The "Win-Continue, Lose-Reverse" rule in oligopolies: robustness of collusive outcomes

Segismundo S. Izquierdo [1,*], Luis R. Izquierdo [2]

[1] Department of Industrial Organization, Universidad de Valladolid, 47011 Spain;
E-mail: segis@eii.uva.es.
[2] Department of Civil Engineering, Universidad de Burgos, 09001 Spain;
E-mail: luis@izquierdo.name.
* Author to whom correspondence should be addressed.

Version: 21/11/2014

Abstract: The so-called "Win-Continue, Lose-Reverse" (WCLR) rule is a simple iterative procedure that can be used to choose a value for any numeric variable (e.g. setting a price or a production level to maximise profit). The rule dictates that one should evaluate the consequences of the last adjustment made to the value (e.g. an increase or a decrease in production), and keep on changing the value in the same direction if the adjustment led to an improvement (e.g. if it led to greater profits), or reverse the direction of change otherwise. Somewhat surprisingly, this simple rule has been shown to lead to collusive outcomes in Cournot oligopolies, even though its application requires no information whatsoever about the choices made by any competing firms or about their results.

In this paper we show that the convergence of the WCLR rule towards collusive outcomes can be very sensitive to small independent perturbations in the cost functions or in the income functions of the firms. These perturbations typically push the process towards the Nash equilibrium of the one-shot game. We also explore the behaviour of WCLR against other strategies and demonstrate that WCLR is easily exploitable. We then conduct a similar analysis of the WCLR rule in a differentiated Bertrand model, where firms compete in prices. As in the Cournot model, our simulations show convergence of WCLR firms to collusive outcomes, high sensitivity to small independent perturbations, and vulnerability to be exploited by other strategies.

Keywords: Cournot, Bertrand, duopoly, oligopoly, Win-Continue, Lose-Reverse, collusion, differentiated, simulation.

1. Introduction and motivation

It is generally recognised that the actual decision-making processes followed by real-world firms when they have to set prices or production levels have often little to do with those assumed in the idealized analytical framework of perfect information[1]. In practice, the use of simple revisable strategies, imitation tactics and rules of thumb seems to be a key ingredient in many decision processes.

Thus, when examining a market, the behaviour of its participants, and the resulting emergent dynamics, it seems valuable to complement the perfect-information analysis with studies that also consider decision procedures that enjoy greater empirical support

[1] This statement does not necessarily imply that market predictions made using the perfect-information model are irrelevant in real life; the famous "as if" theory of Friedman (1953) proves sufficiently accurate and useful in many contexts.

and which may be deemed plausible for the context at hand (Kimbrough and Murphy, 2009).

This point is particularly relevant in markets potentially subject to regulation (e.g. oligopolies) and in situations where the perfect-information theoretical analysis of the social interaction reveals the presence of multiple possible equilibria –as is often the case in indefinitely repeated strategic interactions, including oligopolies in particular. Consequently, several different rules for setting prices or production levels in oligopolies have been analyzed. Bigoni and Fort (2013) provide a recent review of the theoretical and experimental literature on learning in oligopolies.

In this paper we analyse two types of oligopolies: one where firms compete in quantities (a la Cournot), and another where firms compete in prices (a la Bertrand). In both cases, we consider that the market process advances in discrete time steps and at every time step the companies have to simultaneously choose whether to increase or decrease the value of their decision variable (i.e. quantities q_i in the Cournot oligopoly, and prices p_i in the Bertrand oligopoly). The decision rule considered here can be simply stated as: repeat your last action (i.e. an increase or a decrease in your decision variable) if your profits have grown; otherwise, choose the opposite action. This simple rule has been named "Win-Continue, Lose-Reverse" (WCLR) by Huck et al. (2003)[2], who conducted a thorough study of its convergence properties in symmetric Cournot oligopolies.

The WCLR rule adjusts the level of the decision variable in the direction that is expected to make profits grow, according to the observed effect on profits of the last increment/decrement. Note that this gradual adjustment strategy can be considered a type of reinforcement learning rule: an action (i.e. an increase or decrease in production or price) is deemed satisfactory –and therefore repeated– if it provides a profit boost, and it is considered unsatisfactory –and therefore avoided– otherwise (Izquierdo and Izquierdo, 2012).

Mathematically, the WCLR strategy presents some similarities with a gradient ascent optimization method. In fact, if the profits of a company were to depend only on its own price or level of production (as in a monopoly with stable demand and costs), this rule would basically be a gradient ascent method and, under conditions that are well known in the optimization literature (Snyman, 2005), it would lead to the vicinity of a local optimum. In a duopoly, however, the profits of a company depend on its competitor's price or output level, and the application of the WCLR rule by each of the companies independently does not constitute a gradient ascent method for the joint profit of the two companies. Thus, it is interesting to study to which reference point of the strategic game (e.g. collusive outcome, competitive outcome, or one-shot Nash equilibrium) such a simple strategy converges, if it does converge to any at all.

For a Cournot duopoly in which companies vary their production levels q_i by a predefined amount δ (step size), Huck et al. (2003) show that, under rather general conditions, the quantities q_i converge for low values of δ to a small area around the cooperative (collusive) solution. Friedman et al. (in press) have recently proposed a slightly modified version of WCLR to explain experimental data from the lab and they also prove that their WCLR version converges to the collusive solution.

---

[2] The same authors use the name "trial and error" in Huck et al. (2004), where they also present and discuss this learning rule in a discrete-time setup, though the analysis in that paper is focused on a continuous version of the process.

In this paper, we show that the convergence of the WCLR rule to collusive outcomes is not robust to small independent perturbations in the profit functions of the firms (e.g., small independent variations in the cost functions, or small differences on the price received by each company). The existence of such small independent perturbations tends to push the process towards the Nash equilibrium of the one-shot game. We also explore the behaviour of WCLR against other strategies and demonstrate that WCLR is easily exploitable. Finally, we show that all these results extend to duopolies a la Bertrand with differentiated product, where firms compete in prices rather than in quantities.

The structure of the remaining of the paper is very simple: in section 2 we present and discuss the results for the Cournot model, and in section 3 we show that the obtained results also apply to duopolies a la Bertrand. Section 4 ends with the conclusions.

2. Competition in quantities: Cournot model

In this section we analyse a Cournot duopoly in which at every time step t (t = 0, 1, …) each company i (i = 1, 2) chooses a production level or quantity $[q_i]_t$. The market price $[p]_t$ is the same for both companies and it depends on the total quantity produced by the two firms. The amount $[q_i]_t$ is produced on period t with a cost function $C(q)$. The profit for each company on period t is $[\pi_i]_t = [p]_t \cdot [q_i]_t - C([q_i]_t)$. Incremental values are naturally defined as $[\Delta\pi_i]_t \equiv [\pi_i]_t - [\pi_i]_{t-1}$, for $t > 0$, and initial values at time step 0 are $[\Delta\pi_i]_0 = 0$, and $[\Delta q_i]_0 = 0$.

Let us also define $[s_i]_t \equiv \text{sign}([\Delta q_i]_t \cdot [\Delta\pi_i]_t)$. Note that $[s_i]_t$ is equal to +1 if the last changes in $[q_i]_t$ and $[\pi_i]_t$ took place in the same direction, and $[s_i]_t$ is equal to –1 if such changes went in opposite directions.

For each company i, the production levels are calculated as $[q_i]_{t+1} = max([q_i]_t + [\Delta q_i]_{t+1}, 0)$, starting with some initial positive production level $[q_i]_0$ at time step 0. The decision rule WCLR used to calculate the production increments $[\Delta q_i]_{t+1}$ is implemented as follows:

WCLR Rule:

- If $t = 0$ or $[s_i]_t = 0$, then $[\Delta q_i]_{t+1}$ takes one random value out of the set $\{-\delta_i, 0, \delta_i\}$, where δ_i is the step size.

- Otherwise, $[\Delta q_i]_{t+1} = \delta_i \cdot [s_i]_t$.

It is also assumed that the process includes some "noise" such that, with a small probability ε for each company in every period, the company will deviate from the value prescribed above for $[\Delta q_i]_{t+1}$ and will take a random choice out of the set $\{-\delta_i, 0, \delta_i\}$. This "decision noise" can represent occasional mistakes or experimentation.

Huck et al. (2003) prove that, with $\delta_i = \delta$, under rather general conditions, if the step size δ and the noise level ε are sufficiently small (but strictly positive), in the long run the process $[q_1, q_2]_t$ will spend most of the time in a small neighbourhood around the collusive outcome, and their simulations show a quick convergence to that situation. The remaining of this section is devoted to show that this convergence can be very sensitive to small independent perturbations in the profit functions of the firms. The reader can run all the simulations reported here using the online model at http://luis.izqui.org/models/wc-lr-cournot/. The computer model has been implemented in NetLogo (Wilensky, 1999).

2.1 The WCLR rule in the Cournot model with noise

For illustrative purposes we consider a linear inverse demand function: $p = max(100 - (q_1 + q_2), 0)$ and a quadratic cost function: $C[q] = 10q + 0.1q^2$. In this situation, the collusive value for the production of each company, characterized by the first-order conditions $\frac{\partial(\pi_1+\pi_2)}{\partial q_i} = 0$, is $q_i = 21.43$, which corresponds to a price level $p = 57.14$. The Cournot equilibrium, characterized by the equations $\frac{\partial \pi_i}{\partial q_i} = 0$, is $q_i = 28.13$, corresponding to a price level $p = 43.75$.

We also set $\delta_i = 0.1$ and $\varepsilon = 0.01$. Initial levels of production $[q_i]_0$ are set randomly in the range [0, 50], but note that the model is ergodic (since $\varepsilon > 0$); thus, its long-run behaviour does not depend on initial conditions.

Departing from the baseline scenario above, we study the sensitivity of the model to three types of noise:

1. "Decision noise", characterised by the parameter ε, as described above.

