UNIVERSITY OF
CAMBRIDGE

Cambridge Working Papers in Economics

The Strategic Robustness of Mark-up Equilibria

David M. Newbery and Thomas Greve

CWPE 1341 & EPRG 1318

The Strategy Robustness of Marked-up Equilibria

David M. Newbery and Thomas Greve

November 2013

CWPE 1341 & EPRG 1318

UNIVERSITY OF CAMBRIDGE | Energy Policy Research Group

The strategic robustness of mark-up equilibria

EPRG Working Paper 1318

Cambridge Working Paper in Economics 1341

David M Newbery and Thomas Greve

Abstract This paper is an extension of the paper "The Robustness of Agent-Based Models of Electricity Wholesale Markets," EPRG1213 which was motivated by the problem of analysing market power in liberalised electricity markets. That paper examined two particular forms of agent-based models commonly used in electricity market modelling, and showed that while these mark-up equilibria are robust against Nash deviations. This paper extends the earlier results to explain why these equilibria are robust to single firm Nash Cournot deviations but shows they are vulnerable to more sophisticated deviations.

Keywords mark-up equilibria, stability, oligopoly, agent-based modelling, learning

JEL Classification C63, C73, D43, L10, L13, L94

Contact dmgn@cam.ac.uk
Publication November 2013
Financial Support Imperial College London and EPRG

The strategic robustness of mark-up equilibria

David M Newbery*and Thomas Greve†
Imperial College London and EPRG Cambridge

November 12, 2013

Abstract

Mark-up pricing, much discussed in the 1940s, led Grant and Quiggin (1994) to study Nash mark-up equilibria as alternatives to Cournot and Bertrand oligopoly equilibria. Mark-up models are increasingly used in modelling complex markets, where agent-based modelling is an attractive way of finding equilibria. While such learning behaviour can converge to a Nash equilibrium, that does not establish that the equilibrium is robust against more sophisticated strategy choices. This paper examines two particular forms of agent-based models commonly used in electricity market modelling and demonstrates that they are robust to single firm Nash Cournot deviations but not against more sophisticated deviations.

Keywords: mark-up equilibria, stability, oligopoly, agent-based modelling, learning

JEL Classification: C63, C73, D43, L10, L13, L94

1 Introduction

Empirical industry studies in the 1940's were concerned with the apparent mismatch between the theory of profit maximization and the evidence that managers had little if any concept of marginal cost and revenue. Instead they followed rules of thumb in setting prices at a mark-up over average cost (Hall and Hitch, 1939). As Grant and Quiggin (1994) observed, this led to a methodological debate in which Friedman (1953) argued that managers could be following rules

*Department of Electrical and Electronic Engineering, Imperial College London; contact address: Faculty of Economics, Sidgwick Avenue, Cambridge UK CB3 9DE; email: dmgn@cam.ac.uk. This is an extension of Newbery (2012), which has an error in that what were labelled Stackelberg equilibria were Nash equilibria. We are indebted to Rich Gilbert, Robert Ritz and Marta Rocha for comments, and to earlier referees, but remaining errors are ours alone.

†EPRG, Cambridge; contact address: Faculty of Economics, Sidgwick Avenue, Cambridge UK CB3 9DE; email: tg336@cam.ac.uk.

of thumb but still be pushed towards profit-maximizing behaviour, as those who chose non profit-maximizing mark-ups would earn lower profits and lose market share and/or exit. Grant and Quiggin further argued that this argument does not immediately apply to oligopolistic markets in which managers need to anticipate how their rivals will respond to their actions. If other firms were choosing mark-ups on costs, then this would affect the choice of a profit maximizing firm.

Grant and Quiggin (1994) examined the case in which each firm chose its profit-maximizing mark-up over marginal cost (in their model equivalent to choosing a mark-up over average cost) instead of choosing the optimal output level. This relates firms' actions more closely to what they observe, but they are still assumed to set the mark-up where marginal cost is equal to marginal residual revenue (although their opening remarks suggest that firms actually adjust their mark-ups in response to market conditions and so grope towards this profit-maximizing position, in the spirt of Friedman's defence). An alternative approach is take the observation that managers lack full information about their competitive environment more seriously, and explore the consequences of imperfect information. One approach has been to explore the consequences of imperfectly observing the actions of their rivals but nevertheless having otherwise full information about the consequences of any set of actions by the players. Bagwell (1995) explores one-shot Stackelberg games to disentangle the implications of the two standard assumptions - that one agent can move first, and that all other agents perfectly observe her choice. He claimed that the first-mover advantage is eliminated if there is even a slight amount of noise in observing the leader's choice. van Damme and Hurkens (1997) criticized this conclusion by noting that it depended on restricting choices to pure strategy equilibria, and that Bagwell's game always had mixed strategy equilibria close to the Stackelberg equilibrium when the noise was small. Their paper contains an extensive discussion of equilibrium selection, but remains firmly within the standard game-theoretic approach.

An alternative approach is to suppose that the information available to firms is limited to their own costs and market outcomes, which they can observe in a sequence of choices by all the firms. Again there are several ways this can be modelled, depending on what firms can observe. A number of papers have assumed that firms can observe the actions and subsequent profits earned by their rivals, so that they can imitate their behaviour. Vega-Redondo (1997) explores the classic Cournot setting of a market of quantity-setting firms producing a homogenous output in which firms experiment with a different level of output with a ε-(small) probability, and successful firms win out over less successful firms, as Friedman (1953) argued. He concludes that the final resting place of this stochastic dynamic process was the unique Walrasian (i.e. perfectly competitive) outcome. This approach was extended by Schipper (2009), who allowed firms either to imitate the output choices of the most profitable firms, or to optimize against the other firms, but with all firms making small mistakes. This time the long-run state is one in which imitators are better off than optimizers. A subsequent paper by Duersch, Oechssler, and Schipper (2012) extends this idea, which was prompted by observing experimental subjects playing a Cournot duopoly against a computer programmed with a variety of learning algorithms. The computer could easily be beaten, except when it followed the rule of copying the action of the most successful player in the previous round.

Schipper and his colleagues assume that all firms can identify both the actions and the resulting profits of their rivals, while Vega-Redondo adopts a more Darwinian approach in that exit or death is more likely with lower profits. But in a world of strategic rivalry, firms may go to considerable lengths to conceal both profits resulting from specific actions and the actions themselves, and with imperfect competition they may be able to survive without necessarily maximizing profits. Such a world can be explored through agent-based modelling. In such models, agents learn by testing out deviations from past strategy choices to see if they can increase profits, and continue to experiment until further improvements can no longer be found. The model must specify exactly what can and cannot be observed, and in many cases this is restricted to information about the firm's profit and his past choices (which can be computed from past outputs, the cost function and the associated market prices). As such the information is even more imperfect than normally assumed in formal game-theoretic studies, but may avoid some of the evident sensitivities of such studies to the precise form of the information assumed.

Agent-based modelling is widely used for markets characterized by complexities, such as electricity markets, which frequently exhibit market power and where the market design and transmission constraints can make it hard to find analytic solutions and hence characterize equilibrium solutions as functions of underlying parameters. It is therefore attractive to adopt an agent-based modelling strategy that can handle such complexity. Weidlich and Veit (2008) give an excellent survey of agent-based wholesale electricity market models and compare different learning strategies and their results. The aim of this paper is to explore the robustness of these models to more sophisticated behaviour by firms to see if they can benefit by choosing a different strategy than just marking up on marginal cost. The implications of finding that the standard modelling approaches are indeed vulnerable to such deviations is that agent-based models should test any proposed long-run equilibrium against such deviations, and we identify the kind of deviations that are most likely to disturb such equilibria.

