UNIVERSITY OF TECHNOLOGY SYDNEY

ECONOMICS DISCIPLINE GROUP

UTS BUSINESS SCHOOL

WORKING PAPER NO. 8 April 2013

Learning Cycles in Bertrand Competition with Differentiated Commodities and Competing Learning Rules

Mikhail Anufriev
Dávid Kopányiz
Jan Tuinstra

ISSN: 2200-6788
http://www.business.uts.edu.au/economics/

Learning Cycles in Bertrand Competition with Differentiated Commodities and Competing Learning Rules*

Mikhail Anufriev[†1], Dávid Kopányi[‡2] and Jan Tuinstra[§2]

[1]Economics Discipline Group, University of Technology, Sydney, Australia

[2]CeNDEF, University of Amsterdam, The Netherlands

February 27, 2013

Abstract

This paper stresses the importance of heterogeneity in learning. We introduce competition between different learning rules and demonstrate that, though these rules can coexist, their convergence properties are strongly affected by heterogeneity. We consider a Bertrand oligopoly with differentiated goods. Firms do not have full information about the demand structure and they want to maximize their perceived one-period profit by applying one of two different learning rules: least squares learning and gradient learning. We analytically show that the stability of gradient learning depends on the distribution of learning rules over firms. In particular, as the number of gradient learners increases, gradient learning may become unstable. We study the competition between the learning rules by means of computer simulations and illustrate that this change in stability for gradient learning may lead to cyclical switching between the rules. Stable gradient learning typically gives higher average profit than least squares learning, making firms switch to gradient learning. This however, can destabilize gradient learning which, because of decreasing profits, makes firms switch back to least squares learning. This cycle may repeat itself indefinitely.

JEL classification: C63, C72, D43

Keywords: Bertrand competition, heterogeneous agents, least squares learning, gradient learning, endogenous switching, learning cycle

*We are grateful to Florian Wagener for simplifying the proof of Lemma 6.1 and to two anonymous referees for their suggestions and corrections that substantially improved the paper.

[†]email: mikhail.anufriev@uts.edu.au

[‡]email: d.kopanyi@uva.nl

[§]email: j.tuinstra@uva.nl

1 Introduction

There are many situations where decision makers do not have full information about the environment in which they operate. Firms, for example, might not know relevant characteristics of their environment such as how the demand for their good depends on the price they charge, how it is affected by their competitors, who their competitors are and how they act. Learning has a natural role in these situations: agents gather the information resulting from their actions, they evaluate it and take it into account when making a decision.

There exists a wide variety of methods for modeling learning in the economics literature, including different belief-based models, least squares learning and evolutionary methods. Fudenberg and Levine (1998), Evans and Honkapohja (2001) and Cressman (2003) give good overviews of these classes. Different methods may lead to different outcomes. This is illustrated in Offerman et al. (2002), for example: they consider two imitation-based and one belief-based learning rule that lead to different market outcomes theoretically as well as in a laboratory experiment in a Cournot oligopoly with three firms framework they consider. This shows that it is essential to explicitly model the agents' learning behavior. Furthermore, the heterogeneity of agents should also be taken into account. Agents may prefer different learning methods (due to differences in computational abilities, for example) for finding out what the optimal decision is. In fact, Stahl (1996) finds experimental evidence both for heterogeneity among individuals and for switching to rules that performed better in the past. Therefore, it is important to analyze what happens in a heterogeneous environment and how different learning methods affect each other.

The aim of this paper is to analyze the interaction between two different learning methods in a heterogeneous setting where decision makers differ in the method they use. The relevance of this issue is that the convergence properties of a learning method might be affected by the presence of another method. For instance, a method that is capable of finding the Nash equilibrium in a homogeneous setting might lead to a different outcome in a heterogeneous environment. Furthermore, if different methods lead to different outcomes in a homogeneous setting, then it is unclear what will happen in a heterogeneous environment. As a framework of the analysis we consider a Bertrand oligopoly with differentiated goods. Firms do not have full information about the market: they do not fully know the demand function they face. They may use one of two different well known learning methods for deciding on the price of their good. The first method is least squares (LS) learning. With this method firms approximate their demand function with a linear function that depends only on their own price. Then they use this estimated demand function to decide on the price and they update the coefficients in the approximation in every period. Least squares learning is a natural learning method in this setup: when the relation between some variables is unknown,

then it is natural to specify a regression on the variables and to use the estimates for decision making. In our model the approximation the firms may apply is misspecified in two ways: not all the relevant variables, i.e. the prices of the other firms, are included in the regression and the functional form is incorrect. The other learning method we consider is gradient learning. With this method firms use information about the slope of their profit function for modifying their current price. Gradient learning captures the idea that firms systematically change the price of their good in the direction in which they expect to earn a higher profit. Locally stable fixed points of gradient learning correspond to local profit maxima, therefore it is natural to use it in the setting we consider. We analyze LS and gradient learning for the following reasons. Both of them are reasonable methods in the environment we consider and, as we will see, they have been applied in the literature of oligopolistic markets. In the model we assume that firms do not observe either the prices set by other firms or the corresponding demands. Therefore, it is an important criterion that the learning methods should use information only about the firms' own prices and demands. Both LS and gradient learning are appropriate in this sense. Moreover, they result in different market outcomes in a homogeneous setting, so it is not clear what kind of outcome will be observed when some firms apply LS learning while others use gradient learning and when firms are allowed to switch between the learning methods. One method may drive the other one out of the market when endogenous switching between the methods is introduced.

We address the following questions in this paper. How do the two learning methods affect each other in a heterogeneous environment? Do the dynamical properties of the model depend on the distribution of learning methods over firms? If the properties of the methods vary with respect to this distribution, can we observe cycles in which the majority of firms apply the same learning method and later they switch to the other one? Can one method drive out the other one? We study these questions with formal analysis and with computer simulations. We find that the learning methods we consider lead to different market outcomes in a homogeneous setting. With least squares learning, firms move towards a so-called *self-sustaining* equilibrium, as introduced by Brousseau and Kirman (1992), in which the perceived and the actual demands coincide at the price charged by the firms. The learning method does not have a unique steady state; the initial conditions determine which point is reached in the end. In contrast, if gradient learning converges, it leads to the unique Nash equilibrium of the market structure we consider.[1] However, gradient learning may not always converge and then we observe high-period cycles or quasi-periodic dynamics. In the steady states of a heterogeneous setting with fixed learning rules, LS learners

[1] In general, gradient learning may converge to local profit maxima that are not globally optimal. Bonanno and Zeeman (1985) and Bonanno (1988) call this kind of outcomes *local* Nash equilibria. In the market structure we consider there is a unique local profit maximum so gradient learning leads to the Nash equilibrium if it converges.

are in a *self-sustaining equilibrium* in which gradient learners give the best response to the price set by all other firms. The convergence properties of gradient learning depend on the distribution of learning methods over firms: an increase in the number of gradient learners can have a destabilizing effect. When endogenous switching between the learning rules is introduced, then a stable gradient learning does not necessarily drive out LS learning. Some LS learners may earn a higher profit than they would make as a gradient learner and then they would not switch to gradient learning. However, LS learning may drive out gradient learning when the latter is unstable. An interesting cyclical switching can occur when the convergence properties of gradient learning change as the distribution of learning methods over firms varies. When gradient learning converges, gradient learners typically earn more than the average profit of LS learners. This gives an incentive for LS learners to switch to gradient learning. An increase in the number of gradient learners, however, typically destabilizes gradient learning, resulting in low profits for gradient learners so they may start switching back to LS learning. At some point, gradient learning will converge again and the process may repeat itself.

With this paper we demonstrate that different learning methods are likely to coexist and that this coexistence can have substantial consequences for the dynamical properties of the learning methods. The dynamics with coexisting learning rules are more complex than in a homogeneous environment.

Our paper builds upon and contributes to several recent streams of literature on learning in economics. Least squares learning, for example, is widely used in the economics literature. The model we consider is closely related to the model of Kirman (1983) and Brousseau and Kirman (1992). These papers analyze the properties of misspecified LS learning in a Bertrand duopoly with differentiated goods under a linear demand specification. The learning method they use is misspecified as firms focus on their own price effect only. In this paper we generalize some results of Kirman (1983) to the case of n firms under a nonlinear demand specification. Gates et al. (1977) consider least squares learning in a Cournot oligopoly. Each firm regresses its average profits on its outputs and uses the estimated function to determine the output for the next period. The learning method the authors consider differs from ours in two respects. First, each observation has the same weight in our model whereas firms weigh observations differently in Gates et al. (1977). Second, the firms' action is specified as the action that maximizes the one-period expected profit in our paper. In Gates et al. (1977) the next action is the weighted average of the previous action and the action that maximizes the estimated profit function. Tuinstra (2004) considers the same kind of model that is studied in this paper. The firms use a misspecified perceived demand function but a different learning method is applied. When posting a price, firms are assumed to observe the demand for their good and the slope of the true demand function at that price. Then they estimate the demand function by a linear function that matches the demand and the slope the firms faced. For the estimation firms use only the most

recent observation. Firms will then use the new estimates for determining the price in the next period. Tuinstra (2004) analyzes the dynamical properties of this model and shows that complicated dynamical behavior can occur depending on the cross-price effects and the curvature of the demand functions.

Arrow and Hurwicz (1960) analyze the dynamical properties of gradient learning in a general class of n-person games. They derive conditions under which the process converges to an equilibrium and they illustrate their findings for the case of a Cournot oligopoly. Both Furth (1986) and Corchon and Mas-Colell (1996) analyze a price-setting oligopoly in which firms adjust their actions using gradient learning. The uniqueness and the stability of equilibrium points are analyzed in these papers. In this paper we also consider these issues although in a discrete time setting.

The previously discussed papers consider a homogeneous setting in which each agent uses the same learning method. However, it is reasonable to assume heterogeneity in the sense that agents apply different methods. Furthermore, they might switch between these methods. The switching mechanism we apply is related to reinforcement learning as in Roth and Erev (1995) and to the discrete choice model applied in Brock and Hommes (1997). In Roth and Erev (1995) agents have many possible pure strategies and each strategy has a propensity that determines the probability of the pure strategy being applied. These propensities depend on past payoffs. When a particular strategy was used in a given period, then its propensity is updated by adding the realized payoff to the previous propensity. The propensities of the strategies that were not used are not updated. The probability of a pure strategy being applied is proportional to the propensity of the strategy. We also use propensities for LS and gradient learning in the switching mechanism but they are updated differently than in Roth and Erev (1995): when a certain method was used, then the new propensity of that method is a weighted average of its old propensity and the current profit while the propensity of the other method remains unchanged. Furthermore, we determine the probabilities in a different way: we use the discrete-choice probabilities as in Brock and Hommes (1997). This way we can control how sensitive the firms are to differences in the performance measures. In Brock and Hommes (1997) the authors analyze a cobweb model in which agents can use either naive or perfect foresight predictors. The authors show that endogenous switching between the predictors leads to complicated dynamical phenomena as agents become more sensitive to performance differences. Droste et al. (2002) also analyze the interaction between two different behavioral rules. They consider Cournot competition with best-reply and Nash rules. With the best-reply rule, firms give the best response to the average output of the previous period. Nash firms are basically perfect foresight firms that take into account the behavior of the best-reply firms. The model of this paper differs in important aspects from the setup of Droste et al. (2002). First, firms do not know the demand they face in this paper whereas the demand functions are known in Droste et al. (2002). Second, Droste et al. (2002) basically consider

social learning: the firms observe the actions of every firm and they use this information for deciding on the production level. In contrast, firms observe only their own action and the corresponding outcome in this paper. Thus, they use individual data in the learning process instead of industry-wide variables. A consequence of this difference is that firms that apply the same learning method or behavioral rule choose the same action in Droste et al. (2002) but they typically act differently in the model we consider. Third, the switching methods are also different in the two papers. Droste et al. (2002) use replicator dynamics whereas we consider a discrete choice model, augmented with experimentation.