2. "Cost noise", characterised by the parameter ε_{cost} , and implemented by altering each firm's base cost according to the following formula:

$$C_i[q_i] = (10q_i + 0.1q_i^2) \cdot (1 + \varepsilon_{cost} \cdot \mathrm{U}_i[-1,1])$$

where $\mathrm{U}_i[-1,1]$ denotes a continuous uniform random variable with range [–1,1].

3. "Price noise", characterised by the parameter ε_{price} , and implemented by giving each firm i a price p_i according to the following formula:

$$p_i = p \cdot (1 + \varepsilon_{price} \cdot \mathrm{U}_i[-1,1])$$

where p is the price that corresponds to the total level of output using the inverse demand function. This modified model represents small differences in the price that each company gets for its products, which can be due to a number of different reasons, such as random deviations in the quality of the products of a company with respect to the average quality, different times of arrival at the market (which would allow for some variability in demand), different intermediaries with variable commissions, existence of local markets (which would allow for some variability in price), etc.

Fig. 1 below shows a representative run for each of the three types of noise[3]. In the absence of cost noise or price noise, the WCLR rule leads to a narrow area around the collusive outcome, as already shown by Huck et al. (2003). In stark contrast, small independent perturbations in the cost function or in the price function seem to destabilise the collusive outcome and push the simulation towards the Cournot equilibrium. The sensitivity of the model to perturbations in price seems to be greater than the sensitivity to perturbations in cost. This is possibly not surprising, given that, if income is greater than cost (as in the simulations shown below), a 1% variation in income (or price) has a greater effect on profit than a 1% variation in cost. The sensitivity to prices and costs can be shown to be basically the same when income and cost are approximately equal (Izquierdo and Izquierdo, 2015).

[3] Note that the simulation runs with "cost noise" or "price noise" have $\varepsilon = 0.01$ too, as prescribed in the baseline scenario.

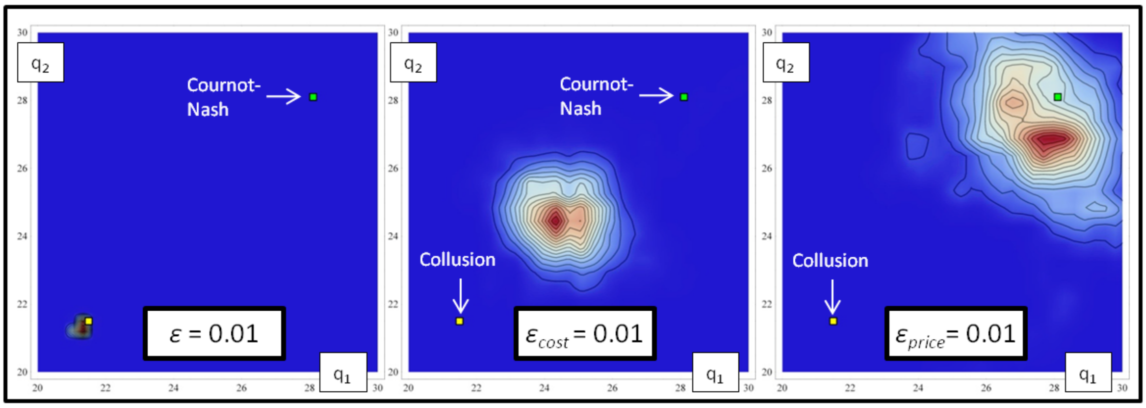

Fig. 1. **Density Histograms of the quantities produced by each firm [q_1, q_2] in one representative simulation run of 100 000 time steps. The left-most histogram shows a baseline scenario. The histogram in the centre corresponds to a simulation run with a 1% cost noise added to the baseline scenario, whilst the right-most histogram shows a simulation run with a 1% price noise added to the baseline scenario**

To study this effect rigorously, we conducted a computational experiment where we explored different values of ε, ε_{cost}, and ε_{price}. For each value of these variables we conducted 100 simulation runs, and for each of the runs we computed the average price in the simulation (taken over 10^5 time steps, and neglecting the first 10^4 time steps). Fig. 2, Fig. 3 and Fig. 4 below show the results obtained.

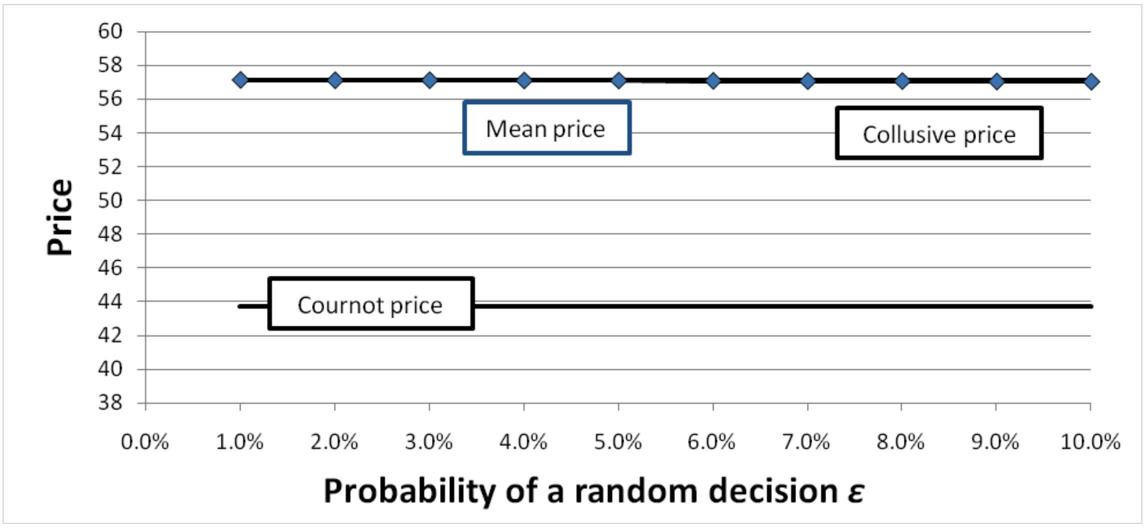

Fig. 2. The blue diamonds show, for each value of the probability of a random decision ε, the mean of 100 prices obtained from 100 independent simulation runs otherwise parameterised as in the baseline case. The price obtained from each simulation run is the average price in that simulation (taken over 10^5 time steps, and neglecting the first 10^4 time steps). The difference between the minimum average price and the maximum average price across simulations was less than 0.1 in all cases

Fig. 2 shows that the WCLR rule leads to collusive outcomes even if the probability of a random decision is fairly high. Fig. 3, in contrast, shows that small perturbations in the cost functions of the firms destabilise the collusive outcome and push the process towards the Cournot-Nash equilibrium of the one-shot game. In the same spirit, Fig. 4 shows that the sensitivity of the model to small perturbations in prices is even higher, and the collusive outcome is completely destabilised in favour of the Cournot-Nash equilibrium for values of the price noise as low as 1%. The same qualitative results are obtained with other noise distributions (see Izquierdo and Izquierdo (2015) for experiments with the normal distribution).

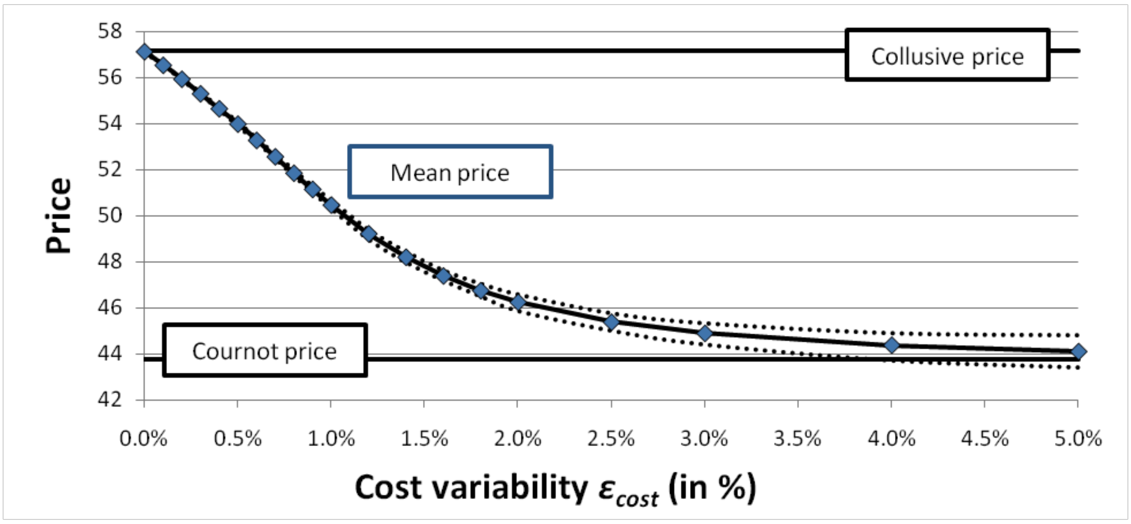

Fig. 3. **The blue diamonds show, for each value of the cost noise parameter ε_{cost}, the mean of 100 prices obtained from 100 independent simulation runs otherwise parameterised as in the baseline. The price obtained from each simulation run is the average price in that simulation (taken over 10^5 time steps, and neglecting the first 10^4 time steps). The dashed lines join the minimum average prices and the maximum average prices across simulations**