2 Learning and mark-up models

Learning can be modelled in several ways: evolutionary learning, reinforcement learning, belief learning, experience-weighted attraction, among others (see Camerer, 2003). Reinforcement and belief learning are two of the most important approaches in individual learning modeling. The latter assumes that players update their beliefs based on past history, whereas the former makes fewer assumptions about the information available to an agent and their cognitive ability. As such, they make concrete the psychological idea of bounded rationality. Reinforcement learning models have been studied in economics (e.g. Roth and Erev, 1995) and in the artificial intelligence literature (e.g. Sutton and Barto, 1998). Q-learning[1] is a reinforcement learning model originally developed in the field of artificial intelligence (Watkins, 1989). An agent using a Q-learning algorithm keeps in memory a Q-value function of the weighted average of the payoffs obtained by playing a certain action in the past. The agent then plays with high probability the action that gives the highest payoff and with a small probability a randomly chosen different action (to test that any optimum found is not just a local maximum), observes the payoff it obtains and then updates its Q-value (Krause et al., 2006; Waltman and Kaymak, 2008).

One such formulation has agents choosing a mark-up on their marginal cost schedule, the natural counterpart to the approach of Grant and Quiggin (Krause et al., 2006). A key question facing such modelers is whether the resulting equilibrium is indeed a Nash equilibrium (where that is unique) in the space of actions allowed in the formulation of the game, and indeed what happens where there are multiple Nash equilibria. This is the question that Krause et al. (2006) address in the context of a simplified electricity market, and answer affirmatively for unique Nash equilibria. Models in which agents choose mark-ups on marginal costs are popular among agent-based models of electricity markets where generators are acutely aware of their marginal (primarily fuel) operating costs when submitting their daily offers into the wholesale and/or balancing markets, and most of the examples surveyed in Weidlich and Veit (2008) have this form.

There are good reasons for studying mark-up models in complex markets, such as electricity wholesale and balancing markets. Standard oligopoly models consider actions to be either quantities (supplies to the market), as in the Cournot formulation, or prices offered to the market (the Bertrand assumption). In the presence of uncertain or varying demand that are characteristic of electricity markets, supply function models, developed by Klemperer and Meyer (1989) and applied to electricity markets by Green and Newbery (1992), are attractive intermediate formulations. In particular, their linear solutions have been influential in motivating the kind of agent-based models considered here. In the simple quadratic cost, linear demand model there are two simple types of deviation from competitive bidding - marking-up the offer schedule by a constant amount, or changing the slope of the offer schedule, as in Hobbs et al. (2000). If

[1] set out in Watkins (1989) and further developed by Littman (1994) and Hu and Wellman (2003), and critically compared to other reinforcement learning models in Weidlich and Veit (2008).

marginal costs are linear, then there is a supply function equilibrium which is a slope mark-up on the marginal cost, and this solution has been widely used in simple analytical models (e.g. Green, 1999).[2]

Given their prevalence in modelling electricity markets, it is clearly timely to examine the robustness of such models. The obvious problem with agent-based modelling, where agents are assumed to learn about the profit consequences of their choices and then adapt these choices to increase profits, is that the action space over which they make choices may be too restrictive and may allow other more sophisticated agents to exploit this type of learning. A good defence of adaptive learning would be to show that the equilibrium of the form of learning were robust against more sophisticated players choosing from a wider set of actions. It is the purpose of this paper to explore the conditions under which this is the case and when not.

Note that this concept of robustness differs from the normal concept of stability in Nash games, which has a long history, summarized in e.g. Okuguchi and Yamazaki (2009). That paper establishes that under the generalized Hahn condition,[3] if a game has a unique Nash equilibrium, then it is globally stable. Stability here means that all players choose strategies from the same strategy set (e.g. outputs, or prices, or mark-ups) and adjust their response in the direction that increases profits. It is extremely useful for defending the study of the Nash equilibrium, as agents will converge on this equilibrium as they adjust their choices to increase profits. In contrast, this paper is concerned with a wider sense of stability or robustness in which there is no temptation for any agent or subset of agents to choose a strategy from a wider strategy set once the Nash equilibrium has been attained.[4] Even here, there are delicacies in describing the interactions between the players.

Agent-based models typically prescribe rules of behaviour that lead the agents to adjust their decision variables in the light of the outcome of the previous period's interactions. In the standard formulation all agents of the same type - for example, generators in electricity markets - would follow the same behavioural rules. A first test of robustness is against deviations by a player choosing actions from a wider strategy set than the remaining players, but otherwise taking the actions of the other players as given (i.e. the game remains Nash). We show that a unique deterministic Nash equilibrium is necessarily robust to single firm deviations, but demonstrate

[2]With finite support to the distribution of demand and no capacity limits as here, there will be a continuum of supply function equilibria, of which one is the linear supply function, which is the optimal response of any firm provided all other firms have chosen that supply function.

[3]The generalised Hahn condition (after Hahn, 1962) holds if the payoff to player i can be written as $U_i(x_i, X)$, where x_i is the strategy choice and $X = \sum x_j$, and when $h_i(x_i, X) \equiv \partial U_i/\partial x_i + \partial U_i/\partial X$, then $\partial h_i/\partial x_i < 0$ and $\partial h_i/\partial X \leq 0$ for all feasible strategies and all i.

[4]Simulation models that used Q-learning have shown that agents can learn to collude (Waltman and Kaymak, 2008) but in this paper we only explore incentives on smaller subsets of firms to follow different strategy choices.

that in the class of models considered here, not to collective deviations. A more stringent requirement is that a sophisticated player acting as a Stackelberg leader cannot profitably commit to a different strategy than the followers, anticipating and taking account of the response of these naive followers, who mechanically choose actions from their limited strategy set but adapt to the leader's choices. We show that the various mark-up models considered are not robust against such deviations.

The next section sets out the model and derives benchmark solutions for the linear-quadratic case, chosen as this has a well-defined slope and for which it is simple to define a slope mark-up.[5] Although demand and cost functions are assumed non-stochastic, the need to model behaviour as learning is motivated by the imperfect knowledge agents have about the shape of the demand schedule and the choices of their competitors - all they observe are the *ex post* price realisations. Section 3 then considers the two mark-up models and examines their robustness against the first type of Nash-Cournot deviation, demonstrating their robustness against single firm deviations but not to multi-firm deviations. Section 4 then constructs a counter-example for beneficial deviation by a single firm in the duopoly case in which a sophisticated leader plays a Stackelberg strategy - in this case by she can profitably commit to an inelastic output level or slope that differs from the mark-up chosen by the more naive follower.

3 The market model

Consider a market of n identical firms, $i = 1, ...n$, producing a non-storable homogenous output (electricity providing an excellent example), each with cost function $C(q_i) = aq_i + \frac{1}{2}cnq_i^2$. Without loss of generality set $a = 0$ by measuring prices relative to the level a. The marginal cost, MC, is cnq_i, which gives an aggregate competitive linear supply schedule cQ, independent of n. The natural demand formulation in a linear model has linear demand, $Q(p) = A - bp$, so $p = (A - Q)/b$. The inverse slope parameter can be set at $b = 1$ by a suitable choice of price units. The perfectly competitive solution is $p_c =$ MC $= cQ$, with

$$p_c = \frac{Ac}{(1+c)}, \quad q_c = \frac{A}{n(1+c)}. \tag{1}$$

In the constant marginal cost case, $c = 0$, $p_c = 0$, but otherwise if $c > 0$ the average (short-run) cost is half the price p_c (both relative to the price level normalization, a).[6]

[5] Grant and Quiggen (1994) consider a model with constant elasticities of supply and demand, and just a mark-up on average or marginal costs, but this does not allow any simple extension to marking up the slope of the firm's supply schedule.