The paper is organized as follows. In Section 2 we present the market structure and derive the Nash equilibrium of the model. Section 3 discusses least squares learning and gradient learning. We analyze the steady states of a homogeneous LS-learning oligopoly both analytically and through simulations. Then we investigate the dynamical properties of the model with only gradient-learning firms. Section 4 combines the two learning methods in a heterogeneous setting. The learning rules are fixed in the sense that firms apply the same learning rule during the whole simulation. We analyze the model both analytically and numerically. We compare the profitability of the two learning methods as the distribution of methods over firms varies. Section 5 focuses on switching. We illustrate cyclical switching between the learning methods when the stability of gradient learning changes with the number of gradient learners. Section 6 concludes. The proofs of the propositions are presented in the Appendix.

2 The market structure

Consider a market with n firms, each producing a differentiated good and competing in prices. The demand for the product of firm i depends on the price of good i and on the average price of the other goods. The demand is given by following nonlinear function:

$$D_i(p) = \max\left\{\alpha_1 - \alpha_2 p_i^{\beta} + \alpha_3 \bar{p}_{-i}^{\gamma}, 0\right\}, \tag{1}$$

where p_i is the price of good i, p is the vector of prices and $\bar{p}_{-i} = \frac{1}{n-1}\sum_{j\neq i} p_j$. All parameters are positive and we further assume that $\beta, \gamma \in (0, 1]$ and $\beta \geq \gamma$. Parameter α_3 specifies the relationship between the products: for $\alpha_3 > 0$ the goods are *substitutes* whereas for $\alpha_3 < 0$ they are *complements*. In this paper we focus on substitutes.[2] The demand is decreasing and convex in the own price and it is increasing and concave in the price of the other goods. The market structure is fully symmetric: firms face symmetric demands and the marginal cost of production is constant, identical across firms and equal to c.

[2]We discuss results for the case of complements too but we do not report the corresponding simulations. The case of complements is discussed in more detail in Kopányi (2013) under a linear demand specification.

We impose some restrictions on the parameter values which ensure that a symmetric Nash equilibrium exists.

Assumption 2.1 *The parameters satisfy* $\alpha_1 - \alpha_2 c^\beta + \alpha_3 c^\gamma > 0$ *and* $-\alpha_2\beta c^\beta + \alpha_3\gamma c^\gamma < 0$.

The first restriction says that the demand is sufficiently large: demands are positive when each firm sets the price equal to the marginal cost. The second restriction ensures that the demand for the goods strictly decreases if every firm increases the price of its good marginally in a symmetric situation where $p_i = p$ for all firms (as long as $D(p) > 0$)[3]. Proposition 2.2 specifies the unique Nash equilibrium of the model. The proof is presented in the Appendix.

Proposition 2.2 *Under Assumption 2.1 the model has a unique Nash equilibrium. In this equilibrium, each firm charges price* p_N *that is characterized by the equation*

$$\alpha_1 - \alpha_2 p_N^\beta + \alpha_3 p_N^\gamma - \alpha_2\beta(p_N - c)p_N^{\beta-1} = 0. \tag{2}$$

The Nash equilibrium price exceeds the marginal cost.

Note that the Nash equilibrium is symmetric and the price is independent of the number of firms. This is due to the fact that the *average* price of other goods determines the demand so the number of firms does not affect the equilibrium.

Firms do not have full information about the market environment. In particular, they do not know the demand specifications, furthermore they cannot observe either the prices or the demands for the other goods. They are assumed to know their own marginal cost. Firms repeatedly interact in the environment described above. They are myopic profit maximizers: they are only interested in maximizing their one-period profit. Firms can apply one of two methods to decide on the price they ask in a given period. These methods are discussed in Section 3.

3 Learning methods

One method that firms may apply is least squares learning. With this method firms use an estimated demand function for maximizing their expected profit. The other method is gradient learning: firms use information about their marginal profit at the current price and they adjust the current price of their good in the direction in which they expect to get a higher profit. Both methods focus on the own price effect without considering the effect of the price change of other goods. Section 3.1 presents LS learning while gradient learning is analyzed in Section 3.2.

[3] I.e. the own price effect dominates in a symmetric situation.

3.1 Least squares learning

With least squares learning firms use past price-quantity observations for estimating a perceived demand function and then they maximize their expected profit, given this perceived demand function. The parameter estimates determine the price they set in the next period. As new observations become available, firms update the parameter estimates and thus the price of their good.

3.1.1 The learning mechanism

Firm i assumes that the demand for its good depends linearly on the price of the good but that it is independent of the price of other goods. The perceived demand function of firm i is of the form

$$D_i^P(p_i) = a_i - b_i p_i + \varepsilon_i, \tag{3}$$

where a_i and b_i are unknown parameters and ε_i is a random variable with mean 0. Notice that firm i uses a misspecified model since the actual demand (1) is determined by all prices, furthermore it depends on prices in a *nonlinear* manner. Kirman (1983) applies the same kind of misspecified LS learning in a Bertand duopoly with differentiated goods. He argues that it is reasonable for firms to disregard the prices of other goods in an oligopolistic setting. When the number of firms is large, it requires too much effort to collect every price, so firms rather focus on their own-price effect and treat the pricing behavior of the other firms as an unobservable error.

For obtaining the coefficients of the perceived demand function the firm regresses the demands it faced on the prices it asked. All past observations are used with equal weight in the regression. Let $a_{i,t}$ and $b_{i,t}$ denote the parameter estimates observed by firm i at the end of period t. These estimates are given by the standard OLS formulas (see Stock and Watson (2003), for example):

$$b_{i,t} = \frac{\left(\frac{1}{t}\sum_{\tau=1}^{t} p_{i,\tau}\right)\left(\frac{1}{t}\sum_{\tau=1}^{t} q_{i,\tau}\right) - \frac{1}{t}\sum_{\tau=1}^{t} p_{i,\tau}q_{i,\tau}}{\frac{1}{t}\sum_{\tau=1}^{t}(p_{i,\tau})^2 - \left(\frac{1}{t}\sum_{\tau=1}^{t} p_{i,\tau}\right)^2}, \tag{4}$$

$$a_{i,t} = \frac{1}{t}\sum_{\tau=1}^{t} q_{i,\tau} + b_{i,t}\frac{1}{t}\sum_{\tau=1}^{t} p_{i,\tau}, \tag{5}$$

where $q_{i,\tau}$ denotes the actual demand for good i in period τ: $q_{i,\tau} = D_i(p_\tau)$.

Note that even though $a_{i,t}$ and $b_{i,t}$ should be positive in order to have an economically sensible perceived demand function, the parameter estimates may be negative in some periods. Furthermore, if $a_{i,t} \leq b_{i,t}c$, then firm i cannot expect to earn a positive profit since the perceived demand becomes zero already at a price that is smaller than the marginal cost c. In these situations we assume that the firm does not use the parameter estimates to choose a price. We will shortly specify how the firm acts in this situation.

When the aforementioned cases do not occur (that is when $a_{i,t} > b_{i,t}c > 0$), then firm i determines the price for the next period by maximizing its expected profit:

$$\max_{p_{i,t+1} \geq c} E_t \left((p_{i,t+1} - c)(a_{i,t} - b_{i,t}p_{i,t+1} + \varepsilon_{i,t+1}) \right) = \max_{p_{i,t+1} \geq c} \{(p_{i,t+1} - c)(a_{i,t} - b_{i,t}p_{i,t+1})\}.$$

The objective function is quadratic in $p_{i,t+1}$ and the quadratic term has a negative coefficient. Then in period $t+1$ the firm asks the perceived profit-maximizing price $p_{i,t+1} = \frac{a_{i,t}}{2b_{i,t}} + \frac{c}{2}$. If the condition $a_{i,t} > b_{i,t}c > 0$ does not hold, then firm i draws the price from the uniform distribution on the set $S = \left\{ p \in \mathcal{R}_+^n : \; p_i > c, \; D_i(p) > 0, \; i = 1, \ldots, n \right\}$.[4] This set is the set of price vectors for which every firm makes a positive profit. Thus, when the perceived demand function is not sensible economically (i.e. $a_{i,t} \leq 0$ or $b_{i,t} \leq 0$), then the firm asks a random price rather than applying an incorrect pricing formula. Also, the firm asks a random but not unprofitable price rather than a price that yields a certain loss.

LS learning is implemented in the following way. For any firm i:

1. $p_{i,1}$ and $p_{i,2}$ are randomly drawn from the uniform distribution on set S.

2. At the end of period 2 the firm uses OLS formulas (4) and (5) to obtain the parameter estimates $a_{i,2}$ and $b_{i,2}$.

3. **a.** In period $t \geq 3$ the firm asks the price $p_{i,t} = \frac{a_{i,t-1}}{2b_{i,t-1}} + \frac{c}{2}$ if $a_{i,t-1} > b_{i,t-1}c > 0$. In every other case $p_{i,t}$ is drawn from the uniform distribution on the set S.

   b. After realizing the demand, the firm updates the coefficients of the perceived demand function using (4) and (5).

4. The process stops when the absolute price change is smaller than a threshold value δ for all firms: $\max_i \{|p_{i,t} - p_{i,t-1}|\} < \delta$.

Notice that the learning process of other firms interferes with the firm's own learning process. As the prices of other goods change, the demand the firm faces also changes. Although the change in the demand for good i is caused not only by the change in its price, firm i attributes the change in its demand solely to changes in its own price and to random noise. Therefore, the firm tries to learn a demand function that changes in every period. Learning is more complicated in the initial periods since prices are more volatile than in later periods when the learning process slows down.

[4] More properly: when firm i chooses a random price, then a price vector p is drawn from the uniform distribution on S and firm i will charge the i-th component of p. Every other firm j chooses a price according to the pricing formula or, if the condition $a_{j,t} > b_{j,t}c > 0$ does not hold, firm j draws another vector.

3.1.2 Equilibria with LS learning

Brousseau and Kirman (1992) show that the misspecified LS learning we consider does not converge in general.[5] Price changes however become smaller over time as the weight of new observations decreases. Thus, the stopping criterion we specified will be satisfied at some point and the learning mechanism stops. We will see that the resulting point is very close to a so-called *self-sustaining equilibrium* in which the actual and the perceived demands of a firm coincide. The set of possible equilibria is infinite, however.

With the method described above firms use a misspecified model since the perceived demand functions (3) differ from the actual demand functions (1). Nevertheless, firms may find that the price they charge results in the same actual demand as the perceived demand function predicts. If this holds for all firms, then the model is in equilibrium since the parameter estimates of the perceived demand functions do not change and firms will ask the same price in the following period. To see that this is the case, note the following. The LS coefficients at the end of period t minimize the sum of squared errors up to period t. If the perceived and the actual demands are equal at p_{t+1}, then the parameter estimates a_{t+1} and b_{t+1} remain the same: under a_t and b_t the error corresponding to the new observation is 0 and the sum of squared errors up to period t is minimized. Thus, the sum of squared errors up to period $t+1$ is minimized by exactly the same coefficients. Brousseau and Kirman (1992) call this kind of equilibrium *self-sustaining equilibrium*: firms charge the optimal price (subject to their perceived demand function) and the corresponding demand coincides with the demand they expected to get, therefore the firms have no reason to believe that their perceived demand function is misspecified. Following their terminology, we refer to such equilibria as self-sustaining equilibria (SSE).