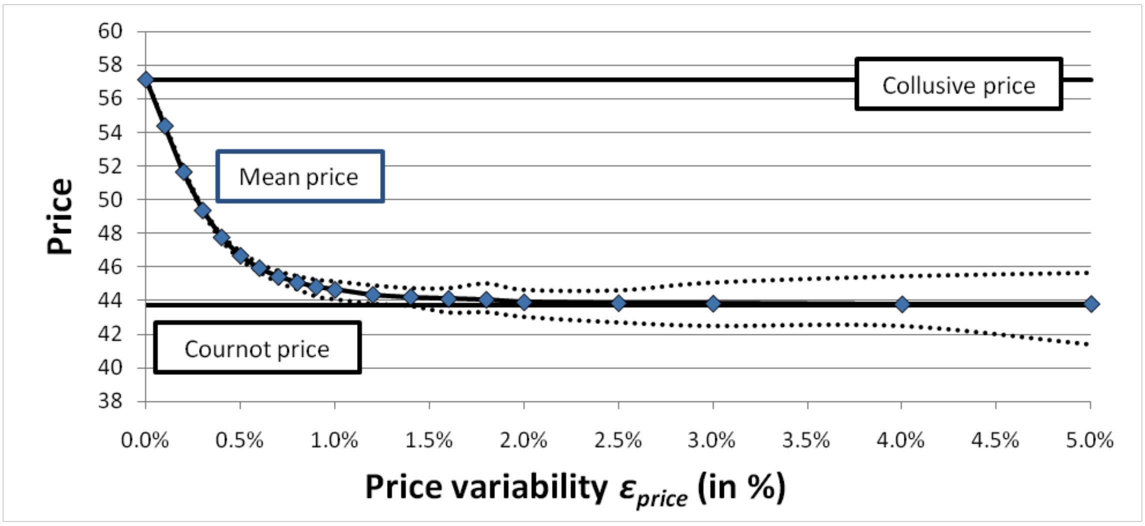

Fig. 4. **The blue diamonds show, for each value of the price noise parameter ε_{price}, the mean of 100 prices obtained from 100 independent simulation runs otherwise parameterised as in the baseline. The price obtained from each simulation run is the average price in that simulation (taken over 10^5 time steps, and neglecting the first 10^4 time steps). The dashed lines join the minimum average prices and the maximum average prices across simulations**

Why is the WCLR rule so robust to "decision noise", but so sensitive to "cost noise" and "price noise"? The answer lies in the fact that these noises produce perturbations that are fundamentally different in nature, in magnitude and in frequency.

Before analysing these three aspects in detail, it is important to understand why the collusive outcome is destabilised. The stability of the collusive outcome induced by the WCLR rule relies on a very specific sequence of coordinated moves conducted by the WCLR firms. Such deterministic pattern of moves is thoroughly described and analysed by Huck et al. (2003). For our purposes, it suffices to understand that this deterministic pattern of moves on which collusion depends is executed by the WCLR firms with exquisite precision and synchrony, like a clockwork dance (as long as there is no noise, naturally). We argue below that if such particular pattern of deterministic moves is broken with some stochasticity, WCLR firms, individually, behave more like maximizers (recall that, after all, WCLR is a gradient optimization method) and, as such, they tend to best respond to each other, i.e. approach Cournot-Nash. The intuition behind this argument is illustrated in Fig. 5 below.

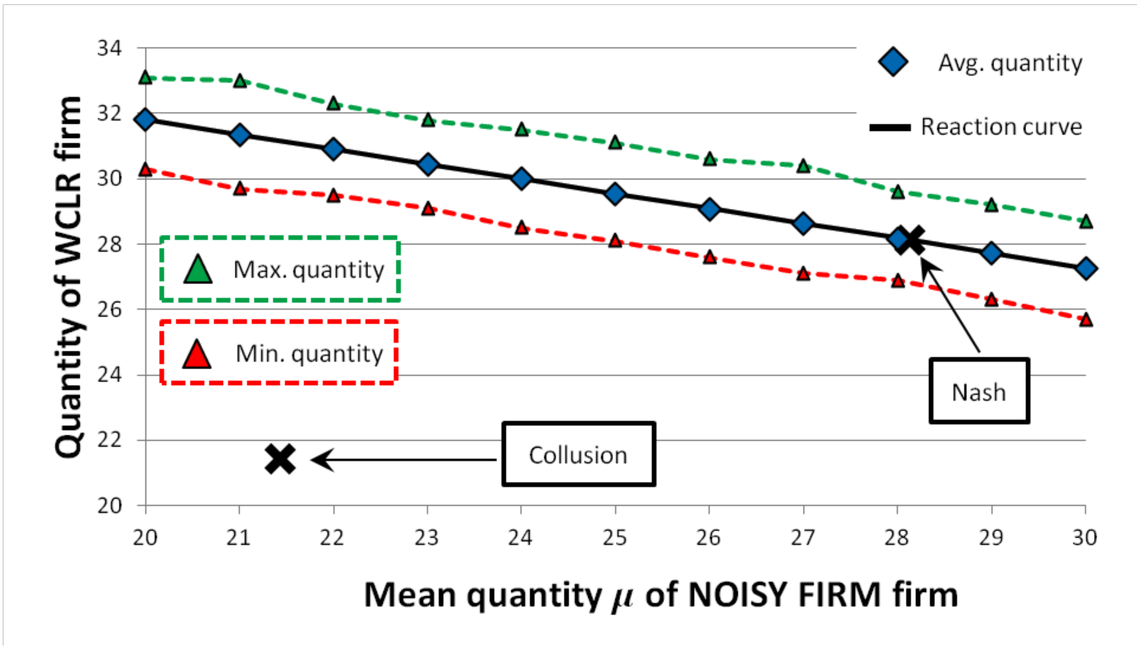

Fig. 5. The graph shows the quantities chosen by a WCLR firm playing against strategy NOISY FIX. The blue diamonds show, for each value of the mean quantity □ produced by NOISY FIX, the average of 100 quantities obtained from 100 independent simulation runs otherwise parameterised as in the baseline. Standard errors are below $2.5 \cdot 10^{-4}$ in all cases. The quantity obtained from each simulation run is the average quantity in that simulation (taken over 10^5 time steps, and neglecting the first 10^4 time steps). The dashed lines join the minimum quantities (red triangles) and the maximum quantities (green triangles) observed across all simulations and across all time steps (neglecting the first 10^4 time steps) for each value of □. The reaction curve is drawn as a solid black line.

Fig. 5 shows the quantities chosen by a WCLR firm playing against a strategy that we call NOISY FIX. NOISY FIX produces one of three possible quantities {□-□, □, □+□} with equal probability in each time step. The environment is otherwise parameterised as in the baseline scenario. Fig. 5 also shows the reaction curve of WCLR as a function of the mean quantity □ produced by NOISY FIX. It is clear that WCLR is able to best respond to NOISY FIX.[4] This suggests that in a noisy environment such that firms end up varying their production level around any certain quantity *somewhat randomly* (as opposed to following a deterministic pattern), WCLR firms will tend to best respond to each other and thus, they will tend to approach the Cournot-Nash equilibrium.

The following explains why, in general, “cost noise” and “price noise” have a much greater potential to break the deterministic so-called *clockwork dance* than “decision noise”. Because of that, they are more effective in pushing the dynamics away from collusion, and in leading them towards the Cournot-Nash equilibrium within a wide parameter range.

Let us first examine the effect of “decision noise”. Note that this type of noise may affect a firm’s decision only with a small probability ε, i.e. in most time steps “decision noise” has no influence whatsoever on the dynamics of the model. Whenever this type

---

[4] Note that WCLR is also able to best respond to strategy FIX, who always sets the same fixed quantity.

of noise does change a firm's decision (i.e. it alters the *clockwork dance*), it has an impact only on the decision at the time step at which it occurs and, as proved by Huck et al. (2003), the process goes back towards collusion automatically in two time steps. This explains why the collusive outcome is so robust to "decision noise". The *clockwork dance* gets altered only sporadically, and when it does, it is fully restored in two time steps.

By contrast, "cost noise" and "price noise" occur more frequently and their effect on the model is fundamentally different from that induced by "decision noise". As for frequency, note that both "cost noise" and "price noise" affect firms' profits every single time step regardless of the magnitude of ε_{cost} or ε_{price} (as long as they are positive), but these noises may not alter firms' decisions. The change in profit caused by either of these noises will make a firm switch its decision if and only if the change is sufficiently large as to modify the sign of $\Delta\pi$. Intuitively, this requirement means that perturbations in prices or in costs have an impact on profits that is at least comparable to the changes in profits caused by one-time-step adjustments in production.

To formally understand the circumstances under which such a requirement may be fulfilled, let us compare the change in profit in one time step in the absence of noise $\Delta^0\pi$, with the one-time-step change in profit strictly induced by the presence of noise $\Delta^\varepsilon\pi$. To be clear, in a model with noise, the total change in profit $\Delta\pi$ would be $\Delta\pi = \Delta^0\pi + \Delta^\varepsilon\pi$. We compute these quantities below:

$$\Delta^0\pi = pq - C[q] - (p_0 q_0 - C_0[q_0])$$

$$\Delta^{\varepsilon_{cost}}\pi = pq - C[q]\cdot(1+\varepsilon_{cost}\cdot \mathrm{U}[-1,1]) - (pq - C[q]) = C[q]\cdot\varepsilon_{cost}\cdot\mathrm{U}[-1,1]$$

$$\Delta^{\varepsilon_{price}}\pi = pq\big(1+\varepsilon_{price}\cdot\mathrm{U}[-1,1]\big) - C[q] - (pq - C[q]) = pq\cdot\varepsilon_{price}\cdot\mathrm{U}[-1,1]$$

where variables with sub index 0 refer to the initial point, whilst p and q refer to the price and the quantity at the final point. To appreciate the relative order of magnitude in a particular setting of the profit increments defined above, we compute them for our baseline scenario. In our setting, at the collusive outcome, we have $|\Delta^0\pi| < 4.3$ for $\delta = 0.1$, and $|\Delta^0\pi| < 0.43$ for $\delta = 0.01$, whilst $\Delta^{\varepsilon_{cost}}\pi = 260.2\cdot\varepsilon_{cost}\cdot\mathrm{U}[-1,1]$ and $\Delta^{\varepsilon_{price}}\pi = 1224.5\cdot\varepsilon_{price}\cdot\mathrm{U}[-1,1]$. This shows that, if step sizes are not too large, both cost and price noise have great potential to frequently change a firm's decision (and thus break the *clockwork dance*). In other words, the noise-to-signal ratio $\Delta^\varepsilon\pi/\Delta^0\pi$ can naturally be very large, considering the effect (on profit) of cost or price variations as the noise, and the effect induced by one-step quantity increments as the signal. A large noise-to-signal ratio means that a great amount of stochasticity is introduced in the model.