[6] In the long run the average cost could be constant, with the location of the short-run cost curve depending on, e.g. capacity, as in McAfee, Preston and Williams (1992), Grant and Quiggen (1996), or Akgün (2004), with the short-run profits covering the fixed capital costs.

3.1 Nash-Cournot oligopoly

Now consider the Nash-Cournot solution in which each identical firm takes the output decisions of the others, $Q_{-i} = \sum_{j \neq i} q_j$, as given, and chooses q_i to maximize profit, $\pi_i(q_i, Q_{-i}) = p(q_i + Q_{-i})q_i - C(q_i)$. The first order condition (f.o.c.) gives the reaction function

$$q_i = (A - Q_{-i})/(2 + cn), \tag{2}$$

so the symmetric oligopoly solution is

$$p_o = \frac{A(\frac{1}{n} + c)}{(1 + c + \frac{1}{n})} > p_c, \quad q_o = \frac{A}{n(1 + c) + 1} < q_c. \tag{3}$$

Note that as $n \to \infty$, so $p_o \to p_c$, and that in the constant marginal cost case, $p_o = A/(1 + n)$, which is strictly positive.

3.2 The robustness of competitive and Cournot equilibrium

As this paper is concerned to examine the robustness of various equilibrium concepts, the logical place to start is with an exploration of the competitive equilibrium described above. In the n-firm case, would it be advantageous for one firm to choose a quantity different from that implied by offering to supply at its marginal cost, MC? The answer is clearly yes, as the supply of the remaining $n - 1$ firms can be subtracted from total demand to give a downward-sloping (and hence well-behaved) residual demand facing the firm, who will then choose output such that MC equals marginal revenue (MR), which will be below price, and hence output would be restricted as in the Nash Cournot case. The appendix provides the algebraic proof and demonstrates that the deviant chooses the same output as if he were in a symmetric Cournot oligopoly, but providing the residual demand is well-behaved so that profit is indeed maximized when the first order condition is satisfied (i.e. MC = MR), this verbal argument is sufficient. However, the Cournot deviant raises the price for all other price-taking firms, who, as they outproduce the deviant, will make higher profits than the deviant, raising the question whether firms would defect from a Cournot oligopoly to become price takers. This is the situation examined by Duersch et al. (2012), who argued that copying the actions of the most successful firm (in this case the price-takers) would lead to a competitive outcome.

More precisely, starting from the n-firm Cournot oligopoly described in (3), would it pay $r \geq 1$ firms to deviate to become price-taking competitive firms and setting output where MC equals price, or at $q_c = p/(nc)$? The residual demand facing the remaining $n - r$ Cournot players will be given from $p = A - (n - r)q_o - rp/(nc)$, or

$$p\gamma = A - (n - r - 1)q_o - q, \ \gamma = \frac{r + nc}{nc},$$

where q is the output choice facing any of the remaining Cournot firms and the output of each other Cournot firm, q_o, is taken as given. The reaction function is now $q = (A-(n-r-1)q_o)/(2+r+nc)$ and the solution for $q = q_o$ is exactly as in (3). However, the price is now

$$\begin{aligned} p &= \frac{A(nc+1+r)}{(1+n(1+c))(1+\frac{r}{nc})} < p_o, \\ q_c &= \frac{A(nc+1+r)}{n(\frac{1}{n}+(1+c))(nc+r)} = q_o\frac{(nc+1+r)}{(nc+r)} > q_o. \end{aligned}$$

Clearly as the deviant receives the same price as the Cournot firms, he makes more profit than the other Cournot firms and might seem to find it more attractive to act competitively than stay with the oligopolists. However, the profit the deviant makes as a competitor facing $r-1$ other competitive firms and $n-r$ Cournot firms is less than the profit he would make as one of $n-r+1$ Cournot firms facing $r-1$ competitive firms, so there is no advantage in deviating.[7]

4 Reinforcement or Q-learning

Under Q-learning, each firm selects a specified type of deviation from competitive bidding, and will continue to adjust the size of the deviation until the resulting market outcome is stationary and it is not profitable to make further adjustments. If there is a unique Nash equilibrium, this process should converge to that equilibrium. In the simple quadratic cost, linear demand model there are two simple types of deviation from competitive bidding: marking-up the offer schedule by a constant amount, or changing the slope of the offer schedule.

4.1 Q-learning with a constant mark-up

Suppose each firm offers its supply at a mark-up above marginal cost, MC,[8] so its offer price $p =$ MC $+m_i$, (so $p-$ MC $= m_i$). As MC $= ncq_i = p - m_i$, the supply schedule in price space is

$$q_i = \frac{(p-m_i)}{nc}, \quad (i = 1, .., i, .., n). \tag{4}$$

Under the mark-up form of Q-learning firm i keeps adjusting its mark-up m_i to improve its profit. The Nash equilibrium in this "multi-agent system" is a set of mark-ups m_i that maximize each firm i's profit, assuming the other firms choose their mark-up independently (the vector of other mark-ups, $\boldsymbol{m}_{-i}$, is thus treated as fixed). The market clearing price, MCP, $p(m_i, \boldsymbol{m}_{-i})$ is the price at which demand equals aggregate supply, found by summing over the q_i from (4). The

[7] This is most simply demonstrated numerically by varying n, r and c. Algebraic demonstration is tedious and unecessary.

[8] In this case the choice of marking up over MC gives a different equilibrium price to marking up over average cost, but producers are often more aware of variable costs than average total cost, so this is not a decisive criticism.

profit maximizing first order conditions (f.o.c.) give the n reaction functions for the mark-ups (derivations are provided in the appendix):

$$m_i(n^2(1+c)^2-1) = Anc + \sum_{j\neq i} m_j, \quad i = 1, ...n.$$

The symmetric solution is therefore

$$m = \frac{Ac}{n(1+c)^2-1}, \tag{5}$$

$$p_m = n(1+c)m = \frac{Ac(1+c)}{(1+c)^2-\frac{1}{n}}, \tag{6}$$

$$q_m = \frac{(p-m)}{nc} = \frac{A(1+c-\frac{1}{n})}{n(1+c)^2-1}. \tag{7}$$

In the constant returns case in which $c = 0$, the Nash equilibrium is the competitive price, $p_m = 0$, in contrast to the Nash-Cournot case, suggesting the competition in mark-ups is more competitive than competition in quantities. As with Nash-Cournot, as $n \to \infty$, $p_m \to p_c$. Proposition 1 establishes that the mark-up equilibrium is always more competitive than the Cournot equilibrium.

Proposition 1 *The symmetric Nash-Cournot oligopoly solution with linear demand and equal quadratic costs always has a higher equilibrium price than the Nash equilibrium in which firms choose their mark-up on marginal cost.*

Proof *Evaluate*

$$\begin{aligned} p_o - p_m &= \frac{A(1+nc)}{n(1+c)+1} - \frac{nAc(1+c)}{n(1+c)^2-1}, \\ \text{Sign}(p_o - p_m) &\propto (1+nc)(n(1+c)^2-1) - nc(1+c)(n(1+c)+1), \\ &= n-1 > 0. \end{aligned}$$

QED.

This confirms for the linear case Grant and Quiggen's (1994) finding for the constant elastic supply and demand case where competition in average mark-ups (in their model equivalent to competition in mark-ups on marginal cost) leads to more competitive outcomes than the Nash-Cournot equilibrium. It also suggests that a sophisticated firm observing other firms following a mark-up strategy might prefer to play a Cournot strategy in order to increase profits, in which case the mark-up equilibrium, once reached, would be vulnerable to deviations. That was certainly the case where the other players were behaving competitively, but the competitive equilibrium with a finite number of firms is not a Nash equilibrium, and provided only one firm considers deviating by choosing output rather than a mark-up, and moves simultaneously with the remaining firms, the equilibrium remains (in a sense to be defined) a Nash equilibrium and the deviant secures no advantage. This can be illustrated intuitively and argued formally in what follows.

