_2} = 40 + \frac{2}{15}\tilde{q} \tag{27}$$

The upstream monopolist will be able to price discriminate each firm since it knows their level. Then the monopolist will just set the objective and maximize with respect to the two choice variables. This leads to two different linear price policies. For firm D_1, this price will be decreasing in the expectation of initial competition, since this firm is already harmed by this variable.

$$\pi_{D_1,1} = 20 - \frac{4}{15}\tilde{q} \tag{28}$$

$$\pi_{D_2,2} = 20 + \frac{1}{15}\tilde{q} \tag{29}$$

$$\dot{Q} = 40 - \frac{3}{4}\tilde{q} \tag{30}$$

$$P(\dot{Q}) = 60 + \frac{3}{15}\tilde{q} \tag{31}$$

This price discrimination leads both downstream firms to face two different effects: one due to competition with its rival, and another due to the price policy. For firm D_1, larger $\tilde{q}$ forces it to reduce production to compensate and increase its margin, but at the same time this makes p_{D_1} to be lower, incentivising this firm to expand production. In the case of firm D_2, it wants to expand its production, since larger $\tilde{q}$ means a reduction in the production of the other firm. Despite this fact, p_{D_2} increases, decreasing the incentives to increase production. Overall, the competition effect dominates in both firms.

$$\Pi_{D_1} = \frac{10}{15}\tilde{q}\left(20 - \frac{4}{15}\tilde{q}\right) \tag{32}$$

$$\Pi_{D_2} = \left(20 + \frac{1}{15}\tilde{q}\right)^2 \tag{33}$$

The profit function of firm D_1 is increasing for low values of $\tilde{q}$ and decreasing for large values. The margin that this firm has is increasing in the expected initial competition, since the total production and the marginal cost decrease for larger values of $\tilde{q}$. At the same time, since the competitive effect dominates, the quantity is decreasing in $\tilde{q}$. Contrary, we see that the profit function of D_2 is always increasing for the feasible set of $\tilde{q}$. This occurs due to the fact that both margin and production are increasing in this variable.

$$\Pi_U = \frac{2}{15}\tilde{q}^2 - 20\tilde{q} + 2000 \tag{34}$$

$$CS = \left(40 - \frac{30}{15}\tilde{q}\right)^2 \frac{1}{2} \tag{35}$$

$$W = -\frac{1}{15}(\tilde{q} + 800)(\tilde{q} - 200) \tag{36}$$

The profits of the monopolist are convex with a minimum in $\tilde{q} = 75$, and a maximum in $\tilde{q} = 0$. One can observe that, for any value of $\tilde{q}$, the profit function of the monopolist under price discrimination is larger or equal than the profit function under uniform pricing. In particular,

being able to price discriminate allows the monopolist to extract rents to the more convenient firm in each scenario. The effect of price discrimination on consumer surplus and total welfare is subject to our definition of feasible set of $\tilde{q}$. For low values, consumer surplus and total welfare are higher under price discrimination while for large values these are reduced compared to the uniform price setting.

Fig. 3: Upstream monopoly, downstream duopoly, and price discrimination
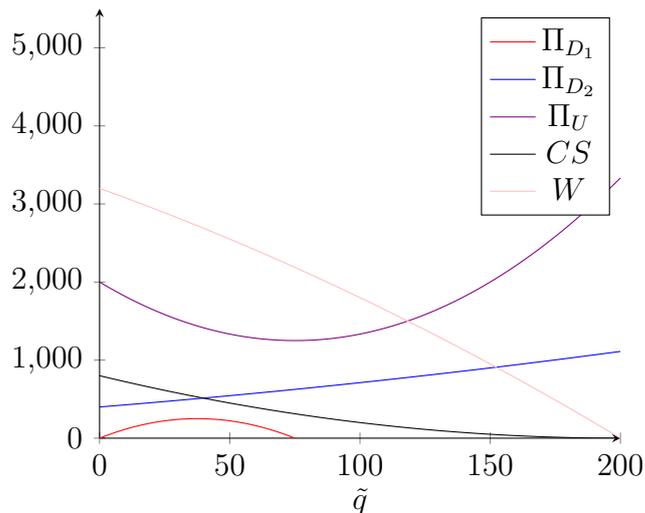
We can see that in this case, for any level of $\tilde{q}$, it pays to accurately best respond over the competitors. Here the firm with more availability of resources to invest is with no doubt firm D_2. One more time, with a convenient expectations generating process that makes firms to behave as if they achieve higher depths of reasoning overtime, ceteris paribus, the price policies would converge since the demand functions of the retailers would tend to look alike. Everything would tend to the Nash outcome in the medium run. If we drop the ceteris paribus and we think on the possibility of the firm accumulating larger profits investing on cost reductions (in scenario with cost structure obviously), the Nash prediction to which it converges is much different, where firm D_2 gets a life time advantage.

4   Concluding Remarks

We presented a Cournot competition game where players are modelled to be boundedly rational, which allows to understand first period competitive outcomes as well as the dynamics towards the Nash prediction. We base the analysis on specific examples, therefore a general approach is left to be done.

The applications show that reaching more depths of reasoning does not necessarily lead firms to obtain higher profits; it only pays to accurately best respond over the opponents in the case with price discrimination and upstream monopoly. The firm with a most favourable position after the first period generally depends on how well it predicts others but also on the actual system of level-0 players' actions.

This also suggests the role of early investments on cost structures as a factor explaining the divergence among initially similar firms as the outcome approaches the Nash prediction.

References

Camerer, C. F., Ho, T. H., and Chong, J. K. (2004). A cognitive hierarchy model of games.

Chong, J. K., Ho, T. H., and Camerer, C. (2016). A generalized cognitive hierarchy model of games. *Games and Economic Behavior*.

Costa-Gomes, M., Crawford, V. P., and Broseta, B. (2001). Cognition and behavior in normal-form games: An experimental study. *Econometrica*.

Costa-Gomes, M. A. and Crawford, V. P. (2006). Cognition and behavior in two-person guessing games: An experimental study. *American Economic Review*.

Hotelling, H. (1929). Stability and Competition. *The Economic Journal*.

Nagel, R. (1995). Unraveling in Guessing Games: An Experimental Study. *The American Economic*.

Nash, J. (1951). Non-Cooperative Games. *Annals of Mathematics*, 54(2):286–295.

Roth, A. E. and Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*.

Stahl, D. O. and Wilson, P. W. (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior and Organization*.