In more general terms, note that –assuming the profit function is continuous– the one-time-step change in profit caused by adjustments in production $\Delta^0\pi$ becomes vanishingly small as the step size goes to zero, i.e. $\lim_{\delta\to 0}\Delta^0\pi = 0$.[5] In contrast, note that neither $\Delta^{\varepsilon_{cost}}\pi$ nor $\Delta^{\varepsilon_{price}}\pi$ approach zero as δ goes to zero. This effectively means that, as the step size δ gets smaller, the potential of "cost noise" and "price noise" to change firms' decisions –and thus affect the dynamics of the model– becomes greater, regardless of the shape of the demand or the cost functions. This fact is illustrated in Fig. 6 below.

---

[5] In our setting, noting that $p = p_0 + \{-2\delta, -\delta, 0, \delta, 2\delta\}$ and $q = q_0 + \{-\delta, 0, \delta\}$, it is possible to derive the following bound: $|\Delta^0\pi| \leq (|p_0 - 10| + 2.2q_0 + 2.1\delta)\cdot\delta$.

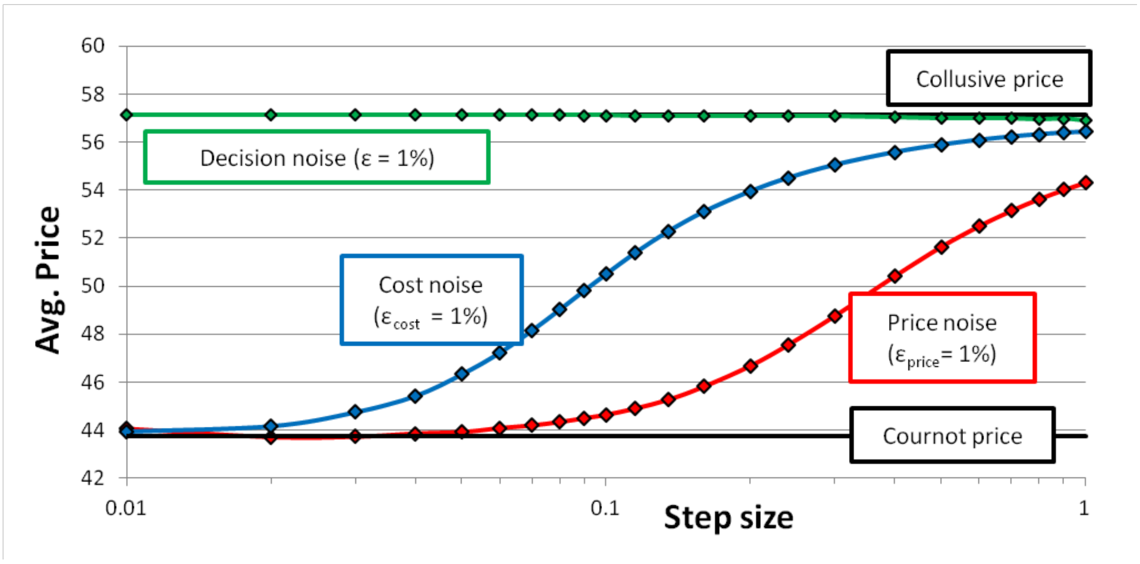

Fig. 6. The diamonds show, for different values of the step size, the average prices obtained with decision noise $\varepsilon = 1\%$ only (green), with cost noise $\varepsilon_{cost} = 1\%$ added (blue) and with price noise $\varepsilon_{price} = 1\%$ added (red). Simulation runs have been otherwise parameterised as in the baseline. The price obtained from each simulation run is the average price in that simulation (taken over 10^5 time steps, and neglecting the first 10^5 time steps). Each diamond represents the average of 100 independent simulation runs. All standard errors are below 0.12. The collusive price and the Cournot price are also included (solid black lines) as a reference

Finally, there is yet another key difference between the three types of noise. Note that "decision noise" does not affect the profit landscape (i.e. the function that maps quantities to profit). This is crucial because the stability of the collusive outcome induced by the WCLR rule relies on a steady profit landscape. To be clear, Huck et al. (2003) prove and explain that, in a steady profit landscape, following any variation in the quantities chosen by the firms, the WCLR rule induces, after at most two time steps, a sequence of moves leading towards the collusive outcome. "Decision noise" breaks the *clockwork dance* at the particular time steps when it alters a firm's decision (an event that occurs with small probability) but, in at most two periods after such a perturbation, the firms are engaged again in a series of coordinated and perfectly synchronized moves towards or around collusion. In contrast, "cost noise" and "price noise" do shake the profit landscape back and forth every time step, altering the relation between $[\Delta q_i]_t$ and $[\Delta \pi_i]_t$ not only in the current period but also in the following one (since $[\pi_i]_t$ is used in the computation of both $[\Delta \pi_i]_t$ and $[\Delta \pi_i]_{t+1}$). Because of this, every single perturbation does not only have the potential to affect the decision at the time step it occurs, but it can also have an direct impact on subsequent decisions. This deeper type of alteration, which transcends the time step at which it occurs, constitutes a greater source of miscoordination that can further disturb the *clockwork dance* on which the stability of the collusive outcome relies.

2.2 Correlated perturbations

In this section, we show that the destabilizing factor of the variability in cost or price is not so much the existence of the perturbations, but the fact that they are somewhat independent or uncorrelated between the firms. To illustrate this, here we consider the effect of *correlated* perturbations. Correlations would be observed in the real world if there were variations in costs or in the demand function that affected both companies in a similar way (for instance, seasonal demand variability). To study such situations, we model a price perturbation for each firm which is composed of both a common factor α·U[–1,1] –with weight α– and an independent factor (1 – α)·U_i[–1,1] –with weight (1 – α)–, according to the formula:

$$p_i = p \cdot (1 + \varepsilon_{price} \cdot R^{\alpha}_{i})$$

where

$$R^{\alpha}_{i} = \alpha \cdot U[-1,1] + (1 - \alpha) \cdot U_i[-1,1].$$

Thus, parameter α is a measure of the correlation between the perturbations of each firm. Extreme value α = 0 represents completely uncorrelated perturbations (as analyzed above), and extreme value α = 1 represents full correlation (where the perturbations for each firm are exactly the same). Fig. 7 below shows that the more correlated perturbations are, the less impact they have on destabilising the collusive outcome. As explained before, perturbations affect the dynamics of the model mainly through the generation of miscoordination that breaks the *clockwork dance* played by the firms; thus, it is natural that the impact of correlated noise, which does not cause so much miscoordination, is less acute than the effect of uncorrelated perturbations.

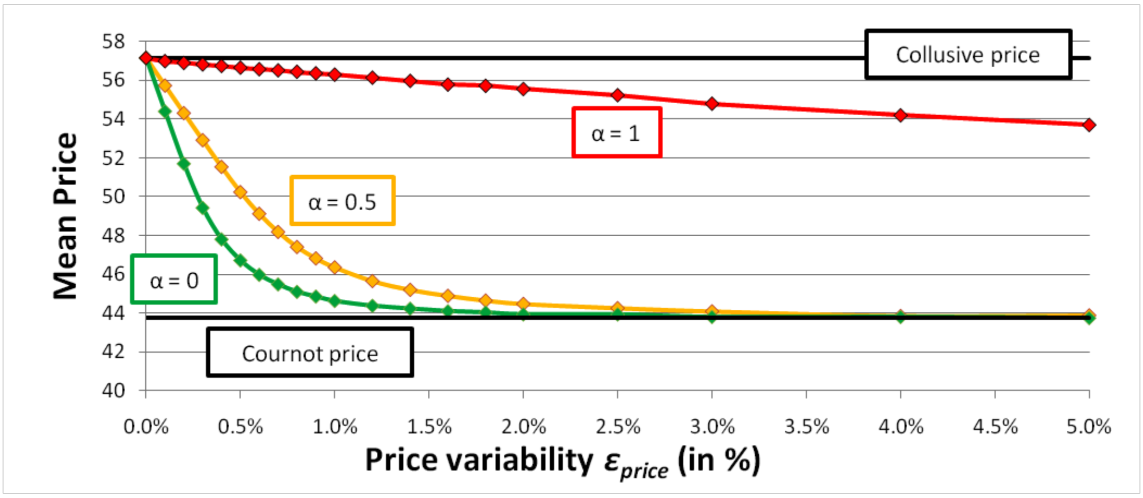

Fig. 7. The diamonds show, for each value of the price noise parameter ε_{price} and different values of α, the mean of 100 prices obtained from 100 independent simulation runs otherwise parameterised as in the baseline. The price obtained from each simulation run is the average price in that simulation (taken over 10^5 time steps, and neglecting the first 10^4 time steps)

2.3 More than two competing firms

The simulation results of Huck et al. (2003) in symmetric oligopolies with more than two competing firms (up to ten) and some small decision noise also showed convergence of the WCLR rule to collusive outcomes. We show in Fig. 8 below that, as in the duopoly case, the existence of small independent perturbations in the price that each company obtains also destabilises the collusive outcome and pushes the process towards the Nash equilibrium of the one-shot game. Uncorrelated perturbations in cost have the same qualitative effect, so they are not shown here.