4.2 Q-learning with a choice of slope

Suppose instead of choosing a mark-up on the linear MC schedule, whose slope is nc, firms choose a linear offer schedule which is increasingly above the true MC schedule (in quantity space). Their supply schedule in price space can be written as

$$q_i = \frac{s_i p}{n}, \quad i = 1, ...n, \tag{8}$$

instead of (4), where lower values of s_i indicate higher mark-ups (and again in the symmetric case aggregate supply will be independent of n). Aggregate demand equals supply gives the MCP, $p(s_i, \boldsymbol{s}_{-i})$:

$$A - p = p\frac{1}{n}\sum s_j, \quad p = \frac{A}{1 + \frac{1}{n}\sum s_j}.$$

Differentiating the profit function $\pi_i(s_i, \boldsymbol{s}_{-i}) = q_i(s_i, \boldsymbol{s}_{-i})p(s_i, \boldsymbol{s}_{-i}) - C(q_i)$ partially w.r.t. s_i, (details in the appendix) gives the reaction function:

$$s_i = \frac{n + \sum_{j \neq i} s_j}{1 + c(n + \sum_{j \neq i} s_j)}, \quad i = 1, ...n.$$

The symmetric equilibrium solves

$$\begin{aligned} 0 &= (n-1)cs^2 - (n - nc - 2)s - n, &(9)\\ s &= \frac{n - nc - 2 + \theta}{2(n-1)c}, \text{ where } \theta = \sqrt{(n-2)^2 + n^2 c(c+2)}, &(10)\\ p_s &= \frac{2A(n-1)c}{(1+c)(n-2) + \theta}, \; q_s = \frac{As}{n(1+s)}. &(11) \end{aligned}$$

In the constant returns case in which $c = 0$, $p_s = 0$, as with the mark-up case, but with a positive slope prices will be above the competitive solution. In the limit as $n \to \infty$, $p_s \to \frac{Ac}{1+c}$, the competitive (and Nash-Cournot) limit. Proposition 3 shows that the slope mark-up equilibrium is more competitive than the Nash Cournot equilibrium but less competitive than the mark-up equilibrium.

Proposition 2 *In a Nash game, if players assume that the strategy space is the choice of the slope of its offer schedule, then with a quadratic cost function, linear demand and identical players, the equilibrium yields a price that lies between the Nash-Cournot and the Nash mark-up price, and hence yields higher profits than choosing the best mark-up over marginal costs.*

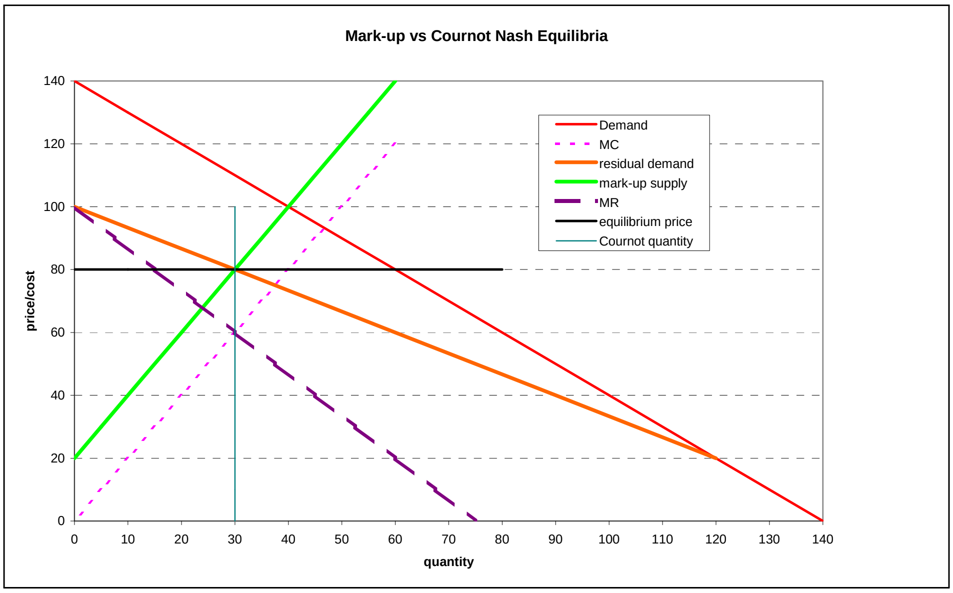
Proof. *Given in the appendix.* ■

4.3 Cournot deviations from mark-up equilibria

Figure 1 shows the result of the deviant, d, choosing output when the remaining firms choose to mark-up on marginal cost. The supply schedule of these firms from (4) is $q_j = (p-m)/(nc)$. The $n-1$ firms' collective supply is then subtracted from total demand to give the residual demand at price p as shown in Figure 1 and given by

$$q_d(p) = Q(p) - (n-1)q_j = (A-p) - \frac{(n-1)(p-m)}{nc} \equiv \alpha - \beta p, \tag{12}$$

where $\alpha = A + (1-\frac{1}{n})\frac{m}{c}$, $\beta = 1 + (1-\frac{1}{n})/c$. The deviant's optimal response is to choose q_d to maximize profit, $pq_d - C(q_d)$, where $p = (\alpha - q_d)/\beta$. The question to address is whether choosing the Nash mark-up given in (5) is robust against a player optimizing against this strategy. Proposition 2 shows, as the figure clearly demonstrates, that the answer is yes.

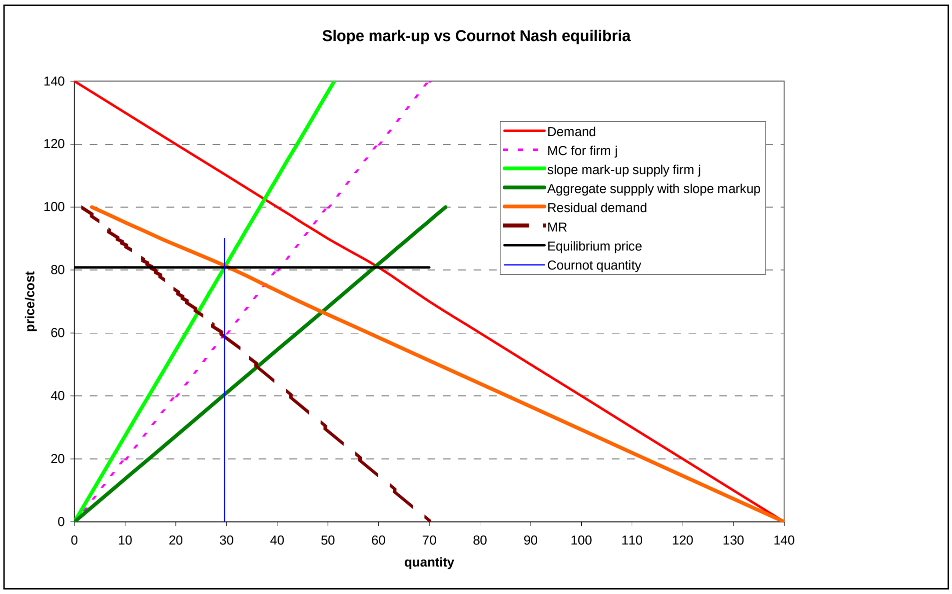


$A = 140,\ c = 1,\ n = 2,\ q_m = 30.$

Proposition 3 *In a deterministic Nash game, if players assume that the strategy space is the choice of the mark-up on its marginal cost or the slope of the supply schedule, then that player will maximize her profits regardless of whether another player chooses the same strategy choice as other players (mark-up on marginal cost or the slope of the supply schedule) or an optimal quantity to supply, and hence the Nash mark-up/slope equilibria are robust against Cournot deviations.*