The left panel of Figure 1 illustrates a disequilibrium of the model. The solid line is the perceived inverse demand function

$$p_i = P_i^P(q_i) \equiv \frac{a_i}{b_i} - \frac{1}{b_i} q_i, \tag{6}$$

the dashed line depicts the actual inverse demand function

$$p_i = P_i(q_i, \bar{p}_{-i}) \equiv \left[\frac{1}{\alpha_2} \left(\alpha_1 + \alpha_3 \bar{p}_{-i}^{\gamma} - q_i \right) \right]^{\frac{1}{\beta}}. \tag{7}$$

The downward-sloping dotted line is the perceived marginal revenue. The quantity that maximizes the expected profit of firm i is given by the x-coordinate of the intersection of the perceived marginal revenue (MR) and the marginal cost (MC). Let q^P denote this quantity. If the firm wants to face a demand equal to q^P, then it has to ask price p which is determined by the value of the perceived inverse demand function

[5] LS learning may converge in many other situations. Marcet and Sargent (1989) derive conditions under which LS learning converges for a wide class of models.

at q^P. However, the firm might face a different demand as the actual and perceived demand functions differ. Let q^A denote the actual demand the firm faces when its price is p. The left panel of Figure 1 shows a situation in which the expected and the actual demands are not the same. This is not an SSE of the model. In this case the firm will add the new observation (p, q^A) to the sample and run a new regression in the next period. This new observation changes the perceived demand function and the firm will charge a different price. In contrast, the right panel of Figure 1 illustrates the situation when $q^P = q^A$, that is the actual and the expected demands coincide at price p. This constitutes an SSE (provided that the corresponding variables of the other firms also satisfy these conditions). The new observation does not change the coefficients of the perceived demand function so the firm will charge the same price in subsequent periods.

We now describe the equilibrium conditions and the set of SSE prices. Variables p_i^*, a_i^*, and b_i^*, $(i = 1, \ldots, n)$ constitute an SSE if the following conditions are satisfied for all firms:

$$p_i^* = \frac{a_i^*}{2b_i^*} + \frac{c}{2}, \tag{8}$$

$$D_i(p^*) = D_i^P(p_i^*). \tag{9}$$

Condition (8) says that firms set the price that maximizes their expected profit subject to their perceived demand function. Condition (9) requires that the actual and the perceived demands are the same at the SSE prices. Since we have 2 independent equations and 3 variables for each firm, we can express a_i^* and b_i^* as a function of the SSE prices. Thus, for given prices we can find perceived demand functions such that the firms are in an SSE. Proposition 3.1 specifies the coefficients of the perceived demand function in

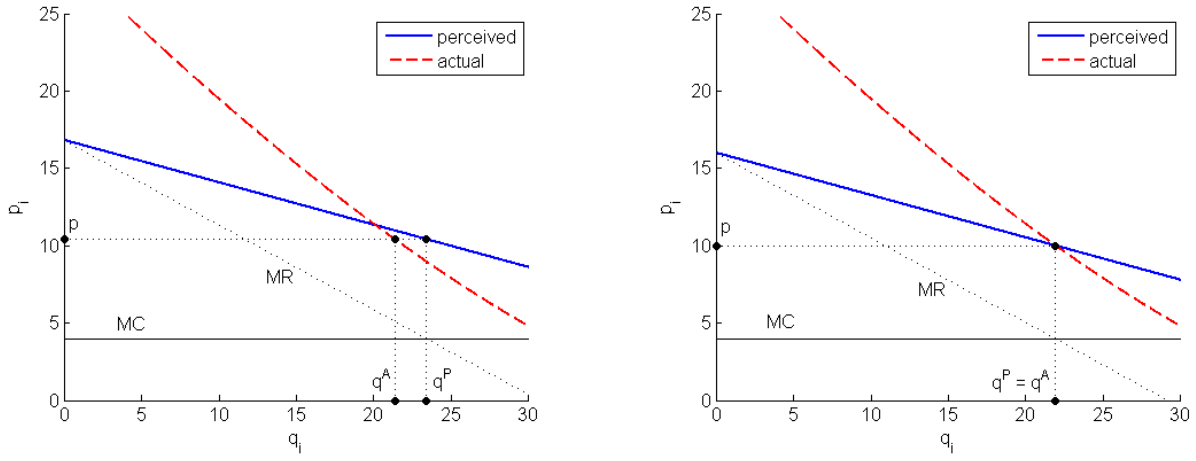


Figure 1: Disequilibrium (left) and self-sustaining equilibrium (right) of the model with LS learning. Parameter values: $\alpha_1 = 35$, $\alpha_2 = 4$, $\alpha_3 = 2$, $\beta = 0.7$, $\gamma = 0.6$, $c = 4$ and $\bar{p}_{-i} = 8$.

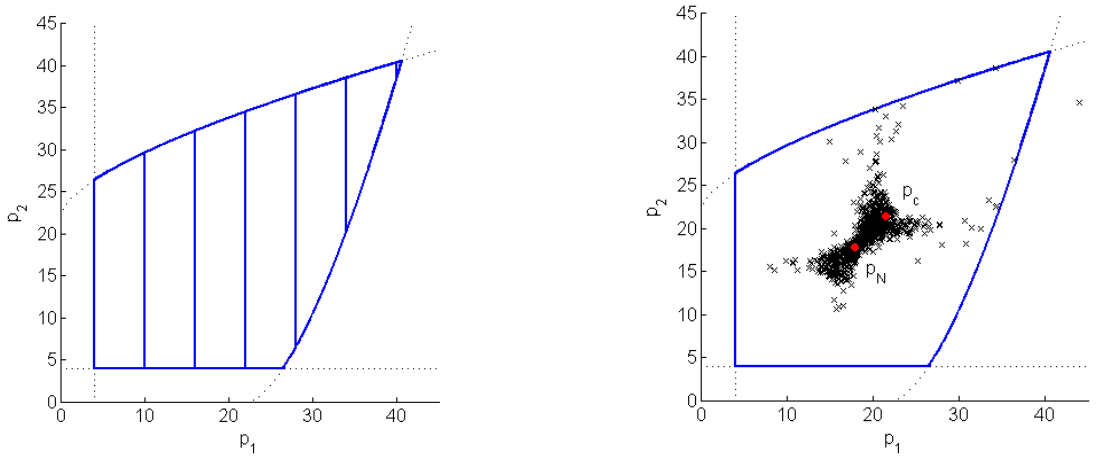


Figure 2: The set of SSE prices for two firms (left panel) and the end prices of simulations with initial prices drawn from the uniform distribution on the set of SSE prices (right panel). Parameter values: $\alpha_1 = 35$, $\alpha_2 = 4$, $\alpha_3 = 2$, $\beta = 0.7$, $\gamma = 0.6$, $c = 4$ and $\delta = 10^{-8}$.

terms of the SSE prices. It also describes the set of SSE prices. The proposition is proved in the Appendix.

Proposition 3.1 *For given prices p_i^* $(i = 1, \ldots, n)$ the model is in an SSE if the coefficients of the perceived demand function of firm i are given by*

$$a_i^* = D_i(p^*)\left(1 + \frac{p_i^*}{p_i^* - c}\right),$$
$$b_i^* = \frac{D_i(p^*)}{p_i^* - c}.$$

The set of SSE prices is described by the conditions $p_i^* > c$ and $D_i(p^*) > 0$, or equivalently

$$c < p_i^* < P_i(0, \bar{p}_{-i}^*) = \left[\frac{1}{\alpha_2}\left(\alpha_1 + \alpha_3(\bar{p}_{-i}^*)^\gamma\right)\right]^{\frac{1}{\beta}}.$$

This set is nonempty and bounded.

The values of a_i^* and b_i^* derived in Proposition 3.1 are in line with Proposition 3 of Kirman (1983): they reduce to the same expression for the case of a duopoly with a linear demand function and zero marginal cost. Note that the set S we use in the LS algorithm coincides with the set of SSE prices. The set of SSE prices always contains the Nash equilibrium as the Nash equilibrium price exceeds the marginal cost and the corresponding demand is positive. The left panel of Figure 2 depicts the set of SSE prices for the case of two firms. The figure corresponds to parameter values $\alpha_1 = 35$, $\alpha_2 = 4$, $\alpha_3 = 2$, $\beta = 0.7$, $\gamma = 0.6$ and $c = 4$. We will use these parameter values in all later simulations too. For these values the Nash equilibrium price is $p_N \approx 17.7693$ with corresponding profit $\pi_N \approx 223.9148$.

In Proposition 3.1 we characterized the set of prices that may constitute an SSE. However, nothing ensures that every point of that set will actually be reached from some initial points. In fact, Kirman (1983) derives the set of points that can be reached with some initial values for the case of two firms and linear demand specification. He shows that this set is smaller than the set of SSE prices.[6]

Note that different firms could use different types of perceived demand functions, in general. For example, some firms could take into account the prices of (some of the) other firms. Furthermore, perceived demand functions could be nonlinear in prices. We discuss the effects of heterogeneity in the perceived demand functions in the concluding remarks.

3.1.3 Simulation results

To illustrate some properties of least squares learning we simulate the model where each firm is an LS learner. We use the aforementioned parameter values with threshold value $\delta = 10^{-8}$ in the stopping criterion. First we illustrate that firms reach a point in the set of SSE prices when there are two firms. We drew 2000 random points from the uniform distribution on the set of SSE prices and ran 1000 simulations using these points as initial prices.[7] In order to save time we limited the number of runs to 10000.[8] The right panel of Figure 2 depicts the end prices of the 1000 runs. We observe that almost all of the final points lie in the set of SSE prices and that they do not fill the whole set. Nevertheless, there is quite a variety in final prices so homogeneous LS learning can lead to many possible outcomes. For the case of 10 firms we observed a variety in the final prices again and that most of the final points lie in the set of SSE prices. This latter result is not robust to the demand parameters: when the set of SSE prices is more expanded towards high prices, then final points fall outside the set of SSE prices more often.[9] The other finding that LS learning may result in many possible outcomes is robust with respect to the demand parameters. These results remain valid even when we add a small noise to the actual demands.

Figure 3 illustrates typical time series of prices and profits for the case of 10 firms. Although the stopping criterion is satisfied at period 9201, we plot only the first 20 periods as the time series do not change much after that. We observe that prices are more volatile in the first few periods but then they

[6] Kirman (1983) does not consider non-negativity constraints on a_i and b_i so any positive price pair can constitute an SSE in his model.

[7] We need two initial points for each simulation. The first 2 points are used as initial values in the first simulation, the third and the fourth are used in the second one etc.

[8] So the simulation stopped at period 10000 even if the stopping criterion was not met. Based on other simulations, this does not affect the outcome substantially.

[9] Simulations show that if we do not consider the non-negativity constraint on demands, then almost all points lie within the set of SSE prices irrespective of the shape of the set.