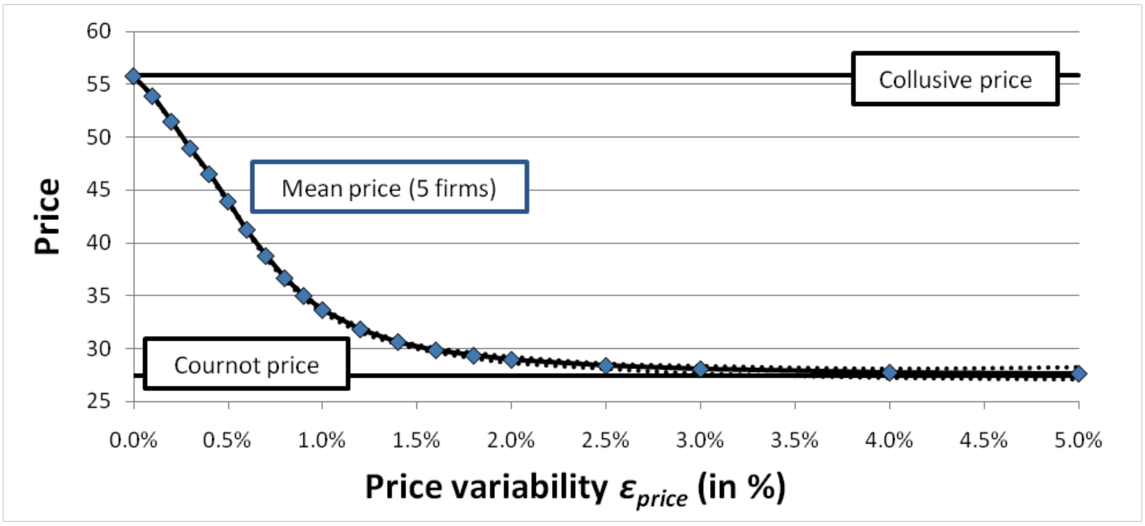

Fig. 8. The blue diamonds show, for each value of the price noise parameter ε_{price}, the mean of 100 prices obtained from 100 independent simulation runs otherwise parameterised as in the baseline, in an oligopoly with 5 competing firms. The price obtained from each simulation run is the average price in that simulation (taken over 10^5 time steps, and neglecting the first 10^4 time steps). The dashed lines join the minimum average prices and the maximum average prices across simulations

It should also be noted that, as the number of competing firms increase, the one-shot Cournot-Nash equilibrium gets closer to the outcome predicted under the assumption of perfect competition, so, as the number of firms increase, the WCLR rule with independent cost or price perturbations leads to market prices and production levels which approach those predicted by the perfect competition theory. Fig. 9 below shows the effect of uncorrelated 2% price perturbations in oligopolies with different number of firms. The results also show an increasing difference between the simulated price and the Cournot price as the number of firms in the market increases, which may be due to the decreasing marginal importance of one firm in the market as the number of firms in the market increases.

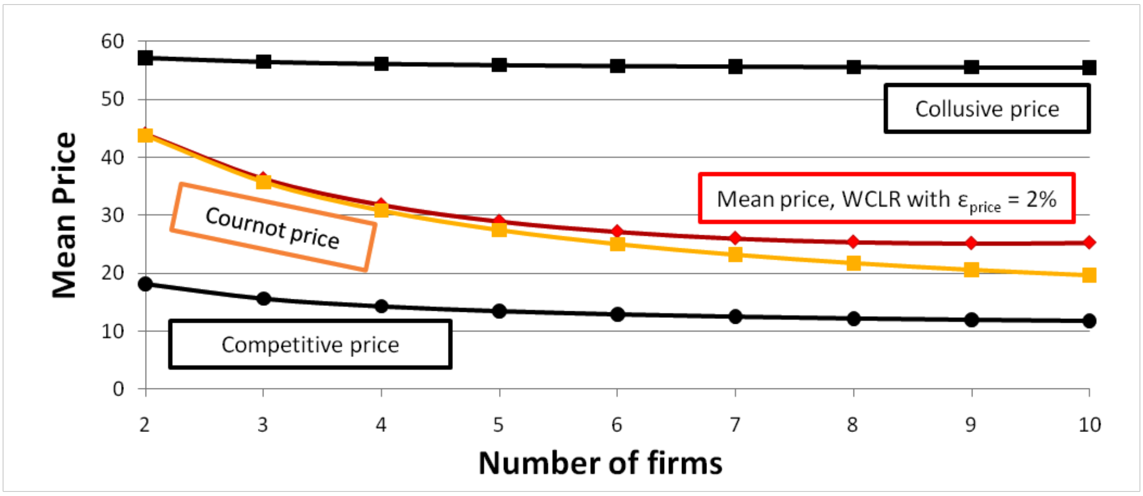

Fig. 9. The diamonds show, for a price noise parameter ε_{price} = 2% and different number of firms, the mean of 100 prices obtained from 100 independent simulation runs otherwise parameterised as in the baseline. The price obtained from each simulation run is the average price in that simulation (taken over 10^5 time steps, and neglecting the first 10^4 time steps)

2.4 Better responses and exploitability of WCLR

This section addresses two related (but distinct) questions:

a) Is WCLR an optimal response to itself?

b) Is WCLR generally able to respond optimally to other strategies?

The answer to both questions is negative.

Let us start with a simple observation which shows that, in general, WCLR is not a best response to itself. Consider the strategy FIX, which simply chooses a production level and keeps it fixed. Strategy FIX will make a WCLR player move towards its reaction curve. Consequently, FIX can choose the quantity corresponding to the first mover of the Stackelberg game, and (assuming a small step size for the WCLR player) it will approximately obtain the corresponding profit, which can be greater than the collusive profit.[6] Thus, it is clear that, in general, WCLR is not a best response to itself.

To illustrate the fact that WCLR does not generally respond optimally to other strategies either, we need to consider strategies slightly more sophisticated than FIX, i.e. strategies that can set different quantities at different time steps. Incidentally, the following analysis will also serve as a second illustration of the fact that there are strategies that can perform better against WCLR than WCLR itself.

Let us focus on the set of strategies that can end up in a cycle containing 3 distinct points when playing against WCLR (assuming no noise). To be clear, we will consider a strategy –henceforth called MAXIMIZER– which, playing against WCLR, can indefinitely cycle through 3 points (1; 2; 3), where MAXIMIZER chooses quantities ($q_1^M; q_2^M;\ q_3^M$) respectively and WCLR chooses quantities ($q_1^W; q_2^W;\ q_3^W$) respectively. The actual sequence in the cycle played by MAXIMIZER and WCLR will be 1,2,1,3;1,2,1,3;1,2,1,3;… The quantities chosen by each strategy in the cycle are defined as the solution of the following constrained optimization problem:

$$Max_{\{q_i^M;q_i^W\}_{i=1,2,3}}\ 2\cdot\pi^M[p_1,q_1^M]+\pi^M[p_2,q_2^M]+\pi^M[p_3,q_3^M]$$

Subject to:

$$\pi^W[p_2,q_2^W]<\pi^W[p_1,q_1^W]; \qquad q_2^W=q_1^W+\delta;$$
$$\pi^W[p_3,q_3^W]<\pi^W[p_1,q_3^W]; \qquad q_3^W=q_1^W-\delta;$$
$$p_i=InverseDemand[q_i^W+q_i^M];\ q_i^M\geq 0; q_i^W\geq 0;\ p_i\geq 0; \qquad i=1,2,3.$$

where super index *M* refers to variables concerning MAXIMIZER, whilst super index *W* refers to variables concerning strategy WCLR. Getting WCLR into the cycle described above from any initial condition is unproblematic as long as q_1^W lies in WCLR's reaction curve.[7]

Fig. 10 below shows that MAXIMIZER can obtain profits very close to monopolistic profits and is able to effectively push WCLR out of the market[8] when playing the aforementioned cycle against WCLR. It is also clear that WCLR is not responding op-

---

[6] The baseline scenario is an example where Stackelberg's first-mover profit is not greater than the collusive profit; however, if the demand function is changed to be $p = max(100 - (q_1 + q_2)^2, 0)$, then Stackelberg's first-mover profit is indeed greater than the collusive profit.

[7] For the sake of clarity, we are ignoring the fact that, strictly speaking, the difference between WCLR's initial quantity and q_1^W should be a multiple of δ. If this were not the case, the optimization problem should be formulated in discrete terms.

[8] In all points shown in Fig. 11, WCLR is obtaining small but strictly positive profits.

timally to MAXIMIZER. To see this, consider strategy OPTIMAL, which reacts optimally to each quantity q_i^M that MAXIMIZER sets in the cycle. Fig. 10 shows that the profits obtained by OPTIMAL against MAXIMIZER are clearly higher than those obtained by WCLR against MAXIMIZER.

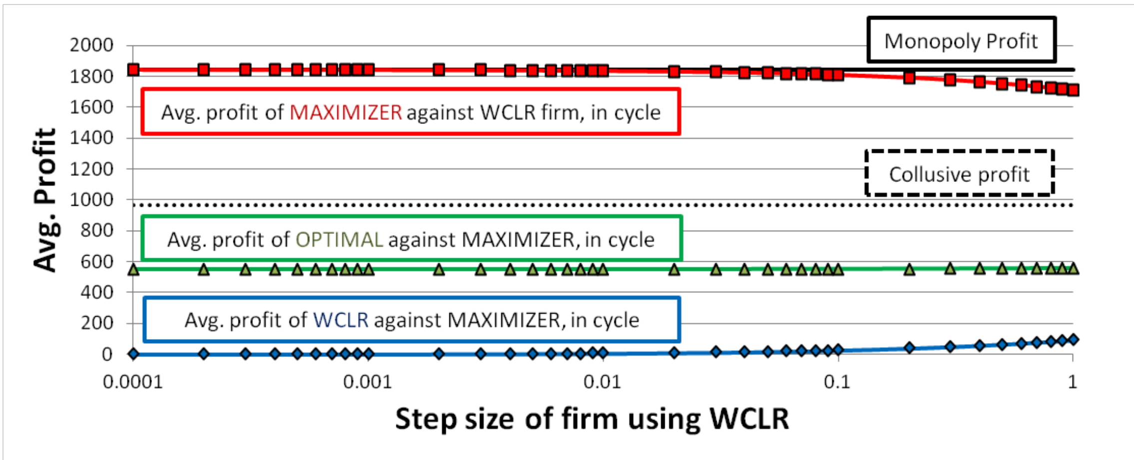

Fig. 10. For the baseline parameterisation without noise, and different step sizes for the WCLR firm, the graph shows the average (per-period) profit of MAXIMIZER against WCLR (red squares), the average (per-period) profit of strategy WCLR against MAXIMIZER (blue diamonds), and the average (per-period) profit of strategy OPTIMAL against MAXIMIZER (green triangles) in the type of cycle described in the main text. The monopolistic profit (solid black line) and the collusive profit (dashed black line) are also included as a reference

Thus, it is clear that WCLR is not generally an optimal response to itself (e.g. MAXIMIZER is definitely better) and that WCLR is generally not able to respond optimally to other strategies (like e.g. MAXIMIZER).