Proof. *Each firm faces the same residual demand schedule and will choose the same optimal output whether they choose the optimal mark-up, the optimal slope, or the optimal quantity (or* *any other choice variable such as price) that corresponds to MC set equal to the marginal residual demand revenue. QED.* ■

Thus the Nash equilibrium behaviour in mark-ups or slopes is robust against a Nash-Cournot deviant who can choose from a broader set of strategies that also includes quantities, at least in a deterministic setting. Figure 2 illustrates this for the case in which all but the deviant firm choose their supply slopes.



$$A = 140,\ c = 1,\ n = 2,\ q_s = 29.59.$$

On reflection, this should not be surprising, as once the other firms have chosen their optimal mark-up (or slope) given the residual demand *they assume faces them*, the deviant faces the same residual demand and hence chooses the same mark-up, which is where the residual marginal revenue meets MC, shown in Figures 1 and 2. One might reasonably argue that the resulting outcome is no longer a true Nash equilibrium, in that while the deviant firm correctly predicts what the other firms will do, these mark-up firms are not correctly predicting what strategy the deviant is following and hence not correctly predicting the residual demand they face, although they are predicting the mark-up she will actually choose. We will explore this further in section 4. It is also important to note that although the equilibrium is robust to deviations, the resulting price depends on the strategy space, in this case mark-ups or supply slopes rather than quantities (and is lower as a result). Note that the argument does not depend on linearity of demand or marginal cost and is a direct consequence of Nash behaviour in a deterministic setting.

4.4 Multi-firm deviations

Suppose that r firms decide to play a Cournot strategy, knowing that $n-r$ firms will continue with their mark-up strategies but the other deviants will choose the same output as the first deviant (but independently, each taking the other's output as given). The market clearing condition is $p = A - q - (r-1)q_o - \frac{n-r}{nc}(p-m)$, or

$$p\gamma = A - q - (r-1)q_o + \frac{(n-r)m}{nc}, \quad \frac{dp}{dq} = -\frac{1}{\gamma}, \quad \gamma = 1 + (1 - \frac{r}{n})/c,$$

where q_o is the output of the each other Cournot deviant, taken as given, and q is the output choice to be made. As before the f.o.c. for the deviant is given by $p =$ MC $-qdp/dq$ or

$$p\gamma = ncq\gamma + q = A - q - (r-1)q_o + \frac{m(n-r)}{nc}.$$

Set $q_o = q$ and substitute for m to give

$$\begin{aligned}(r+1+nc\gamma)q_o &= A\frac{n^2(1+c)^2 - r}{n^2(1+c)^2 - n}.\\ \gamma p &= A - rq_o + \frac{A(n-r)}{n^2(1+c)^2 - n}.\end{aligned}$$

As a numerical example, let $n = 5$, $r = 2$, $c = 1$, $A = 100$, so that the symmetric mark-up equilibrium has $q_i = 9A/95 = 9.47$, $p = 52.6$. The deviants' output will be $98A/1045 = 9.378$, which is smaller, so the equilibrium will be different. The price will be $p = 52.8$ and so the profit of a deviating firm rather than conforming to the original strategy will be 274.8 rather than 274.2, or 0.22% higher. Thus there is a (small) incentive for a subset of more than one (very) sophisticated firms to deviate from a Nash mark-up equilibrium.

5 Robustness to Stackelberg deviations

Although a single deviant was unable to improve on her profits by choosing quantities rather than mark-ups, knowing that the remaining firms were acting on the (mistaken) assumption that all firms were choosing their mark-up facing the same residual demand schedule, there remains a question whether this is a consistently formulated equilibrium. If all firms observe is the consequences of their choices in the market price, then they are correctly choosing the optimal choice of mark-up (or output). If they are basing their choice of mark-up on assumptions about the shape of the residual demand they face then the assumed residual demand will be incorrect in the face of a Cournot deviant. One way round this inconsistency is to suppose that the deviant firm's strategy choice is known to the remaining firms, who nevertheless continue to choose their mark-up (and similarly the deviant knows that the other firms will behave that way). In a learning context, this would require the leader to stick to her output strategy, while the followers learned that they could then improve their profits by adapting to the new environment. The resulting equilibrium is most simply modelled as the outcome of a Stackelberg game in which the deviant is the leader who can commit, in this case to her output, and to which the followers respond. As the aim is to demonstrate that mark-up equilibria are not robust to this more sophisticated deviation, this section considers the simpler duopoly case ($n = 2$). The first step reproduces the classic Cournot Stackelberg oligopoly. The leader can commit to her output level, q_l, before the other (the follower) makes his choice, q_f, so that the leader can optimize against the follower's reaction function (2). The resulting equilibrium has

$$q_l = \frac{A(1+2c)}{2(1+4c+2c^2)} > q_f = \frac{A\left(1+6c+4c^2\right)}{4(1+c)(1+4c+2c^2)}. \tag{13}$$

These expressions are somewhat opaque, but simplify in the constant returns case in which $c = 0$ to the familiar solution $q_l = A/2$, $q_f = A/4 = p$, and the profit of the leader is $3A^2/16$, larger than the follower's profit, who receives only one-third as much or $A^2/16$, and also higher than under the symmetric Nash Cournot equilibrium of $A^2/9$. If $c = 1$, then $q_l = 12A/56 > q_o = A/5 > q_f = 11A/56$ and the leader's profit is 0.45% higher than in the symmetric duopoly.

5.1 Stackelberg quantity deviations from the mark-up equilibrium

The relevant question for studying the stability of agent-based learning is whether choosing the optimal Nash mark-up is robust against a more sophisticated player who in a sequential setting can stick to her optimal output while the other player continues to mark up on marginal cost but learns the optimal mark-up (effectively the correct position of his residual demand schedule). The previous duopoly example demonstrates that mark-up behaviour it is not robust (in a deterministic setting at least) against a Cournot (quantity-fixing) deviation and that is also the case when the follower is choosing his mark-up. Thus if the leader commits to a quantity, q_l, the follower chooses a mark-up m given the residual demand schedule $q_f(q_l) = A - p - q_l$. The follower's supply schedule is given by (4) with $n = 2$, so market clearing yields

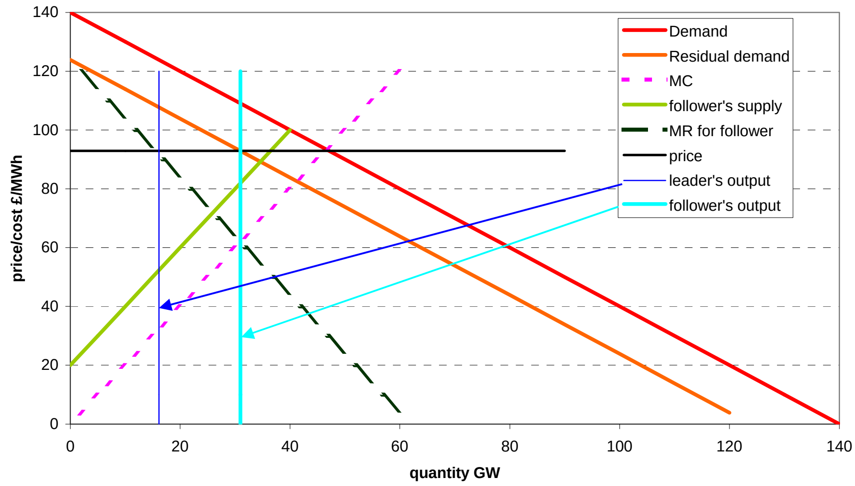
$$p = A - q_l - \frac{p-m}{2c}, \text{ or } p = \frac{m + 2(A - q_l)c}{1+2c}. \tag{14}$$

The follower chooses m to maximize profit (see appendix) resulting in a reaction function

$$q_f = \frac{p-m}{2c} = \frac{p}{1+2c}.$$

The leader's optimal response (and the follower's output) are exactly as in the Stackelberg Cournot equilibrium (13) in which both agents choose quantities, as shown in Fig. 3.