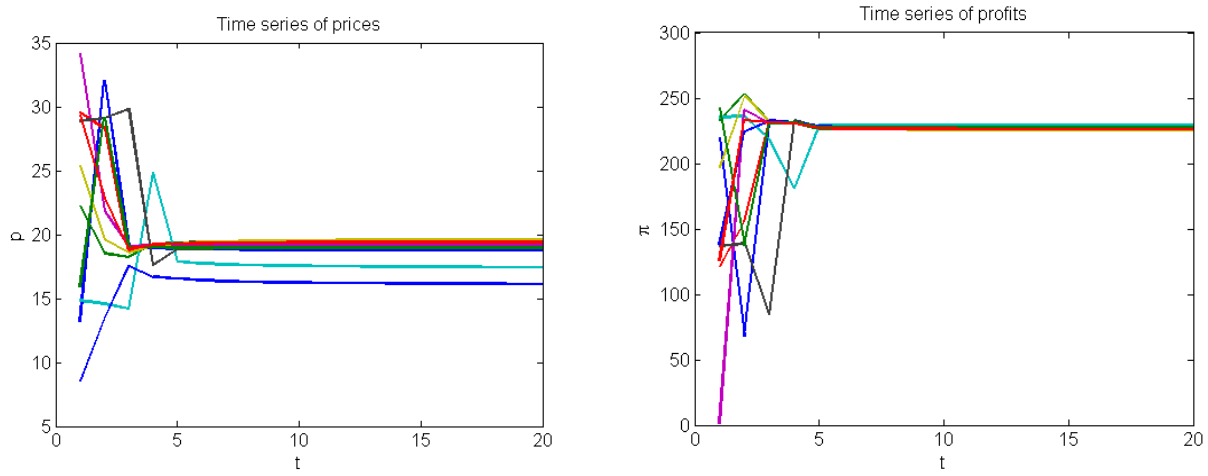


Figure 3: Time series of prices (left) and profits (right) in an oligopoly with 10 LS learners. Parameter values: $\alpha_1 = 35$, $\alpha_2 = 4$, $\alpha_3 = 2$, $\beta = 0.7$, $\gamma = 0.6$, $c = 4$ and $\delta = 10^{-8}$. Nash equilibrium values: $p_N \approx 17.7693$ and $\pi_N \approx 223.9148$.

start to settle down. In this particular simulation, end prices are between 16.3 and 19.6, while profits lie between 225.4 and 229.3. Note that profits exceed the corresponding Nash equilibrium profit 223.9. We analyzed the distribution of end prices by simulating the model with initial prices drawn from the uniform distribution on the set of SSE prices. As the number of firms increases, a higher proportion of end prices lies close to the mode that exceeds the Nash equilibrium price.

3.2 Gradient learning

Let us now turn to the other method that firms may apply for deciding on prices. Instead of assuming a specific form for the demand function and estimating its parameters, firms use information about the slope of their profit function.[10] Knowing the slope at the current price, firms adjust the price of their good in the direction in which they expect to get a higher profit.

3.2.1 The learning mechanism

The price charged by firm i in period $t+1$ is given by

$$p_{i,t+1} = \max\left\{p_{i,t} + \lambda \frac{\partial \pi_i(p_t)}{\partial p_{i,t}}, c\right\}, \tag{10}$$

[10] For analytically calculating the slope firms would need to know the actual demand function and the prices asked by other firms. Nevertheless, with market experiments they can get a good estimate of the slope without having the previously mentioned pieces of information. Thus, it is not unreasonable to assume that firms know the slope of their profit function.

where the derivative of the profit function is $\alpha_1 - \alpha_2 p_{i,t}^{\beta} + \alpha_3 \bar{p}_{-i,t}^{\gamma} - \alpha_2 \beta (p_{i,t} - c) p_{i,t}^{\beta-1}$. Formula (10) shows that the price adjustment depends on the slope of the profit function and on parameter λ. In Section 3.2.2 we will see that the stability properties of this learning rule depend heavily on the value of λ.

We augment this method with an additional rule. Note that if a firm sets a too high price for which the demand is zero, then (10) gives the same price for the next period since the slope of the profit function is zero at that point. However, it should be clear for the firms that the zero profit may result from charging a too high price, so it is reasonable to lower the price. Therefore, we add the following rule to gradient learning: if a firm faced zero demand in two consecutive periods, then it lowers its previous price by λ_0. This rule ensures that firms cannot get stuck in the zero profit region. We assume that λ_0 takes the same value as λ in all simulations.[11]

Gradient learning is implemented in the following way. For every firm i:

1. $p_{i,1}$ and $p_{i,2}$ are drawn from the uniform distribution on the set S.[12]

2. In period $t \geq 3$:

   - If $D_i(p_{t-2}) \neq 0$ or $D_i(p_{t-1}) \neq 0$, then $p_{i,t} = \max\left\{p_{i,t-1} + \lambda \frac{\partial \pi_i(p_{t-1})}{\partial p_{i,t-1}}, c\right\}$.
   - If $D_i(p_{t-2}) = D_i(p_{t-1}) = 0$, then the price is given by $p_{i,t} = \max\{p_{i,t-1} - \lambda_0, c\}$.

3. The process continues until all price changes are smaller in absolute value than a threshold value δ.

Similarly to the case of LS-learning firms, the firms' learning processes interfere with each other. Although a firm moves in the direction that is expected to yield a higher profit, it may actually face a lower profit after the price change since the profit function of the firm changes due to the price change of other firms. Nevertheless, if gradient learning converges, then this disturbance becomes less severe as there will be only small price changes in later periods.

3.2.2 Equilibrium and local stability

Let us now investigate the dynamical properties of gradient learning. In the first part of the analysis we will not consider non-negativity constraints on prices and demands and we disregard the augmenting rule. We will discuss the effects of these modifications after deriving the general features of the learning rule.

[11] The exact value of λ_0 affects only the speed of return from the zero profit region, it does not affect the convergence properties of the method.

[12] Although it would be sufficient to take one initial value for the simulations, we take two initial values so that gradient learning would be more comparable with LS learning in a heterogeneous setting. We take initial values from the same set for the same reason.

The law of motion of prices is given by

$$p_{i,t+1} = p_{i,t} + \lambda \frac{\partial \pi_i(p_t)}{\partial p_{i,t}}.$$

The system is in a steady state if the derivative of the profit function with respect to the own price is zero for all firms. Under the demand specification we consider, this condition characterizes the Nash equilibrium, so the Nash equilibrium is the unique steady state of the model with only gradient learners.

Let us now analyze the stability of the steady state. Proposition 3.2 summarizes the dynamical properties of the gradient-learning oligopoly. The proof of the proposition can be found in the Appendix.

Proposition 3.2 *The Nash equilibrium price p_N is locally stable in the gradient-learning oligopoly if*

$$\lambda \left\{ \alpha_2 \beta p_N^{\beta-1} \left[2 + (\beta - 1) \frac{p_N - c}{p_N} \right] + \alpha_3 \gamma p_N^{\gamma-1} \frac{1}{n-1} \right\} < 2.$$

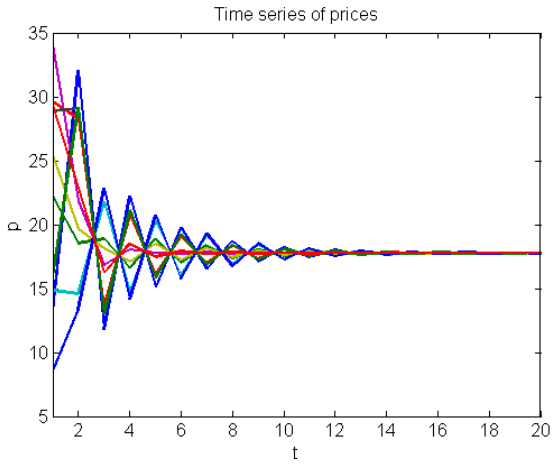
When the expression on the left hand side equals 2, $n-1$ eigenvalues become -1, the remaining eigenvalue lies within the unit circle.

According to Proposition 3.2, the steady state is locally stable if parameter λ is sufficiently small. The steady state loses stability through a *degenerate flip bifurcation*: multiple eigenvalues exit the unit circle through the value -1. In general, many different cycles with period multiple of 2 could be created with this kind of bifurcation.[13] We numerically investigate the occurring dynamics in the next section. Note that the coefficient of λ in the stability condition is decreasing in n as p_N is independent of n. Thus, an increase in the number of firms has a stabilizing effect.

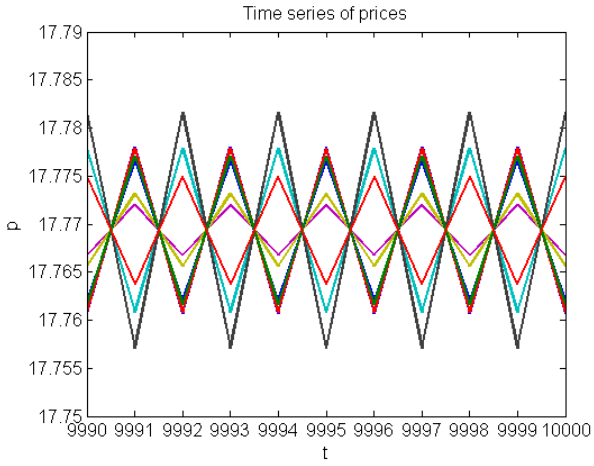
So far we have not considered the effect of the constraints $p_i \geq c$, $D_i(p) \geq 0$ and the augmenting rule that lowers too high prices. For discussing these effects let us first consider a linear demand function. In that case the system is linear so there are three kinds of possible dynamics if we do not consider any constraints: convergence to a steady state, to a 2-cycle or unbounded divergence. Unbounded divergence is no longer possible when we impose the constraints on prices and demands. These constraints and the augmenting rule drive prices back towards the region where the demands are positive. Therefore, we may observe high-period cycles, quasi-periodic or aperiodic dynamics for high values of λ.

In the nonlinear setting we consider, the non-negativity constraint on prices must be imposed since a negative price yields a complex number as demand. The effect of the constraints and the augmenting rule is the same as for a linear demand function: unbounded divergence cannot occur, we observe high-period cycles, quasi-periodic dynamics or aperiodic time series instead.

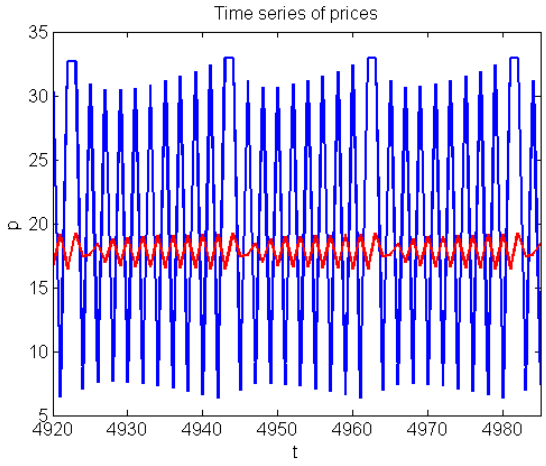
[13] More details about degenerate flip bifurcations can be found in Mira (1987), Bischi et al. (2000) and Bischi et al. (2009).