Interestingly, WCLR itself can exploit other WCLRs that play with smaller step sizes. This is clearly illustrated in the leftmost part of Fig. 11 below, where it is shown that a WCLR player with step size $\delta_1 = 0.1$ can obtain profits close to monopolistic against a WCLR firm with step size $\delta_2 = 0.001$.

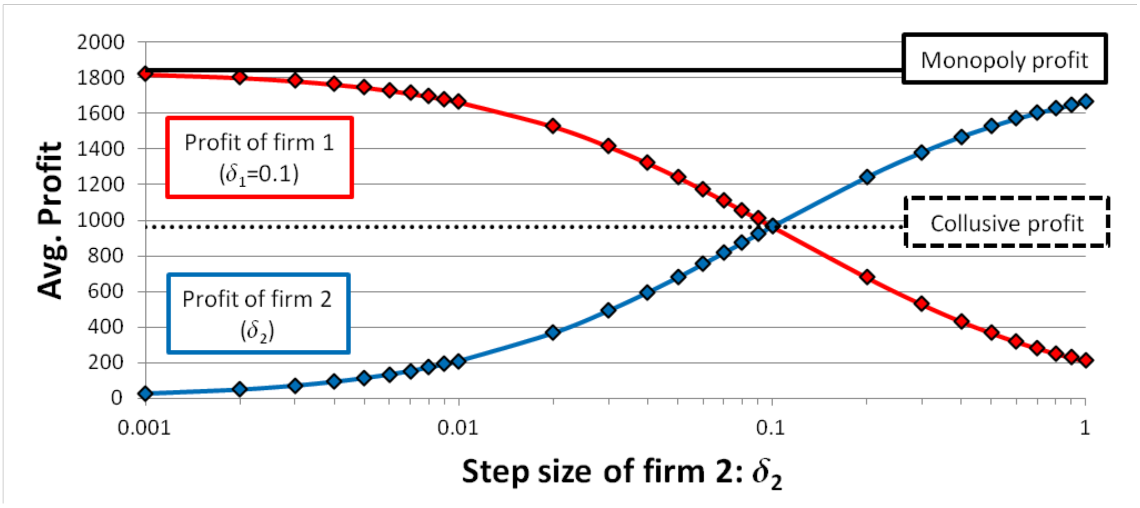

Fig. 11. The red diamonds show the average profit of WCLR firm 1 (with step size $\delta_1 = 0.1$) and the blue diamonds show the average profit of WCLR firm 2 (with step size δ_2) when playing against each other, for different step sizes δ_2 for WCLR firm 2. The profits shown for each value of δ_2 are the averages of 100 profits obtained from 100 independent simulation runs otherwise parameterised as in the baseline. The profit obtained from each simulation run for each firm is the average profit in that simulation taken over 10^5 time steps and neglecting the first 10^5 time steps. Standard errors are below 0.25 in all cases. The monopolistic profit (solid black line) and the collusive profit (dashed black line) are also included as a reference

3. Competition in prices: Duopoly with differentiated products

In this section, we explore the behaviour of the WCLR rule in a (Bertrand-like) duopoly with differentiated products, in which the decision variable is the price level, and where each company faces a demand function that depends on both its price and its competitor's.

The considered model is a duopoly in which at every time step t (t = 0, 1, …) each company i (i = 1, 2) chooses a price level $[p_i]_t$. Each company faces a market demand $[q_i]_t$ which depends on the prices chosen by both companies. The amount $[q_i]_t$ is produced on period t with a cost function $C(q)$. The profit for each company on period t is $[\pi_i]_t = [p_i]_t \cdot [q_i]_t - C([q_i]_t)$. Incremental values are naturally defined as $[\Delta\pi_i]_t \equiv [\pi_i]_t - [\pi_i]_{t-1}$, and initial values at time step 0 are $[\Delta\pi_i]_0 = 0$, and $[\Delta p_i]_0 = 0$.

Similarly to the Cournot model, let us define $[s_i]_t \equiv \text{sign}([\Delta p_i]_t \cdot [\Delta\pi_i]_t)$. Note that $[s_i]_t$ is equal to +1 if the last changes in $[p_i]_t$ and $[\pi_i]_t$ took place in the same direction, and $[s_i]_t$ is equal to –1 if such changes went in opposite directions.

For each company i, the price levels are calculated as $[p_i]_{t+1} = max([p_i]_t + [\Delta p_i]_{t+1}, p_{min})$, starting with some initial positive price level $[p_i]_0$ at the initial time step[9]. The decision rule WCLR used to calculate the price increments $[\Delta p_i]_{t+1}$ is implemented as follows:

WCLR Rule:

- If t = 0 or $[s_i]_t = 0$, then $[\Delta p_i]_{t+1}$ takes one random value out of the set $\{-\delta_i, 0, \delta_i\}$, where δ_i is the step size.

- Otherwise, $[\Delta p_i]_{t+1} = \delta_i \cdot [s_i]_t$.

The boundary situation where the firm is not selling anything ($[q_i]_t = 0$) is taken into account by making such a firm reduce its price ($[\Delta p_i]_{t+1} = -\delta_i$). It is also assumed that the process includes some "noise" such that, with a small probability ε for each company in every period, the company will deviate from the value prescribed above for $[\Delta p_i]_{t+1}$ and will take a random choice among $\{-\delta_i, 0, \delta_i\}$. This "Decision noise" can represent occasional mistakes or experimentation.

In the following section, we show by means of simulation that, similarly to the Cournot case, with $\delta_i = \delta$, if the step size δ and the noise level ε are small, the process $[p_1, p_2]_t$ quickly converges to a small neighbourhood around the collusive outcome, and tends to remain on that area. We also show, however, that this convergence can be very sensitive to small independent perturbations in the profit functions of the firms. The reader can run all the simulations reported here using the online model at http://luis.izqui.org/models/wc-lr-bertrand/. The computer model has been implemented in NetLogo (Wilensky, 1999).

3.1 The WCLR rule in Bertrand competition with differentiated products

As in the analysis of the Cournot model, we focus here on concrete representative examples for illustrative purposes. In particular, our baseline scenario for Bertrand competition will use the quadratic cost function $C[q] = 10q + 0.1q^2$ and the symmetric linear demand functions: $q_i = max(100 - p_i + 0.5\, p_j, 0)$, $i \neq j$.

---

[9] The minimum price p_{min} is included to ensure that firms do not sell below their minimum marginal cost.

The reference theoretical points for this model are the competitive outcome (p_i = 27.27), characterized by the equations $p_i = \frac{\partial Ci}{\partial q_i}$ (together with the demand functions), the one-shot Nash equilibrium (p_i = 81.25), characterized by the equations $\frac{\partial \pi_i}{\partial p_i} = 0$, and the collusive outcome (p_i = 109.52), characterized by the equations $\frac{\partial (\pi_i + \pi_j)}{\partial p_i} = 0$.

A parallelism with the Cournot case (Huck et al., 2003), extended to allow for different step sizes, would lead to a reference theoretical point for the WCLR rule at a point characterised by the equations $\frac{\partial \pi_i}{\partial p_i} = -\frac{\delta_j}{\delta_i}\left|\frac{\partial \pi_i}{\partial p_j}\right|$, which, in the symmetric case (in which $\frac{\partial \pi_1}{\partial p_2} = \frac{\partial \pi_2}{\partial p_1}$ for $p_1 = p_2$) and with equal step size $\delta_1 = \delta_2$, coincides with the collusive solution.

As a baseline scenario, we set δ_i = 0.1 and ε = 0.01. Initial prices $[p_i]_0$ are set randomly in the range [0, 100], but note that the model is ergodic (since $\varepsilon > 0$); thus, its long-run behaviour does not depend on initial conditions. Departing from this baseline scenario, we also consider here a "Cost noise", characterised by the parameter ε_{cost}, and a "Demand noise", characterised by the parameter ε_{demand}. To be clear, the introduction of noise modifies the cost function and the demand functions in the following way:

$$C_i[q_i] = (10q_i + 0.1q_i^2) \cdot (1 + \varepsilon_{cost} \cdot \mathrm{U}_i[-1,1])$$

$$q_i = max(\,(100 - p_i + 0.5\,p_j\,)\,(1 + \varepsilon_{demand} \cdot \mathrm{U}_i[-1,1])\,,\,0) \quad i \neq j$$

Fig. 12 below shows a representative run for each of the three types of noise[10]. It indicates that the same patterns that apply to the Cournot model seem to apply also to the WCLR rule in this price-setting duopoly: convergence to an area around the collusive outcome if there are no independent perturbations (other than "decision noise"), and displacement towards the one-shot Nash equilibrium if there are (small) independent perturbations in demand or in costs. Just like in the Cournot model, the sensitivity of the model to perturbations in demand seems to be greater than its sensitivity to perturbations in costs. This is due to the same reason explained before: income is significantly greater than cost for both firms in the region of interest, so a certain percentage change in demand (and, therefore, in income) induces a greater change in profit than the same percentage change in costs. Consequently, under such positive-profit circumstances, demand variability constitutes a greater source of miscoordination than cost variability, and it therefore affects the dynamics of the model more strongly. The reader can use the online model to confirm that if fixed costs are chosen to make income and cost similar in magnitude, then the sensitivity of the model to these two types of noise –"demand noise" and "cost noise"– is more alike.

---

[10] Note that the simulation runs with "cost noise" or "demand noise" have ε = 0.01 too, as prescribed in the baseline scenario.