Stackelberg quantity vs mark-up equilibrium



$$A = 140,\ c = 1,\ n = 2,\ q_l = 16.2,\ q_f = 31.0$$

Proposition 4 *If one player assumes that the strategy space is the choice of the mark-up on its marginal cost, then with a quadratic cost function, linear demand and two identical players, the other player will find it profitable to commit to choosing a (different) optimal quantity to supply.*

Proof. *From equation* (13) $q_l > q_f$.*Since the leader could have chosen the same output as the follower but chose not to, she must be making higher profits.* ■

5.2 Stackelberg quantity deviations from the slope mark-up equilibrium

If the leader offers a fixed quantity q_l and the follower offers the supply schedule (8), $q_f = sp/2$, the market clearing price is

$$p = \frac{2(A - q_l)}{2 + s}, \quad \frac{ds}{dp} = -\frac{(2 + s)}{p}.$$

The follower's problem is to maximize $\pi_f = p^2(2s - cs^2)/4$ for which the f.o.c. is

$$\begin{aligned} p(2s - cs^2) &= p(1 - cs)(2 + s), \\ s &= \frac{2}{1 + 2c}. \end{aligned} \tag{15}$$

It may seem surprising that this does not depend on the leader's choice, but the follower's actual mark-up will be lower the larger the output of the leader and hence the lower the price, which gives the leader the advantage. The leader chooses q_l to maximize profit $\pi_l = q_l p(q_l) - cq_l^2$, given that $p = 2(A - q_l)/(2 + s)$. The profit maximizing solution is again the Cournot Stackelberg equilibrium (13).

Proposition 5 *If one player assumes that the strategy space is the choice of the slope of its offer, then with a quadratic cost function, linear demand and two identical players, the other player will find it profitable to commit to choosing a (different) optimal quantity to supply.*

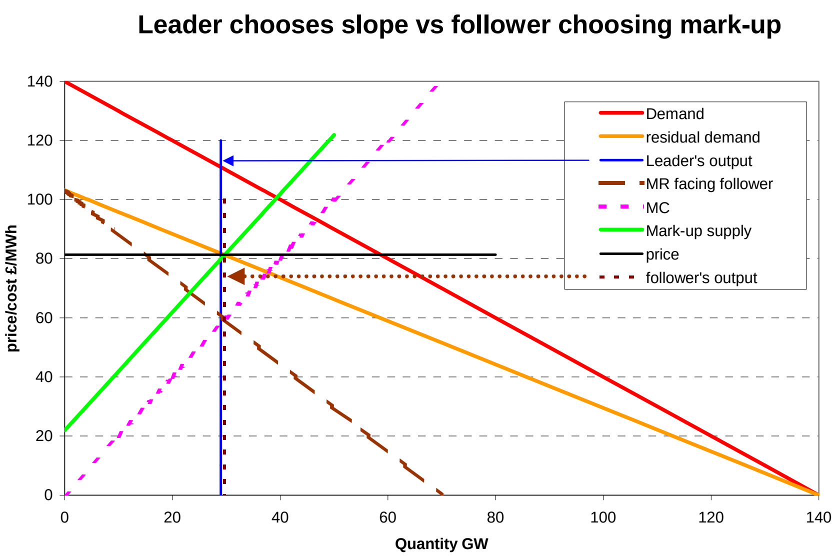
Proof. *From equation* (13) $q_l > q_f$*.Since the leader could have chosen the same output as the follower but chose not to, she must be making higher profits.* ■

5.3 Stackelberg slope deviations from the mark-up equilibrium

If the leader chooses a slope supply schedule given by (8) with $n = 2$, and the follower chooses his mark-up m, the market clearing price is given by

$$p = A - \frac{sp}{2} - \frac{p-m}{2c}, \text{ or } p = \frac{2Ac+m}{2c+1+sc}. \tag{16}$$

Fig 4 illustrates and the appendix demonstrates that this can be a profitable deviation, although not as profitable as choosing output rather than slope when confronting mark-up followers.



$$A = 140, \ c = 1, \ n = 2, \ q_l = 29.0, \ q_f = 29.7$$

6 Conclusion

Agent-based models are attractive in attempting to model outcomes in complex markets where some agents can act strategically, and there has been considerable interest in whether adaptive or Q-learning will lead to Nash equilibria, as these would seem natural equilibrium concepts.

However, as with all attempts to model strategic behaviour, the resulting equilibrium is sensitive to the action space from which agents choose. Standard oligopoly models consider actions to be either quantities (supplies to the market), as in the Cournot formulation, or prices offered to the market (the Bertrand assumption). In the presence of uncertain or varying demand, supply function models, developed by Klemperer and Meyer (1989) and applied to electricity markets by Green and Newbery (1992), are attractive intermediate formulations, and their linear solutions[9] have been influential in motivating the kind of agent-based models considered here.

All these specifications assume a unitary or owner-managed firm pursuing maximum profit but under managerial capitalism the ultimate owners need to motivate managers. Ritz (2008) argues that rewarding managers for increasing their market share is consistent with the evidence and can be a useful in pursuing more collusive strategies. Ritz concludes that though competing for market shares seems more aggressive, it is indeed more "robust" to strategic manipulations. Following the same line, Vickers (1985) and Fershtman and Judd (1987) study the strategic distortion of preferences. These models compare different proximate objective functions with the same choice variable (and with the same goal of ultimately maximizing profits). In contrast, our paper has the same ultimate objective function – profit maximization – but compares different choice variables.

While the choice of action space in optimizing models is normally guided by the market structure and the actions that agents have, the choice of action space in agent-based models is normally guided by tractability, where a choice of a single parameter (such as the mark-up over marginal cost or the slope of the supply schedule) considerably simplifies the problem. This paper has shown that the two mark-up strategies considered are more competitive than Nash-Cournot behaviour in the linear supply and demand duopoly case, with the Nash choice of the optimal slope of the offers yielding lower prices than the Cournot duopoly prices but higher prices than the optimum mark-up on linear marginal costs. While these mark-up equilibria are robust against Nash deviations by single firms choosing quantities instead of mark-ups (so they are in that sense Nash equilibria), they are not robust to either group deviations or to more sophisticated single firm Stackelberg deviations in which the deviant maintains her output and the remaining players adapt to that and find the corresponding mark-up equilibrium output levels. The deviant player makes higher profits following this Cournot Stackelberg strategy (or the slope mark-up against a fixed mark-up strategy), casting doubt on the robustness of Nash mark-up equilibria.

The implication of this finding for agent-based modelling is that it would be sensible to

[9] Supply function models typically have a continuum of solutions, one of which may be linear, providing there are no relevant capacity constraints. Where capacity constraints are important, there may be unique but non-linear solutions.

test any resulting numerically found convergence results for robustness against different choice variables by one or more (large) agents, who would need to maintain their choice for sufficiently many iterations to induce responses by the remaining agents. A finding that the resulting equilibria were close to the original solution would attest to its robustness, but if not the proposed solution would remain suspect. We plan to explore such questions in simple and then more complex agent-based modelling situations.

References

Akgün, U. (2004) "Mergers with Supply Functions', *The Journal of Industrial Economics*, Vol. 52 No. 4 pp. 535-46.

Bagwell (1995). "Commitment and Observability in Games", *Games and Economic Behavior*, 8, 271-280.

Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*, Princeton University Press, Princeton, New Jersey.

Duersch, P., J. Oechssler, and B.C. Schipper (2012). "Unbeatable imitation," *Games and Economic Behavior*, 76(1), 88-96.

Fershtman, C. and K. Judd (1987), "Equilibrium incentives in oligopoly", *American Economic Review*, 77, 927-940.

Friedman, Milton. (1953). *Essays in Positive Economics*, University of Chicago Press, Chicago.

Grant, S. and Quiggin, J. (1994), "Nash equilibrium with markup-pricing oligopolists", *Economics Letters* 45, 245-51

Green R. (1999). "The electricity contract market in England and Wales", *The Journal of Industrial Economics*, Vol:47, pp. 107-124

Green, R. J. and D. M. Newbery (1992). "Competition in the British Electricity Spot Market", *Journal of Political Economy,* 100 (5), pp. 929-53.

Hahn, F.H. (1962) "The Stability of the Cournot Oligopoly Solution," *Review of Economic Studies*, Vol 29, 329-333.

Hall, R. and Hitch, C. (1939). "Price Theory and Business Behaviour", *Oxford Economic Papers* Vol. 2, pp. 12-45.

Hobbs, B. C. Metzler, and J.-S. Pang, (2000). "Strategic gaming analysis for electric power systems: an MPEC approach", *Power Systems, IEEE Transactions on*, vol. 15, no. 2, pp. 638–645.

Hu, J. and M. Wellman, (2003). "Nash Q-learning for general-sum stochastic games", *Journal of Machine Learning Research*, vol. 4, pp. 1039–1069.

Klemperer, P. D. and M.A. Meyer, (1989). "Supply Function Equilibria in Oligopoly under Uncertainty", *Econometrica*, 57 (6), pp. 1243-1277.

Krause, T., G. Andersson, D. Ernst, E. Vdovina-Beck, R. Cherkaoui, and A. Germond, (2006). "A comparison of Nash equilibria analysis and agent-based modelling for power markets", *International Journal of Electrical Power & Energy Systems*, 28(9), pp. 599-607.

Littman, M. (1994). "Markov games as a framework for multiagent reinforcement learning", in *Proceedings of the Eleventh International Conference on Machine Learning.* pp.157-163, San Francisco, CA: Morgan Kaufmann.

McAfee, J. Preston and M .A. Williams (1992). "Horizontal mergers and Antitrust Policy", *The Journal of Industrial Economics*, Vol. 40 No. 2 pp. 181-7.

Newbery, D.M.G. (2012) "The Robustness of Agent-Based Models of Electricity Wholesale Markets", EPRG Working Paper 2012, available at http://www.eprg.group.cam.ac.uk/eprg1213/eprg-wp1213_complete/.
McAfee, J. Preston and M .A. Williams (1992). "Horizontal mergers and Antitrust Policy", *The Journal of Industrial Economics*, Vol. 40 No. 2 pp. 181-7.
Ritz, R. A. (2008), 'Strategic incentives for market share', *International Journal of Industrial Organization*, 26, 586-597.

Roth, A. E. and I. Erev (1995), 'Learning in extensive form games: experimental data and simple dynamic models in the intermediate term', *Games and Economic Behavior*, 8, 164-212.

Schipper, B.C. (2009). "Imitators and optimizers in Cournot oligopoly", *Journal of Economic Dynamics and Control*, 33, 1981-1990.

Sutton, R. S. and A. G. Barto (1998). *Reinforcement learning: An introduction*, MIT Press, Cambridge, MA.

Waltman, L. and U. Kaymak (2008), "Q-learning agents in a Cournot oligopoly", *Journal of Economic Dynamics and Control*, 32, 3275-3293.

Okuguchi, K and T. Yamazaki (2009) "Global Stability of Nash Equilibrium in Aggregative Games," Working Paper 77, Nigata University, at http://ecows.econ.niigata-u.ac.jp/~tyamazak/wp/WP_No77

Van Damme and Hurkens (1997). "Games with Imperfectly Observable Commitment", *Games and Economic Behavior*, 21, 282-308.

Vega-Redondo, F. (1997). "The evolution of Walrasian behavior", *Econometrica*, 65, 375-384.

Vickers, J. (1987), "Delegation and the theory of the firm," *Economic Journal*, 95, 138-147.

Watkins, C. (1989). "Learning from Delayed Rewards," Ph.D. dissertation, Cambridge University, Cambridge, England.

Weidlich, A. and D Weit (2008). "A critical survey of agent-based wholesale electricity market models", *Energy Economics,* 30, 1728-59 at http://dx.doi.org/10.1016/j.eneco.2008.01.003

7 Appendix

7.1 Deviations from competitive equilibrium

The supply function of each competitive firm j is given by $ncq_j = p$, so residual demand is $q = A - p - (n-1)q_j = A - p(\frac{nc+n-1}{nc})$. The deviant firm d then maximizes profit, for which the f.o.c. gives

$$q_d = \frac{A}{1+n+nc}, \quad p_d = \frac{Ac(1+c)}{(1+c)^2 - \frac{1}{n^2}} > p_c = \frac{Ac}{1+c}.$$

Note that the optimal output is the same as the Nash Cournot output give in (3). Profit for the deviant is

$$\pi = \frac{A^2c(n(1+c)^2 + 2(1+c) + 1/n)}{2(1+n+nc)^2((1+c)^2 - 1/n^2)} > \pi_c = \frac{A^2c}{2n(1+c)}.$$

This last inequality can be demonstrated by subtracting competitive profit from deviant profit, which, after simplification, has the sign of

$$c(1+c+1/n)^2 \geq 0.$$

Thus, provided the competitive equilibrium has a positive price (i.e. $c > 0$), it is more profitable to set quantity (or offer whatever the market demands at the price p_d above) than to act competitively. This is equivalent to observing that the competitive equilibrium here is not a Nash equilibrium with finitely many firms.

7.2 Q-learning with a constant mark-up

If firm i sets mark-up m_i, then the market clearing price, MCP, $p(m_i, \boldsymbol{m}_{-i})$ solves, after substituting q_i from (4),

$$\begin{aligned} p &= A - \sum q_i = \frac{(Anc + \sum m_i)}{n(1+c)}, \text{ so } \frac{\partial p}{\partial m_i} = \frac{1}{n(1+c)}, \\ ncq_i &= p - m_i, \text{ so } \frac{\partial q_i}{\partial m_i} = \frac{1}{nc}(\frac{\partial p}{\partial m_i} - 1). \end{aligned}$$

The f.o.c. from maximizing profit w.r.t. m_i gives

$$\begin{aligned} \frac{\partial \pi_i}{\partial m_i} &= (p - \mathrm{MC})\frac{\partial q_i}{\partial m_i} + q_i \frac{\partial p}{\partial m_i} = \frac{m_i}{nc}(\frac{\partial p}{\partial m_i} - 1) + q_i \frac{\partial p}{\partial m_i}, \\ &= \frac{-m_i(1 - \frac{1}{n} + c)/c + q_i}{n(1+c)} = 0, \quad \text{so } q_i = \frac{m_i}{nc}(n - 1 + nc), \\ ncq_i &= p - m_i = m_i(n - 1 + nc), \end{aligned}$$

$$\begin{aligned} nm_i(1+c) &= p = \frac{(Anc + m_i + \sum_{j\neq i} m_j)}{n(1+c)}. \\ m_i(n^2(1+c)^2 - 1) &= Anc + \sum_{j\neq i} m_j. \end{aligned}$$

The symmetric equilibrium has $m_i = m$:

$$\begin{aligned} m(n^2(1+c)^2 - 1) &= Anc + (n-1)m, \\ m(n(1+c)^2 - 1) &= Ac. \end{aligned}$$