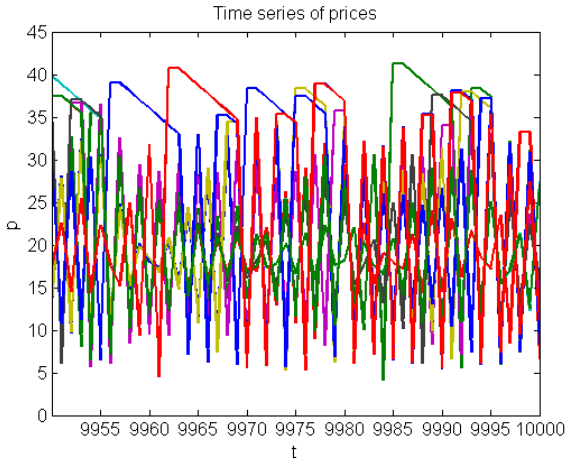
(a) Convergence to the Nash equilibrium for $\lambda = 0.8$



(b) Convergence to a 2-cycle for $\lambda \approx \lambda^*$



(c) Quasi-periodic dynamics for $\lambda = 0.9344$



(d) Aperiodic dynamics for $\lambda = 1$

Figure 4: Typical time series of prices in an oligopoly with 10 gradient learners for different values of λ. Other parameter values: $\alpha_1 = 35$, $\alpha_2 = 4$, $\alpha_3 = 2$, $\beta = 0.7$, $\gamma = 0.6$, $c = 4$ and $\delta = 10^{-8}$. Nash equilibrium price: $p_N \approx 17.7693$.

Similarly to LS learning, firms could use different types of gradient learning as well. This can be implemented with different adjustment parameter λ among firms. Heterogeneity in parameters of the learning algorithms is well established by the laboratory experiments in different settings, see e.g. Erev et al. (2010) and Anufriev et al. (2013). In the concluding remarks we discuss the effect of individual heterogeneity with respect to the adjustment parameter λ.

3.2.3 Simulation results

We run simulations for illustrating the possible dynamics of the model with only gradient learners. We use the same parameter values as before. Figure 4 illustrates typical time series of prices: convergence to

the Nash equilibrium for $\lambda = 0.8$ in panel (a), time series that resemble a 2-cycle for $\lambda = \lambda^* \approx 0.9391$ in panel (b), quasi-periodic dynamics for $\lambda = 0.9344$ in panel (c) and aperiodic dynamics for $\lambda = 1$ in panel (d). These patterns can occur for different demand parameters too but for different values of λ.

In line with Proposition 3.2, we observe convergence to the Nash equilibrium price when λ is sufficiently small. Starting with initial prices in a small neighborhood of the Nash equilibrium, we can observe time series converging to a 2-cycle around the point of bifurcation. Furthermore, by varying initial conditions we observe a co-existence of several attracting 2-cycles. For higher values of λ we observe high-period cycles, quasi-periodic or aperiodic dynamics.

It turns out from simulations that the lower λ is, the larger the range of initial prices for which convergence can be observed in simulations. So when the steady state is locally stable and λ is close to the bifurcation value, then we observe convergence only for a small set of initial prices. When initial prices lie outside of this set, then we observe high-period cycles, quasi-periodic or aperiodic dynamics.

4 Heterogeneous oligopoly with fixed learning rules

In this section we combine the learning methods discussed in Sections 3.1 and 3.2 and we consider the case of a heterogeneous oligopoly in which some firms use least squares learning while others apply gradient learning. Firms use a fixed learning method and they cannot change the rule they use. We will see that the main features of the two methods remain valid even in the heterogeneous setting: when λ is sufficiently small, then LS learners get close to an SSE in which gradient learners give the best response to the prices set by the other firms.

4.1 Steady states and stability

Consider a market with n_L LS learners and $n - n_L$ gradient learners where $0 < n_L < n$. Let us assume without loss of generality that the first n_L firms are the LS learners. We discussed in Section 3.1.2 that the steady states of an LS-learning oligopoly are characterized by a *self-sustaining equilibrium*. The same conditions must hold for LS learners in a steady state of a heterogeneous oligopoly: their actual and perceived demands must coincide at the price they ask (given the prices of other firms), otherwise they would update their perceived demand function and the price of their good would change in the next period. At the same time, the slope of the profit function of gradient learners must be zero in a steady state otherwise the price of their good would change. Proposition 4.1 characterizes the steady state of the heterogeneous oligopoly with fixed learning rules. We leave the proof to the reader, it can be proved with very similar steps as in the proof of Proposition 2.2.

Proposition 4.1 *In a steady state of the system, LS learners are in an SSE and gradient learners give the best response to the prices set by other firms. The price p_G set by gradient learners is characterized by*

$$\alpha_1 - \alpha_2 p_G^{\beta} + \alpha_3 \left[\frac{1}{n-1} \left(\sum_{s=1}^{n_L} p_s^* + (n - n_L - 1) p_G \right) \right]^{\gamma} - \alpha_2 \beta (p_G - c) p_G^{\beta - 1} = 0,$$

where p_s^* denotes the price of LS learner s.

Later we will illustrate with numerical analysis that there is a unique solution p_G to the above equation. Since, at a steady state, gradient learners give the best-response price, steady states are similar to a Stackelberg oligopoly outcome with LS learners as leaders and gradient learners as followers. It is, however, not a real Stackelberg outcome because LS learners do not behave as real leaders since they do not take into account the reaction function of gradient learners when setting the price of their good. Nevertheless, LS learners may earn a higher profit than gradient learners in a steady state.[14]

Let us now turn to the stability of the steady states. As LS learners always settle down at a certain price since the weight of a new observation decreases as the number of observations increases, stability depends mainly on the dynamical properties of gradient learning. Proposition 4.2 presents these properties. The proof can be found in the Appendix.

Proposition 4.2 *Under the assumption that LS learners have reached their steady state price, the dynamical properties of gradient learning are as follows. For $0 < n_L < n - 1$ the price p_G set by gradient learners is locally stable if $0 < \lambda M_1 < 2$ and $0 < \lambda M_2 < 2$, where*

$$M_1 = \alpha_2 \beta p_G^{\beta - 1} \left[2 + (\beta - 1) \frac{p_G - c}{p_G} \right] - \alpha_3 \gamma \left(\frac{n_L \bar{p}^* + (n - n_L - 1) p_G}{n - 1} \right)^{\gamma - 1} \frac{n - n_L - 1}{n - 1}, \tag{11}$$

$$M_2 = \alpha_2 \beta p_G^{\beta - 1} \left[2 + (\beta - 1) \frac{p_G - c}{p_G} \right] + \alpha_3 \gamma \left(\frac{n_L \bar{p}^* + (n - n_L - 1) p_G}{n - 1} \right)^{\gamma - 1} \frac{1}{n - 1}, \tag{12}$$

and $\bar{p}^* = \frac{1}{n_L} \sum_{s=1}^{n_L} p_s^*$ is the average LS price.

For $n_L = n - 1$ the price set by the gradient learner is locally stable if

$$\lambda \alpha_2 \beta p_G^{\beta - 1} \left[2 + (\beta - 1) \frac{p_G - c}{p_G} \right] < 2.$$

[14] Gal-Or (1985) shows that there is a second mover advantage in a price-setting duopoly with substitute goods. Under a linear demand specification, this result can be extended to a higher number of firms too. We expect this to hold also for the nonlinear demand specification we consider when the demand functions are not too far from the linear case. However, since LS learners do not charge the optimal leaders' price, this deviation from the optimal price may hurt the gradient learners even more and they may earn a lower profit than the LS learners.

Note that the previous proposition concerns the stability of the price set by gradient learners and not those of the *steady states*. Although LS learners get close to an SSE and the price set by gradient learners is locally stable for low values of λ, we cannot say that the steady state is locally stable. A small perturbation of a steady state leads to different LS prices and this changes the best-response price too. If, however, the LS prices remained the same, then gradient learners would return to the best-response price after a small perturbation. Note further that the proposition establishes sufficient conditions for *local* stability only. Thus, we might not observe convergence to a steady state for large perturbations of gradient learners' price.

The distribution of learning methods over firms affects the stability of the price set by gradient learners as n_L appears in the aforementioned stability conditions. It is, however, not clear analytically how stability changes with respect to n_L because a change in n_L affects the average LS price $\bar{p}^*$, which can take many different values. For further analyzing this issue, we use numerical calculations. First we check the direct effect of n_L on stability and then we analyze how an increase in n_L affects the average LS price.

Although LS prices are unknown, we can make use of the fact that the set of SSE prices is bounded: the minimal SSE price is c and the maximal SSE price $\hat{p}$ is defined by $\alpha_1 + \alpha_3 \hat{p}^\gamma - \alpha_2 \hat{p}^\beta = 0$ (as shown in the proof of Proposition 3.1). Thus, we have $c \leq \bar{p}^* \leq \hat{p}$. Taking values for $\bar{p}^*$ from this range, we can calculate p_G, M_1 and M_2 numerically. The left panel of Figure 5 shows that p_G is unique (given the average LS price and the number of LS learners). Note that there is a value of $\bar{p}^*$ for which the best-response price is the same irrespective of the number of LS learners. This price equals the Nash equilibrium price since the best response to the Nash equilibrium price is the Nash equilibrium price itself.

It turns out from the calculations that only M_2 is relevant for stability: M_1 is always positive and $M_1 < M_2$ as $\alpha_3 > 0$. Using (12) we can calculate for any $\bar{p}^*$ and any n_L the threshold value of λ for which the gradient learners' price loses stability. Using these threshold values, we depict the stability-instability region for different values of n_L in the right panel of Figure 5. We can see from the graph that for a given average LS price, the region of stability is increasing (decreasing) in n_L if the average LS price is larger (smaller) than the Nash equilibrium price and if there are more than one gradient learners (i.e. $n_L < n - 1$).[15] Thus, for a fixed average LS price, an increase in the number of LS learners has a (de)stabilizing effect if the average LS price is larger (smaller) than the Nash equilibrium price and if there are at least 2 gradient learners. For $n_L = n-1$ the stability condition becomes different. When the average LS price exceeds the Nash equilibrium price, then the change in the stability region is still monotonic, but it is no longer monotonic when the average LS price is lower than the Nash equilibrium price.

[15] Remember that there is a different stability condition for the case $n_L = n - 1$.

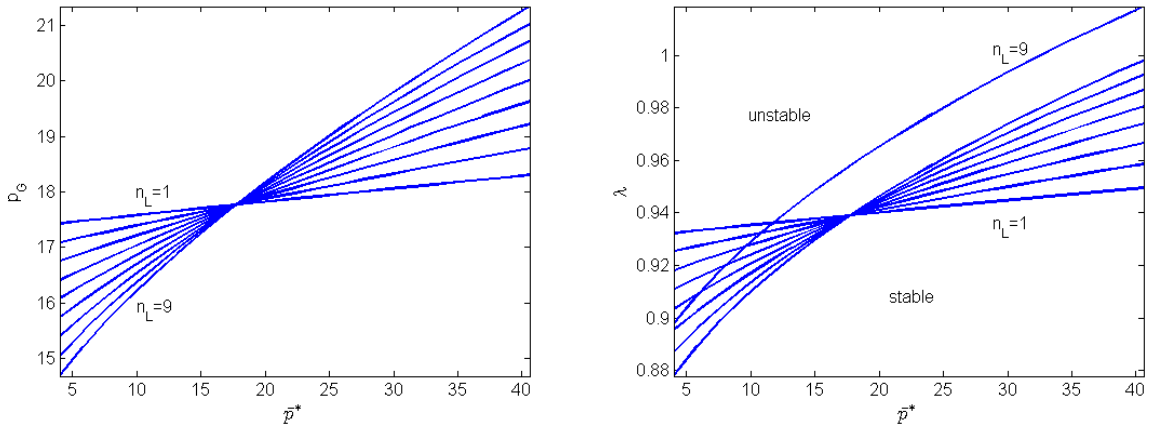


Figure 5: The gradient learners' price as a function of the average price of n_L least squared learners (left panel) and boundaries of the stability regions in the coordinates $(\bar{p}^*, \lambda)$. For the corresponding number n_L of least square learners the gradient learning algorithm is locally stable for pairs below the boundary. Parameter values: $\alpha_1 = 35$, $\alpha_2 = 4$, $\alpha_3 = 2$, $\beta = 0.7$, $\gamma = 0.6$ and $c = 4$.