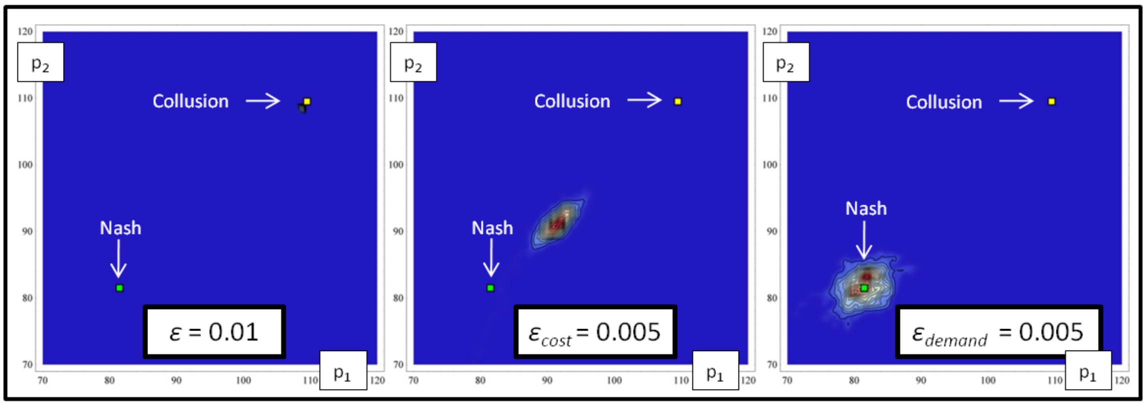

Fig. 12. Density Histograms of the prices set by each firm [p_1, p_2] in one representative simulation run of 100 000 time steps in the model with competition in prices. The left-most histogram shows a baseline scenario. The histogram in the centre corresponds to a simulation run with a 0.5% cost noise added to the baseline scenario, whilst the right-most histogram shows a simulation run with a 0.5% demand noise added to the baseline scenario

In the same spirit as before, we conducted a computational experiment where we systematically explored different values of ε_{cost}. Fig. 13 below shows that, as anticipated, small perturbations in the cost functions of the firms destabilise the collusive outcome and push the process towards the Nash equilibrium of the one-shot game. Given the symmetry in the simulation setup, we present the average prices of one firm only.

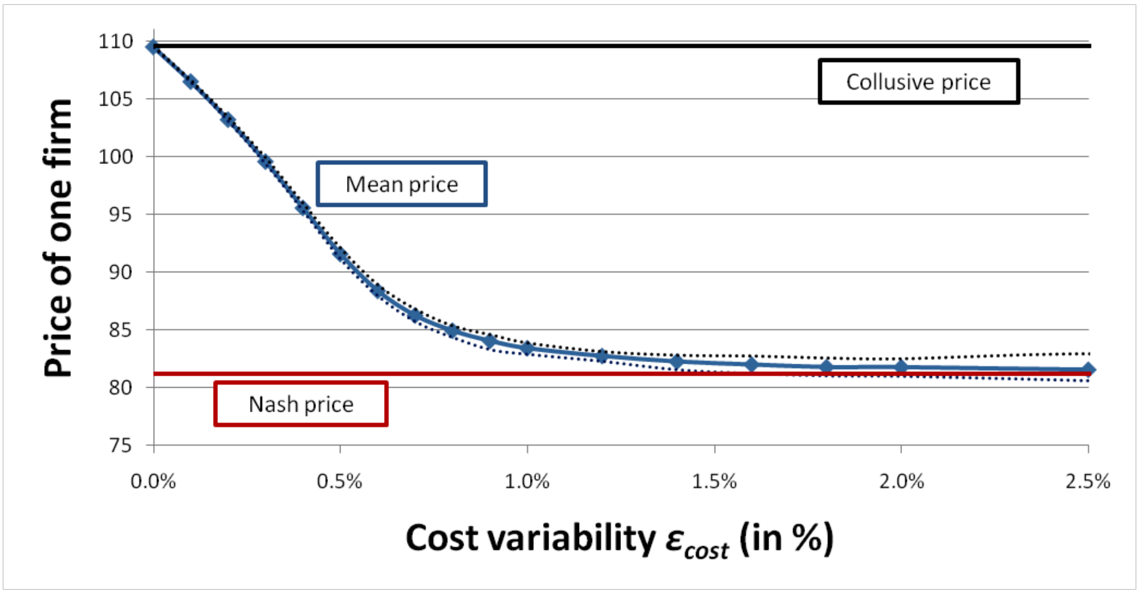

Fig. 13. The blue diamonds show, for each value of the cost noise parameter ε_{cost}, the mean of 100 prices obtained from 100 independent simulation runs otherwise parameterised as in the baseline. The price obtained from each simulation run is the average price of one of the firms in that simulation (taken over 10^5 time steps, and neglecting the first 10^4 time steps). The dashed lines join the minimum average prices and the maximum average prices across simulations

3.2 Correlated perturbations

Finally, to fully understand the mechanism through which demand and cost perturbations affect the dynamics of the WCLR price-setting duopoly with differentiated products, we also include here a parameter α that allows us to modulate the correlation between the demand perturbations received by each of the firms. Specifically, we model a demand perturbation for each firm which is composed of both a common factor $\alpha \cdot U[-1,1]$ –with weight α– and an independent factor $(1 - \alpha) \cdot U_i[-1,1]$ –with weight $(1 - \alpha)$–, according to the formula:

$$q_i = max(\ (100 - p_i + 0.5\ p_j\ )\ (1 + \varepsilon_{demand} \cdot R^{\alpha}_{i})\ ,\ 0)$$

where $i \neq j$ and

$$R^{\alpha}_{i} = \alpha \cdot U[-1,1] + (1 - \alpha) \cdot U_i[-1,1].$$

Thus, parameter α is a measure of the correlation between the demand perturbations of each firm. Extreme value α = 0 represents completely uncorrelated perturbations, whilst extreme value α = 1 represents full correlation between the demand perturbations received by each firm (potentially due, for instance, to seasonal demand variability in both products, or global factors affecting the demand of both products in the same way). Fig. 14 below shows that the more correlated perturbations are, the less impact they have on destabilising the collusive outcome. Given the symmetry in the simulation setup, we present the average prices of one firm only.

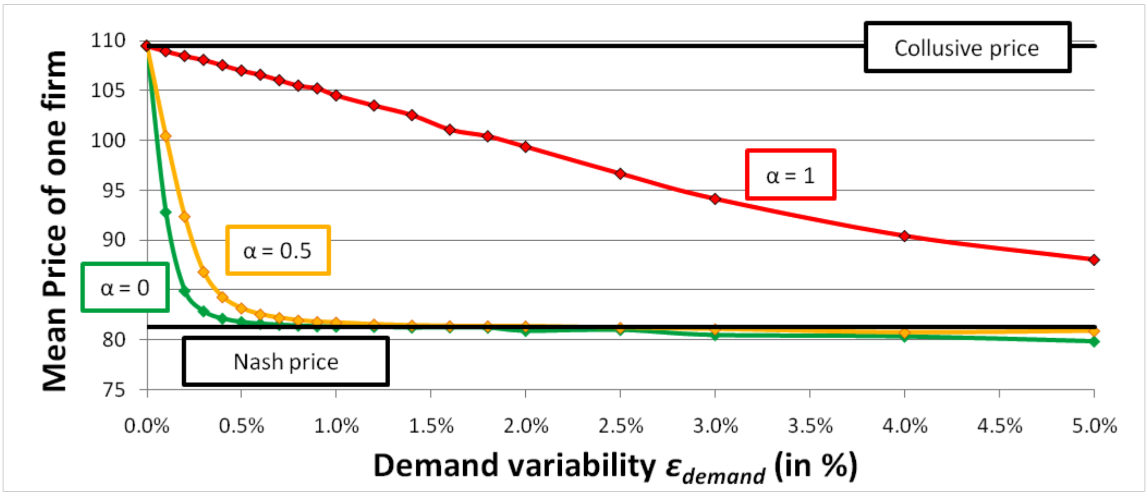

Fig. 14. The diamonds show, for each value of the demand noise parameter ε_{demand} and different values of α, the mean of 100 prices obtained from 100 independent simulation runs otherwise parameterised as in the baseline. The price obtained from each simulation run is the average price of one of the firms in that simulation (taken over 10^5 time steps, and neglecting the first 10^4 time steps)

As explained before, perturbations affect the dynamics of the model mainly through the generation of miscoordination between the firms; thus, it is natural that the impact of correlated noise, which does not cause so much miscoordination, is less acute than the effect of uncorrelated perturbations.

3.3 Better responses and exploitability of WCLR

In the same spirit as in section 2.4, we show here that WCLR is not generally an optimal response to itself or other strategies in a duopoly a la Bertrand either. For that, we also design here a MAXIMIZER strategy which, playing against WCLR without any noise, can indefinitely cycle through 3 points (1; 2; 3), where MAXIMIZER chooses prices $(p_1^M; p_2^M;\ p_3^M)$ respectively and WCLR chooses prices $(p_1^W; p_2^W;\ p_3^W)$ respectively. The actual sequence in the cycle played by MAXIMIZER and WCLR will be 1,2,1,3;1,2,1,3;1,2,1,3…, and the prices chosen by each strategy in the cycle are defined as the solution of the following constrained optimization problem:

$$Max_{\{p_i^M;p_i^W\}_{i=1,2,3}}\ 2 \cdot \pi^M[p_1^M, q_1^M] + \pi^M[p_2^M, q_2^M] + \pi^M[p_3^M, q_3^M]$$

Subject to:

$\pi^W[p_2^W, q_2^W] < \pi^W[p_1^W, q_1^W];$ $\qquad\qquad p_2^W = p_1^W + \delta;$

$$\pi^W[p_3^W, q_3^W] < \pi^W[p_1^W, q_1^W]; \qquad p_3^W = p_1^W - \delta;$$

$$q_i^M = Demand^M[p_i^M, p_i^W]; \qquad q_i^W = Demand^W[p_i^W, p_i^M]; \qquad i = 1,2,3;$$

$$p_i^M \geq 0; q_i^M \geq 0;\ p_i^W \geq p_{min}[q_i^W]; q_i^W \geq 0;$$

where super index M refers to variables concerning MAXIMIZER, whilst super index W refers to variables concerning strategy WCLR. Getting WCLR into the cycle described above from any initial condition is unproblematic as long as q_i^W lies in WCLR's reaction curve.