7.3 Slope mark-up equilibrium

$$p = \frac{nA}{n + \sum s_j}, \quad \frac{\partial s_i}{\partial p} = -\frac{nA}{p^2}.$$

Differentiate the profit function $\pi_i(s_i, \boldsymbol{s}_{-i}) = q_i(s_i, \boldsymbol{s}_{-i})p(s_i, \boldsymbol{s}_{-i}) - cnq_i^2/2 = p^2(s_i - \frac{1}{2}cs_i^2)/n$ partially w.r.t. s_i to give

$$\begin{aligned} p(2s_i - cs_i^2) &= nA(1 - cs_i), \\ 2s_i - cs_i^2 &= (n + \sum s_j)(1 - cs_i), \\ s_i &= \frac{n + \sum_{j\neq i} s_j}{1 + c(n + \sum_{j\neq i} s_j)}. \end{aligned}$$

The symmetric equilibrium $s_i = s$ solves

$$\begin{aligned} 0 &= (n-1)cs^2 - (n - nc - 2)s - n, \\ s &= \frac{n - nc - 2 + \theta}{2(n-1)c}, \quad \theta = \sqrt{(n-2)^2 + n^2c(c+2)}, \\ p_s &= \frac{A}{1+s} = \frac{2A(n-1)c}{(1+c)(n-2) + \theta}, \quad q_s = \frac{As}{n(1+s)}. \end{aligned}$$

Proof of Proposition 2 *With a constant number of firms, if a firm's output under the Nash slope equilibrium is less than under the Nash mark-up output, then the price will be higher. The sign of the difference between the Nash mark-up output and the Nash slope output is given by*

$$\begin{aligned} \mathrm{Sign}(q_m - q_s) &= \mathrm{Sign}(\frac{1}{q_s} - \frac{1}{q_m}) \propto \frac{n(1+s)}{s} - \frac{n(1+c)^2 - 1}{1 + c - \frac{1}{n}}, \\ &\propto n(1+c)(1 - sc) - 1, \\ &\propto n^2(1+c)^2(1 - (1-\varphi)^{\frac{1}{2}}) - 2(n-1), \end{aligned}$$

where $\theta = n(1{+}c)\sqrt{1-\varphi}$ and $\varphi = \frac{4(n-1)}{n^2(1+c)^2}$. *But by expansion,* $(1{-}(1{-}\varphi)^{\frac{1}{2}}) \equiv \gamma > \frac{1}{2}\varphi = \frac{2(n-1)}{n^2(1+c)^2}$, *so* $\mathrm{Sign}(q_m - q_s) > 0$.

The other part requires us to show that $\mathrm{Sign}(q_s - q_o) > 0$.

$$
\begin{aligned}
\mathrm{Sign}(\frac{1}{q_o} - \frac{1}{q_s}) &\propto 1 + n(1+c) - n - \frac{n}{s}, \\
&\propto s(1+nc) - n, \\
&\propto (n - 2 - nc + \theta)(1+nc) - 2nc(n-1), \\
&= n(1-c) - n^2c(1+c) - 2 + (1+nc)n(1+c)(1-\varphi)^{\frac{1}{2}}, \\
&= 2(n-1) - (1+nc)n(1+c)\gamma, \\
&= n - 1 - \varepsilon > 0, \quad \text{where } \varepsilon = \frac{(\gamma - \varphi/2)(1+nc)n^2(1+c)}{2(n-1)}.
\end{aligned}
$$

Thus in the case of a symmetric oligopoly it is possible to rank the equilibrium prices $p_o > p_s > p_m$. QED

7.4 Stackelberg quantity choice facing mark-up players

From (14), $\partial p/\partial m = 1/(1+2c)$ and the follower maximizes profit $pq_f - cq_f^2$ by his choice of m as before, noting that $\partial q_f/\partial m = (\partial p/\partial m - 1)/2c$ and $p-$ MC $= m$, giving f.o.c

$$
\begin{aligned}
\frac{\partial \pi_f}{\partial m} &= m\frac{\partial q_f}{\partial m} + q_f\frac{\partial p}{\partial m} = 0, \\
m &= \frac{-q_f \partial p/\partial m}{\partial q_f/\partial m} = \frac{2cq_f \partial p/\partial m}{1 - \partial p/\partial m} = \frac{p-m}{2c}.
\end{aligned}
$$

The leader's optimal response is to choose q_l to maximize profit, $pq_l - cq_l^2$, where $p = A - q_l - q_f = A - q_l - p/(1+2c)$, so

$$
p = \frac{(A - q_l)(1+2c)}{2(1+c)}.
$$

Then $2(1+c)\pi = (A - q_l)(1+2c) - 2c(1+c)q_l^2$, for which the f.o.c is

$$
\begin{aligned}
0 &= (1+2c)(A - 2q_l) - 4c(1+c)q_l, && (17) \\
q_l &= \frac{A(1+2c)}{2(1+4c+2c^2)}, && (18) \\
2(1+c)p &= (A - q_l)(1+2c) = A(1+2c)(1 - \frac{1+2c}{2(1+4c+2c^2)}), \\
p &= \frac{(1+2c)(1+6c+4c^2)A}{4(1+c)(1+4c+2c^2)}, \\
q_f &= \frac{p}{1+2c} = \frac{(1+6c+4c^2)A}{4(1+c)(1+4c+2c^2)}. && (19)
\end{aligned}
$$

7.5 Stackelberg quantity deviations from the slope mark-up equilibrium

The leader's profit is $\pi = pq_l - cq_l^2$, so the f.o.c. for profit maximization is

$$
p - 2cq_l + q_l\frac{\partial p}{\partial q_l} = 0, \quad \text{where } \frac{\partial p}{\partial q_l} = -\frac{2}{2+s}.
$$

Substituting for s from (15) and rearranging gives (13).

7.6 Stackelberg slope deviations from the mark-up equilibrium

From (16), $\partial p/\partial m = 1/(1+2c+sc)$ and the follower maximizes profit $pq_f - cq_f^2$ by his choice of m as before, noting that $\partial q_f/\partial m = (\partial p/\partial m - 1)/2c$ and $p-$ MC $= m$, giving f.o.c

$$\begin{aligned} \frac{\partial \pi_f}{\partial m} &= \frac{m}{2c}(\frac{\partial p}{\partial m} - 1) + q_f \frac{\partial p}{\partial m} = 0, \\ m &= p\frac{\partial p}{\partial m} = \frac{p}{1+2c+sc}, \text{ but} \\ p &= \frac{2Ac+m}{2c+1+sc}, \end{aligned}$$

so

$$p = \frac{2Ac(1+2c+sc)}{c(2+s)(2+2c+sc)}.$$

The leader's profit function is proportional to $(2s - cs^2)p^2$ for which the f.o.c. can be written

$$\begin{aligned} \frac{d\log p}{ds} &= \frac{cs-1}{s(2-cs)}, \text{ but} \\ \frac{d\log p}{ds} &= \frac{c}{1+2c+sc} - \frac{1}{2+s} - \frac{c}{2+2c+sc}, \end{aligned}$$

from which the optimal value of s can be determined. For example if $c = 1$, $A = 100$, the symmetric slope equilibrium has $s = \sqrt{3} - 1 = 0.732$ and $q = 21.1$, but in this case the solution is $s = 0.7142$, which is a slightly steeper slope. Instead of the symmetric price being $A/\sqrt{3} = 57.7$, the new price is 58.1 and the leader's output is 20.7, profit 774. Had the leader accepted the mark-up equilibrium, the price would have been 57.1, the output 21.4 and profit 765, which is less. If the leader had chosen output rather than slope the equilibrium price would have been 58.9, quantity 21.4 and profit 800, higher than choosing the slope.