As the number of LS learners changes, the average LS price changes too. Since $\bar{p}^*$ may change in any direction, we cannot say unambiguously whether a change in n_L has a stabilizing or a destabilizing effect on the price set by gradient learners. For analyzing how the average LS price changes as n_L varies, we run 1000 simulations for each value of n_L between 1 and 9 with initial prices drawn from the set of SSE prices. Let $\bar{p}_{i,j}$ denote the average LS price in run i with j LS learners, where $i = 1, \ldots, 1000$ and $j = 1, \ldots, 9$.[16] We used $\lambda = 0.937$ in these simulations: for this value of λ the convergence property of gradient learning changes as n_L varies. The simulations show that the average LS price $\bar{p}_{i,j}$ exceeds the Nash equilibrium price in $64 - 71\%$ of the runs for different number of LS learners. We analyzed the range in which the average LS price varies. For a fixed run i we considered the minimal and the maximal average LS price over the different number of LS learners: $\min_j \bar{p}_{i,j}$ and $\max_j \bar{p}_{i,j}$. This gave the interval $\left[\min_j \bar{p}_{i,j}, \max_j \bar{p}_{i,j}\right]$ in which the average LS price varies in a given run i as the number of LS learners changes. We obtained an interval for each of the 1000 runs this way. Then we considered the length of these intervals $\left(\max_j \bar{p}_{i,j} - \min_j \bar{p}_{i,j}\right)$ and calculated the mean and the standard deviation of them over the 1000 runs. The 95% confidence interval of the length is $[0.9723, 1.2656]$. Thus, as the number of LS learners changes, the average LS price does not vary much (5.47%-7.12% compared to the Nash equilibrium price of $p_N \approx 17.7693$). We compared the means of the average LS prices for different number of LS learners too: $\frac{1}{1000} \sum_{i=1}^{1000} \bar{p}_{i,j}$. The minimal value was 18.3097 (for $j = 5$) and the maximal value was 18.3778 (for $j = 2$). The means for different number of LS learners do not differ from each other significantly at the

[16] We used the same initial prices in a fixed run i for the different values of j.

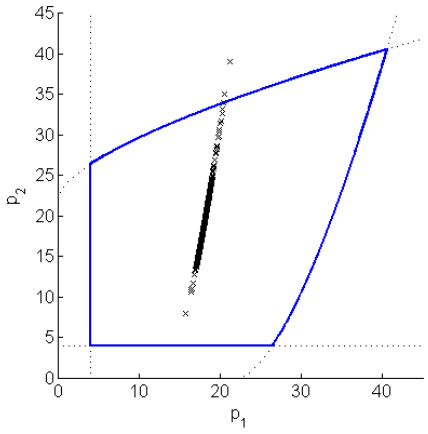


Figure 6: The end points of the simulations with firm 1 as gradient learner and firm 2 as LS learner. Parameter values: $\alpha_1 = 35$, $\alpha_2 = 4$, $\alpha_3 = 2$, $\beta = 0.7$, $\gamma = 0.6$, $c = 4$, $\lambda = \lambda_0 = 0.5$ and $\delta = 10^{-8}$.

5% level. Based on the findings that average LS prices are typically larger than the Nash equilibrium price and that they do not vary much as the number of LS learners changes, we conclude that an increase in the number of LS learners has typically (but not necessarily) a destabilizing effect on gradient learning.

The stability analysis becomes much simpler for the case of complements ($\alpha_3 < 0$). In that case it is easy to see that $0 < M_2 < M_1$ so the relevant stability condition becomes $M_1 < 2$. Numerical calculations show that the value of M_1 monotonically decreases in the number of LS learners irrespective of the average LS price. Thus, an increase in the number of LS learners has a stabilizing effect on the best response price when the average LS price is fixed. Overall, the relation between stability and the distribution of learning methods over firms is stronger for complements than for substitutes.

4.2 Simulation results

First we simulate the model for 2 firms with firm 1 as gradient learner and firm 2 as LS learner. We used $\lambda = \lambda_0 = 0.5$ in the simulations. The price set by the gradient-learning firm is locally stable for this choice of λ. We run 1000 simulations with initial prices drawn from the uniform distribution on the set of SSE prices. Figure 6 depicts the end prices and the set of SSE prices. The LS learner indeed gets close to an SSE in almost all cases: 99.7% of the points lie in the set of SSE prices. The structure of the end points also confirms that the gradient learner gives the best response price: the points lie close to the reaction curve of the gradient learner.

Figure 7 compares the profitability of the two learning methods. The left panel shows the average LS and gradient profits (with 95% confidence interval) for different number of LS learners. For drawing this graph, we simulated the model 1000 times for each number of LS learners with initial prices drawn from

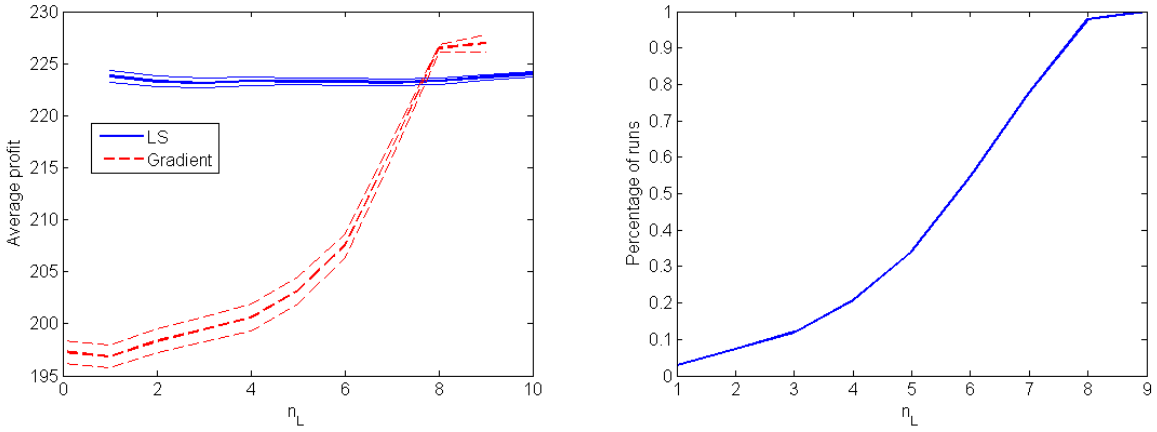


Figure 7: The average LS and gradient profits profits (with 95% confidence interval) (left panel) and the percentage of runs in which gradient learners earn a higher average profit. Parameter values: $\alpha_1 = 35$, $\alpha_2 = 4$, $\alpha_3 = 2$, $\beta = 0.7$, $\gamma = 0.6$, $c = 4$ and $\lambda = \lambda_0 = 0.937$.

the uniform distribution on the set of SSE prices. We let each simulation run for 2000 periods and for each firm we considered the average of its profits over the last 100 periods as the profit of the firm in the given simulation.[17] Thus, for the case of k LS learners, we had $1000k$ observations for LS profits and $1000(10-k)$ observations for gradient profits. We calculated the average and the standard deviation of these values separately for LS and gradient learners. The confidence interval is calculated as $mean \pm 2stdev/\sqrt{1000k}$ and $mean \pm 2stdev/\sqrt{1000(10-k)}$ for LS and gradient learners respectively. The left panel shows that gradient learning yields significantly lower average profit than LS learning when the number of LS learners is low. In contrast, it gives significantly higher profits when the number of LS learners is high enough.

The right panel of Figure 7 depicts for each number of LS learners the percentage of the 1000 simulations in which the average gradient profit was larger than the average LS profit. The graph shows that gradient learning becomes more profitable than LS learning more often as the number of LS learners increases. Since profitability is closely related to the convergence properties of gradient learning, this illustrates that an increase in the number of LS learners has typically a stabilizing effect.[18]

Based on this change in the stability of gradient learning, we conjecture a cyclical switching between

[17] 2000 periods are typically enough for profits to converge when gradient learning converges. We take the average over the last 100 periods in order to get a better view on the profitability of the methods. When gradient learning converges, then profits do not vary much in the last periods. When the price is unstable, gradient profits change more or less periodically, so averaging over the last few profits describes the profitability of the method better than considering only the last profit.

[18] When gradient learning does not converge, then it gives low average profit as the price fluctuates between too low and too high values. Therefore, when the average gradient profits are high, the best-response price must be locally stable and gradient learning must converge.

the learning methods when firms are allowed to choose which method they want to apply. Conjecture 4.3 summarizes our expectation. In the following section we will investigate if cyclical switching occurs.

Conjecture 4.3 *When firms are sensitive to profit differences, changes in the convergence properties of gradient learning may lead to cyclical switching between the learning rules. When gradient learning converges, LS learners have an incentive to switch to gradient learning as it typically yields a higher profit. This increase in the number of gradient learners, however, may destabilize the best-response price, resulting in lower gradient profits. Then firms switch to LS learning, so gradient learning may converge again and the cycle may repeat itself.*

5 Endogenous switching between learning mechanisms

We introduce competition between the learning rules in this section. We extend the model by allowing for endogenous switching between the two methods: firms may choose from the two learning rules in each period. For deciding about the rules, firms take into account their performance: the probability of choosing a specific method is positively related to the past profits realized while using that method. Section 5.1 specifies the switching mechanism, the simulation results are discussed in Section 5.2. The simulations confirm that the cyclical switching we conjectured occurs.

5.1 The switching mechanism

The switching mechanism is based on reinforcement learning as in Roth and Erev (1995) and it is related to the discrete choice model as in Brock and Hommes (1997). The mechanism is augmented with experimentation too. Every firm i has a performance measure for each of the rules. These measures determine the probability of firm i applying a certain rule. Performances depend on past realized profits. Let $l_{i,t}$ ($g_{i,t}$) denote the performance of LS (gradient) learning perceived by firm i at the end of period t. The performance measure for LS learning is updated in each period in the following way:

$$l_{i,t} = \begin{cases} (1-w)l_{i,t-1} + w\pi_{i,t} & \text{if firm } i \text{ used LS learning in period } t \\ l_{i,t-1} & \text{otherwise} \end{cases},$$

where $w \in (0,1]$ is the weight of the latest profit in the performance measure. The performance of gradient learning is updated analogously. The initial performances are the first profits that were realized using the method in question for each firm. Thus, performance measures are basically weighted averages of past profits realized by the given method where weights decay geometrically.

These performance measures determine the probability of applying a learning method in the following way. Firm i applies LS learning in period $t+1$ with probability

$$P_{i,t+1}^{\mathrm{LS}} = (1-2\eta)\frac{1}{\exp\left[\omega\left(g_{i,t}-l_{i,t}\right)\right]+1}+\eta, \tag{13}$$

where $\omega \geq 0$ measures how sensitive the firms are to differences in the performance measures and η is the probability of experimentation. The higher ω is, the higher the probability of applying the method with the higher performance. For $\omega = 0$ firms choose both methods with 50% probability. When $\omega = +\infty$, then firms choose the method with the higher performance with probability $1-\eta$. The interpretation of (13) is that the choice is based on the performance difference between the methods with probability $1-2\eta$ and the firm randomizes with equal probabilities between the methods with probability 2η.