Fig. 15 below shows that MAXIMIZER can obtain profits much higher than collusive profits and is able to effectively push WCLR out of the market[11] when playing against WCLR. It is also clear that WCLR is not responding optimally to MAXIMIZER. To see this, consider strategy OPTIMAL, which reacts optimally to each price p_i^M that MAXIMIZER sets in the cycle. Fig. 15 shows that the average profit obtained by OPTIMAL against MAXIMIZER is clearly higher than the average profit obtained by WCLR against MAXIMIZER.

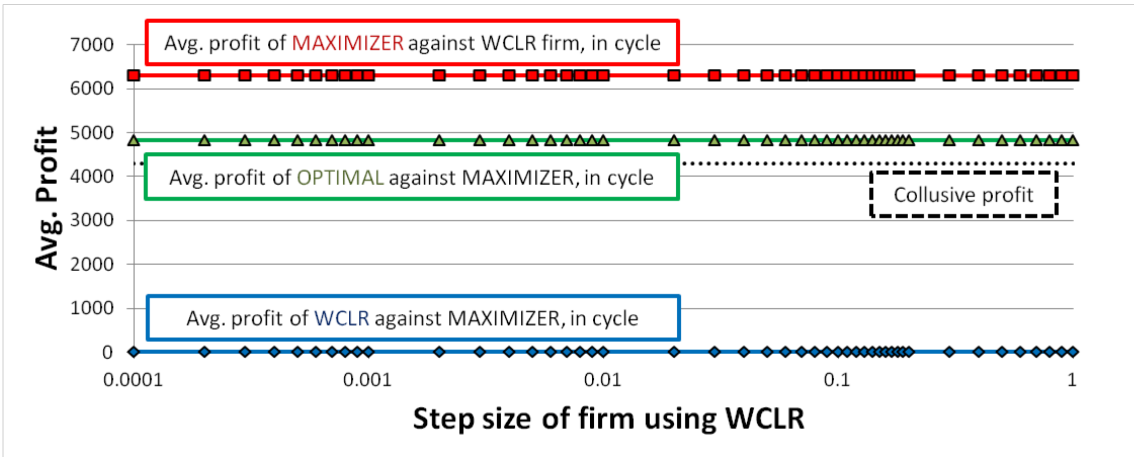

Fig. 15. For the baseline parameterisation without noise, and different step sizes for the WCLR firm, the graph shows the average (per-period) profit of MAXIMIZER against WCLR (red squares), the average (per-period) profit of strategy WCLR against MAXIMIZER (blue diamonds), and the average (per-period) profit of strategy OPTIMAL against MAXIMIZER (green triangles) in the type of cycle described in the main text. The collusive profit (dashed black line) is also included as a reference

Thus, it is clear that, as in the Cournot case, WCLR is not generally an optimal response to itself (e.g. MAXIMIZER is definitely better) and that WCLR is generally not able to respond optimally to other strategies (like e.g. MAXIMIZER). Also as in the Cournot model, WCLR can exploit other WCLRs that play with smaller step sizes. This is clearly illustrated in Fig. 16 below, where it is shown that the ratio between step sizes needed to exploit another WCLR player is smaller in the Bertrand case than in the Cournot model (see Fig. 11).

[11] In all points shown in Fig. 16, WCLR is obtaining small but strictly positive profits.

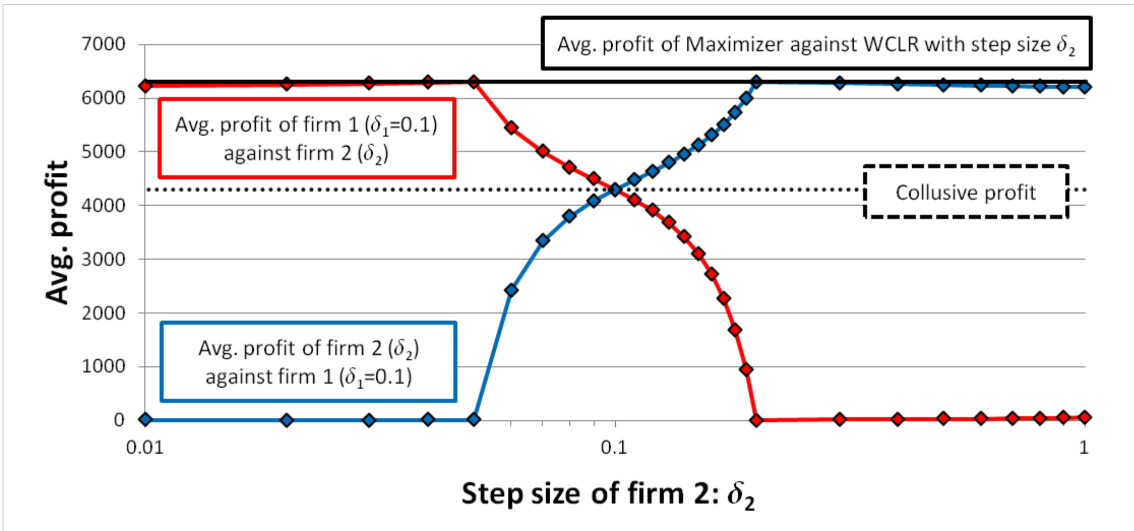

Fig. 16. The red diamonds show the average profit of WCLR firm 1 (with step size δ_1 = 0.1) and the blue diamonds show the average profit of WCLR firm 2 (with step size δ_2) when playing against each other, for different step sizes δ_2 for WCLR firm 2. The profits shown for each value of δ_2 are the averages of 100 profits obtained from 100 independent simulation runs otherwise parameterised as in the baseline. The profit obtained from each simulation run for each firm is the average profit in that simulation taken over 10^5 time steps and neglecting the first 10^5 time steps. Standard errors are below 0.2 in all cases. The profits obtained by MAXIMIZER against firm 2 (solid black line) and the collusive profit (dashed black line) are also included as a reference

4. Conclusions

The results obtained by Huck et al. (2003) indicate that the simple, individual, "sensible" and not forward-looking decision rule WCLR ("Win-Continue, Lose-Reverse") can lead to collusion-like outcomes in Cournot oligopolies, even though each company is independently trying to maximize its own profit, and is acting based only on its own past information. Similar results were obtained by Waltman and Kaymak (2008) considering a more involved learning algorithm (Q-learning). In this paper, we have shown that the convergence of the WCLR rule to collusion-like outcomes also extends to duopolies with differentiated products where firms compete in prices. In principle, these results could raise important concerns about the fairness of fining firms in oligopolies for apparently carrying out collusive practices, since one could always allege that observed collusion-like outcomes could just be the unintended result of using this type of uncoordinated independent (and thus legitimate) decision rule.

However, this paper has shown that small independent variations in the cost functions, or small uncorrelated perturbations in the price obtained by each firm (in the Cournot case) or in the particular demand of each differentiated product (in the Bertrand-like case), can all destabilize the convergence of the WCLR rule to collusive outcomes, pushing the outcomes towards the Nash solution of the one-shot game. Besides, previous simulation results (Keen and Standish, 2006) had already indicated that introducing variability in the step sizes used by each company in each period could also push the process towards the Cournot-Nash solution in markets where firms compete in quantities, and we observe the same effect in the Bertrand-like case. The convergence of the WCLR rule to collusive outcomes also rests on the synchrony or simultaneity of the decisions taken by each company; a process in which each company sequentially modifies its decision variable and obtains the market response before the other company has altered its decision would lead again[12] to the Nash equilibrium of the one-shot game, both in the Cournot and in the Bertrand-like cases. Finally, this paper has also shown that the rule WCLR is easily exploitable by unsophisticated strategies in both types of oligopoly.

Consequently, our results throw substantial doubts on the validity of arguments that try to justify collusive-like outcomes as the unintended result of firms applying the "innocent-looking" uncoordinated decision rule "Win-Continue, Lose-Reverse".

Acknowledgments

This work has received financial support from the Spanish Ministry of Science and Innovation (CSD2010-00034, SIMULPAST) and from *bolsas de viaje asistencia congresos 2014* Universidad de Valladolid.

References

Bigoni, M., Fort, M., 2013. Information and learning in oligopoly: An experiment. Games Econ.Behav. 81(1), 192-214.

Friedman, M. (1953). Part I - the methodology of positive economics. *Essays in positive economics.* (pp. 3-43) University of Chicago Press.

Friedman, D., Huck, S., Opread, R. & Weidenholzer, S. (in press). From imitation to collusion: Long-run learning in a low-information environment. *Journal of Economic Theory*, http://dx.doi.org/10.1016/j.jet.2014.10.006.

Huck, S., Normann, H. T., Oechssler, J., 2003. Zero-knowledge cooperation in dilemma games. J.Theor.Biol. 220(1), 47-54.

Huck, S., Normann, H. T., Oechssler, J., 2004. Through trial and error to collusion. Int.Econ.Rev. 45(1), 205-224.

Izquierdo, L. R., & Izquierdo, S. S. (2012). Reinforcement learning. In N. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 2796-2799) Springer US.

Izquierdo, S., & Izquierdo, L. (2015). The "win-continue, lose-reverse" rule in Cournot oligopolies: Robustness of collusive outcomes. In F. Amblard, F. J. Miguel, A. Blanchet & B. Gaudou (Eds.), *Advances in Artificial Economics* (pp. 33-44) Springer International Publishing.

Keen, S., Standish, R., 2006. Profit maximization, industry structure, and competition: A critique of neoclassical theory. Physica A: Statistical Mechanics and its Applications. 370(1), 81-85.

Kimbrough, S. O., Murphy, F. H., 2009. Learning to collude tacitly on production levels by oligopolistic agents. Comput.Econ. 33(1), 47-78.

Snyman, J. A., 2005. *Practical mathematical optimization: An introduction to basic optimization theory and classical and new gradient-based algorithms*, Springer.

Waltman, L., Kaymak, U., 2008. Q-learning agents in a cournot oligopoly model. J.Econ.Dyn.Control. 32(10), 3275-3293.

Wilensky, U. 1999. NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

[12] In a "normal" or non-pathological scenario, and considering the most direct implementation - see the discussion by Huck et al. (2003).