The model with endogenous switching is implemented as follows.

1. $p_{i,1}$ and $p_{i,2}$ are drawn from the uniform distribution on the set S, for each i.

2. In period 3, k randomly chosen firms apply LS learning, the other firms use gradient learning. LS and gradient prices are determined by the learning mechanisms discussed in Section 3.

3. In period 4:

   a. Firms try the other method: all LS learners switch to gradient learning and vice versa. Prices are determined by the two learning mechanisms. The initial performances are $l_{i,4} = \pi_i^{\mathrm{LS}}$ and $g_{i,4} = \pi_i^{\mathrm{grad}}$.[19]

   b. Firms choose a method for the following period: firm i applies LS learning in period 5 with probability $P_{i,5}^{\mathrm{LS}}$.

4. In period $t \geq 5$:

   a. Prices are determined by the two learning mechanisms. The performance measures l_t and g_t are updated.

   b. Firm i chooses LS learning for period $t+1$ with probability $P_{i,t+1}^{\mathrm{LS}}$.

5. The process stops when a predefined number of periods T is reached.

In the simulations of the following section we use $w = 0.5$, $\omega = 25$ and $\eta = 0.005$. We simulate the model for $T = 10000$ periods.

Erev et al. (2010) find evidence for inertia in subjects' choices in market entry experiments: subjects tend to choose the same action unless there is a substantial drop in the corresponding payoff. Anufriev

[19] π_i^{LS} (π_i^{grad}) denotes the profit of firm i that was earned while using LS (gradient) learning in period 3 or 4.

and Hommes (2012) also find substantial evidence of inertia in forecasting behavior by estimating the individual learning model of heuristic switching on data from learning to forecast experiments. Inertia could be incorporated in the switching mechanism, we discuss its effects in the concluding remarks.

5.2 Learning cycles

First we shortly discuss the results of simulations when the convergence properties of gradient learning do not change as the distribution of learning methods over firms varies and then we illustrate cyclical switching. When gradient learning always converges (i.e. for any number of LS learners), then LS learning need not be driven out even if the firms are very sensitive to performance differences. Some firms may earn a high LS profit and they apply LS learning not only due to experimentation but in many consecutive periods. However, the number of LS learners is typically low. In contrast, when gradient learning diverges fast and firms are sufficiently sensitive to performance differences, then gradient learning is driven out by LS learning. Firms apply gradient learning only due to experimentation, each firm uses LS learning in almost every period. When firms are less sensitive to differences in the performance measures, then gradient learning is used more often but LS learning is applied in the vast majority of the periods.

Now let us consider the case when the convergence properties of gradient learning change as the distribution of learning methods over firms varies. First we illustrate cyclical switching in a duopoly because it is easier to see what drives the firms' switching behavior when the number of firms is low. Then we show that cyclical switching occurs for higher number of firms too. We use the same demand and cost parameters as before. Figure 8 depicts typical time series of prices and the corresponding performance measures for the case of two firms. We used $\lambda = 0.85$ and $k = 1$ in the simulation. Gradient learning is stable for this value of λ only if one firm uses it. In the first third of the illustrated periods firm 1 uses mainly LS learning while firm 2 is a gradient learner. Firm 1 tries gradient learning in one period but it immediately switches back to LS learning as the latter performs better. This change in the price of firm 1 drives away the price of firm 2 from the best-response price and it takes a few periods until the gradient learner reaches the optimal price again. Later firm 1 tries gradient learning again and this induces a change in prices after which the firm becomes a gradient learner. When both firms apply gradient learning, prices start an oscillating divergence. At some point the performance of gradient learning becomes worse than that of LS learning and firm 1 switches back to LS learning. This ends the first oscillating part. Gradient learning, however, becomes more profitable again for firm 1 and another oscillating part starts. This part ends in the same way: firm 1 switches back to LS learning after which the price set by firm 2 starts to converge. The last oscillating part starts by firm 2 switching to LS learning. The price set by firm 2

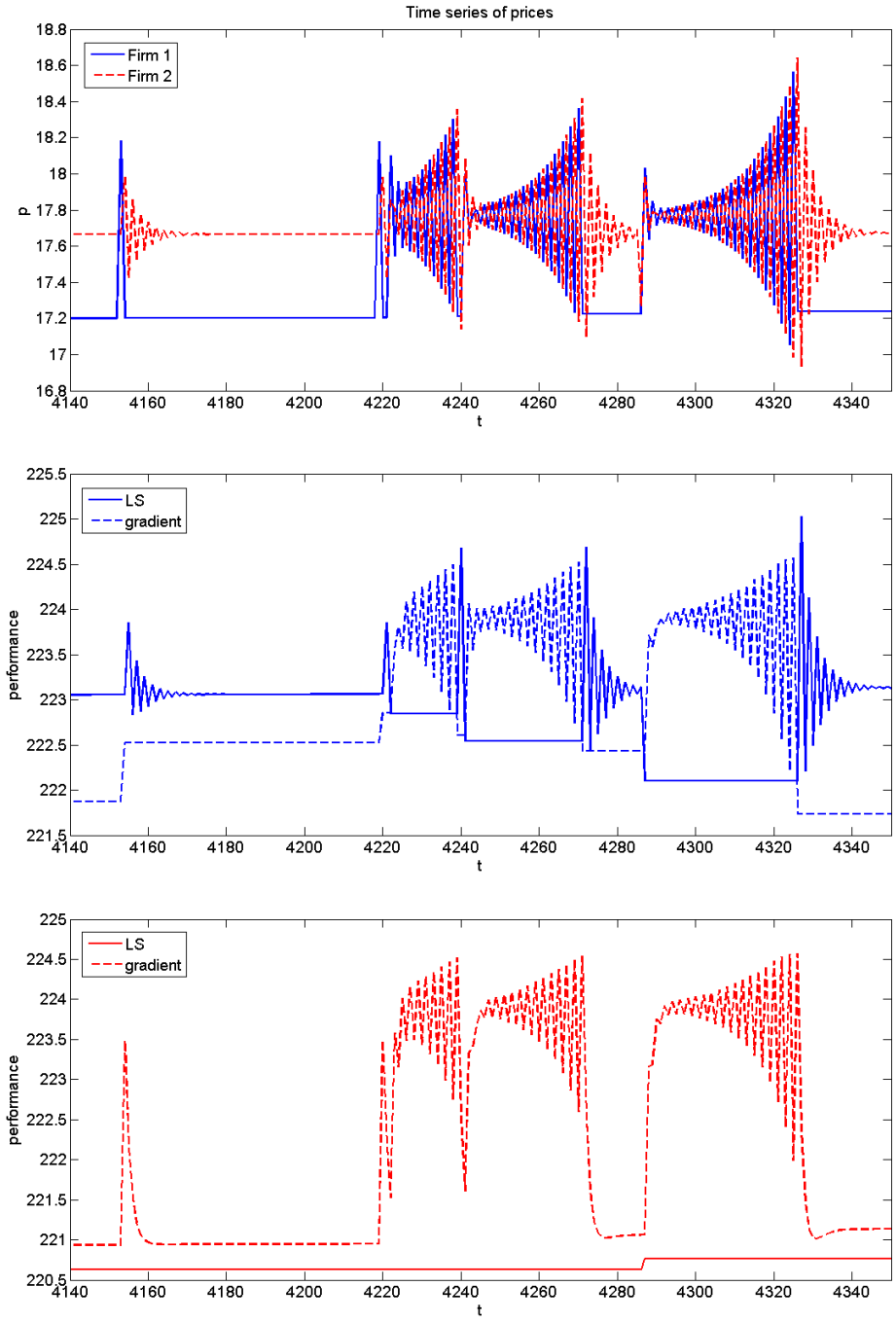


Figure 8: Cyclical switching in a duopoly. Time series of prices (upper panel), the performance measures of firm 1 (middle panel) and firm 2 (lower panel). Parameter values: $\alpha_1 = 35$, $\alpha_2 = 4$, $\alpha_3 = 2$, $\beta = 0.7$, $\gamma = 0.6$, $c = 4$, $\lambda = \lambda_0 = 0.85$ and $k = 1$.

decreases which yields a lower profit for firm 1. Because of this firm 1 switches to gradient learning.

Cyclical switching can occur for higher number of firms too. Figure 9 illustrates this for 10 firms with $\lambda = 0.95$ and $k = 5$. We can observe both diverging and converging phases for gradient learners which

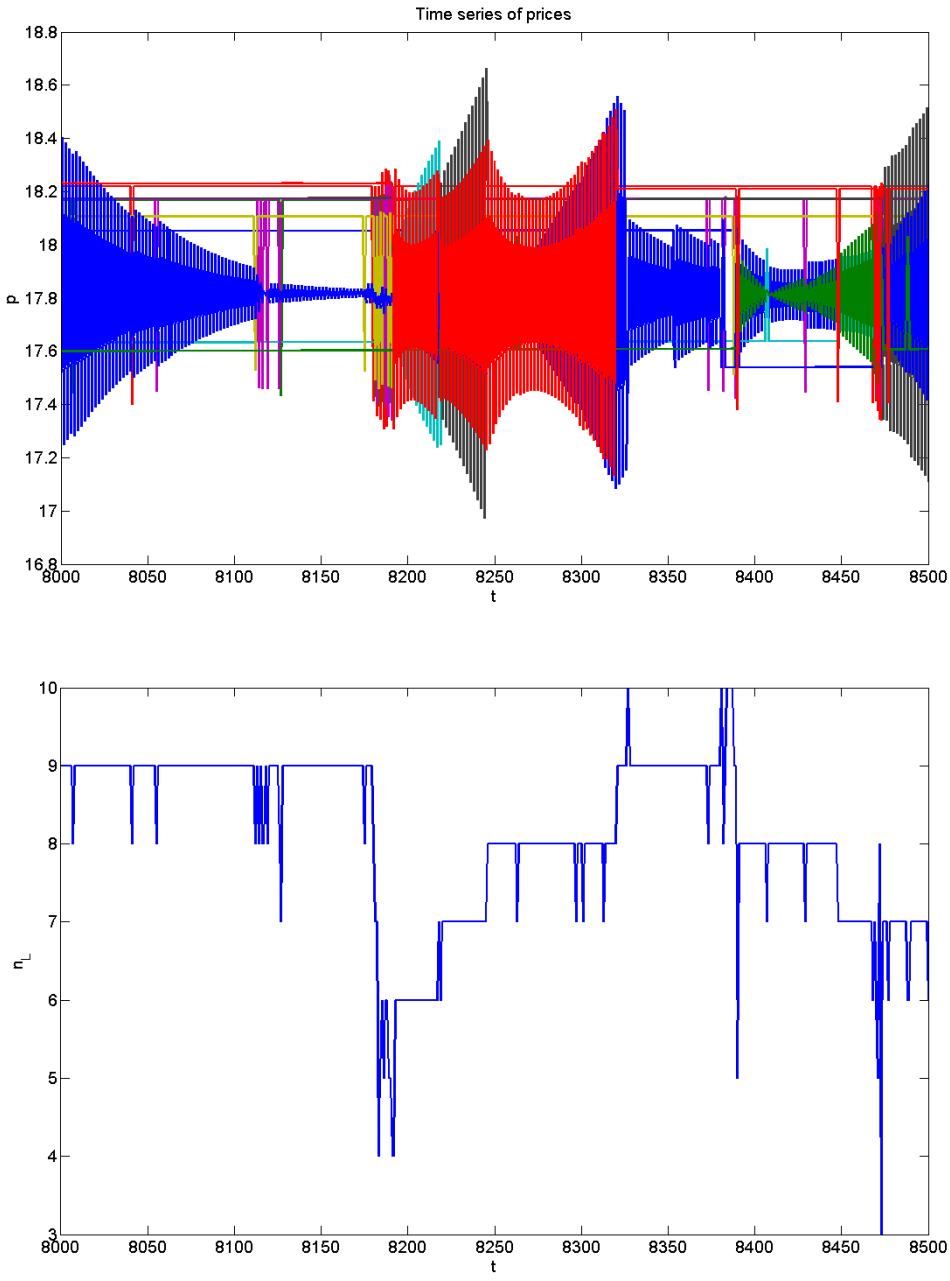


Figure 9: Cyclical switching with 10 firms. Time series of prices (upper panel) and number of LS learners (lower panel). Parameter values: $\alpha_1 = 35$, $\alpha_2 = 4$, $\alpha_3 = 2$, $\beta = 0.7$, $\gamma = 0.6$, $c = 4$, $\lambda = \lambda_0 = 0.95$ and $k = 5$.

shows that the stability of the method changes. This change is related to the number of LS learners. In the first periods, the number of LS learners is high and we observe that the gradient learners' prices converge. Then some LS learners switch to gradient learning, which is reflected in the drop in n_L. As the time series of prices show, gradient learning becomes unstable. After that the number of LS learners starts increasing gradually until it reaches the level $n_L = 9$, for which the gradient price shows a converging pattern again.

Then firms start switching to gradient learning again, which destabilizes the price. We could find evidence for cyclical switching for the case of complements too.

Note that cyclical switching may occur only if the value of parameter λ is such that gradient learning converges when there are few gradient learners and it diverges otherwise. For any parameter values that satisfy Assumption 2.1, we can find values of λ for which the gradient learners' price is locally stable when the number of gradient learners is low and unstable otherwise, for a given average LS price. Note, however, that this change in stability does not ensure that cyclical switching occurs: local stability does not imply that convergence occurs for *any* initial values. Nevertheless, for any parameter values that satisfy Assumption 2.1, there exist values of λ for which cyclical switching may in general occur, but it may be harder to find such values of λ for some parameter values than for others.

6 Concluding remarks

Due to the richness of possible learning methods, agents may prefer to use different ways of learning about their environment. In this paper we demonstrate that several learning methods can coexist, in the sense that there is no clear winner in the profit-driven evolutionary competition between the methods. We stress that this coexistence may have a substantial effect on the dynamical properties of the learning methods and that the dynamics with heterogeneity in learning methods is much more complex than under homogeneous learning.

In this paper we have analyzed the interaction between least squares learning and gradient learning in a Bertrand oligopoly with differentiated goods where firms do not know the demand specification and they use one of the two methods for determining the price of their good. These learning methods have been widely used for modeling learning behavior in oligopolistic markets, but mainly in a homogeneous setup. The methods that we have chosen are not exceptional in the sense that other learning methods may lead to similar results: best response learning, for example, would yield similar outcomes in the current model as a stable gradient learning.

We have analyzed four different setups. In a pure LS-learning oligopoly firms move towards a *self-sustaining equilibrium* in which their expected and actual demands coincide at the prices they charge. The set of SSE prices contains infinitely many points including the Nash equilibrium of the model. The initial conditions determine which point is reached in the long run. We formally prove that firms reach the Nash equilibrium when every firm applies gradient learning and the method converges. When gradient learning does not converge, then it leads to high-period cycles, quasi-periodic or aperiodic dynamics. In a heterogeneous oligopoly with firms applying a fixed learning method, we have analytically derived that

the dynamical properties of gradient learning depend on the distribution of learning methods over firms. Numerical analysis shows that an increase in the number of LS learners can have a stabilizing effect. When gradient learning converges, then LS learners move towards a *self-sustaining equilibrium* in which gradient learners give the best response to the prices of other firms. When endogenous switching between the learning methods is introduced in the model, then a stable gradient learning may not always drive out LS learning: some LS learners may find LS learning to be more profitable for them. LS learning, however, may drive out gradient learning when the latter never converges. When the convergence properties and the profitability of gradient learning changes as the distribution of learning methods over firms varies, a cyclical switching between the learning methods may be observed. Gradient learners tend to switch to LS learning when gradient learning does not converge and thus gives low profits. This decrease in the number of gradient learners can stabilize the method, resulting in higher profits. This can give an incentive for LS learners to switch back to gradient learning. Gradient learning, however, may lose its stability again and the cycle may repeat itself.

The previous analysis can be extended in several ways. Observations could have different weights in the LS formulas. Since observations of the early periods are less informative about the demand function due to the volatility of the prices of other firms, it might be reasonable to introduce a weighting function that gives less weight to older observations. Erev et al. (2010) find specific behavioral regularities in market entry experiments where subjects need to make decisions based on their experience. Since firms have similar tasks in our model, implementing some of these regularities can make the learning methods and the switching behavior empirically more relevant. For example, we could consider inertia in the switching behavior: firms tend to keep their current learning method unless there is a large drop in its profitability. The results of our paper remain valid since inertia does not affect the main driving factor of cyclical switching: the stability of gradient learning. However, it would take longer time to observe cyclical switching as there would be less switching in the model due to inertia. Another factor that can be considered is individual differences within the class of learning methods. These individual differences could be incorporated in the current model using different kinds of perceived demand functions such as higher-order polynomials in LS learning and different values of adjustment parameter λ for different firms in gradient learning. In case of LS learning, firms would still reach a self-sustaining equilibrium when the perceived demand functions are misspecified. If the adjustment parameter λ varies among firms, it can be shown that the stability of gradient learning depends on the *distribution* of adjustment coefficients. If the stability condition is satisfied, then every gradient learner reaches the best response price. If, however, the stability condition is not satisfied, then none of the prices of gradient learners converge. In this latter case, the amplitude of the price oscillation and the profitability of the method is negatively related to the value of the adjustment

parameter. Even when we implement heterogeneity in the adjustment parameter of gradient learners, we can observe switching between the learning methods. Firms with very high and very low values of λ almost always use LS learning, whereas firms with intermediate values of λ keep switching between the learning methods as the convergence properties of gradient learning vary. It might also be reasonable to change the information that is available to firms. If LS learners observe the price of some but not all of the goods, then they may take this extra information into account for estimating the demand parameters. The step size parameter of gradient learning could be endogenized such that the method would stabilize itself automatically when prices start to diverge. Other learning methods such as best-response learning, fictitious play or imitation could also be applied in the current setup. We could add less myopic decision makers to the model such as perfect foresight firms or firms that are actively learning, i.e. firms that want to maximize a discounted sum of profits.

The analysis can be extended to different market structures as well. It might be interesting to analyze what happens under Cournot competition when the quantities set by firms are strategic substitutes. Moreover, learning in more complex environments where firms make not only a price or quantity choice but they also need to make investment, quality or location decisions, can be studied as well along the lines outlined in this paper.

References

Anufriev, M., Bao, T., and Tuinstra, J. (2013). Switching behavior in the lab: Estimation of the intensity of choice. *CeNDEF working paper*.

Anufriev, M. and Hommes, C. (2012). Evolutionary selection of individual expectations and aggregate outcomes in asset pricing experiments. *American Economic Journal: Microeconomics*, 4(4):35–64.

Arrow, K. and Hurwicz, L. (1960). Stability of the gradient process in n-person games. *Journal of the Society for Industrial and Applied Mathematics*, 8(2):280–294.

Bischi, G. I., Chiarella, C., Kopel, M., and Szidarovszky, F. (2009). *Nonlinear oligopolies: stability and bifurcations*. Springer.

Bischi, G. I., Mammana, C., and Gardini, L. (2000). Multistability and cyclic attractors in duopoly games. *Chaos, Solitons & Fractals*, 11(4):543–564.

Bonanno, G. (1988). Oligopoly equilibria when firms have local knowledge of demand. *International Economic Review*, 29(1):45–55.

Bonanno, G. and Zeeman, E. C. (1985). Limited knowledge of demand and oligopoly equilibria. *Journal of Economic Theory*, 35(2):276–283.

Brock, W. and Hommes, C. (1997). A rational route to randomness. *Econometrica*, 65:1059–1095.

Brousseau, V. and Kirman, A. (1992). Apparent convergence of learning processes in mis-specified games. In Dutta, B., (editor), *Game Theory and Economic Applications*. Springer-Verlag.

Corchon, L. and Mas-Colell, A. (1996). On the stability of best reply and gradient systems with applications to imperfectly competitive models. *Economics Letters*, 51(1):59–65.

Cressman, R. (2003). *Evolutionary Dynamics and Extensive Form Games*. The MIT Press.

Droste, E., Hommes, C., and Tuinstra, J. (2002). Endogenous fluctuations under evolutionary pressure in Cournot competition. *Games and Economic Behavior*, 40(2):232–269.

Erev, I., Ert, E., and Roth, A. E. (2010). A choice prediction competition for market entry games: An introduction. *Games*, 1(2):117–136.

Evans, G. and Honkapohja, S. (2001). *Learning and Expectations in Macroeconomics*. Princeton University Press.

Fudenberg, D. and Levine, D. (1998). *The Theory of Learning in Games*. The MIT press.

Furth, D. (1986). Stability and instability in oligopoly. *Journal of Economic Theory*, 40(2):197–228.

Gal-Or, E. (1985). First mover and second mover advantages. *International Economic Review*, 26(3):649–653.

Gates, D., Rickard, J., and Wilson, D. (1977). A convergent adjustment process for firms in competition. *Econometrica*, 45(6):1349–1363.

Kirman, A. (1983). On mistaken beliefs and resultant equilibria. In Frydman, R. and Phelps, E. (editors), *Individual Forecasting and Collective Outcomes*. Cambridge University Press.

Kopányi, D. (2013). Heterogeneous learning in Bertrand competition with differentiated goods. In Teglio, A., Alfarano, S., Camacho-Cuena, E., and Ginés-Vilar, M. (editors), *Managing market complexity: The approach of artificial economics*. Springer-Verlag.

Marcet, A. and Sargent, T. (1989). Convergence of least squares learning mechanisms in self-referential linear stochastic models. *Journal of Economic Theory*, 48(2):337–368.

Mira, C. (1987). *Chaotic Dynamics*. World Scientific.

Offerman, T., Potters, J., and Sonnemans, J. (2002). Imitation and belief learning in an oligopoly experiment. *Review of Economic Studies*, 69(4):973–997.

Roth, A. and Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, 8(1):164–212.

Stahl, D. (1996). Boundedly rational rule learning in a guessing game. *Games and Economic Behavior*, 16:303–330.

Stock, J. and Watson, M. (2003). *Introduction to Econometrics*. Addison Wesley.

Tuinstra, J. (2004). A price adjustment process in a model of monopolistic competition. *International Game Theory Review*, 6:417–442.

Appendix

The proof of Proposition 2.2

The profit of firm i is given by $\pi_i(p) = (p_i - c)\left(\alpha_1 - \alpha_2 p_i^\beta + \alpha_3 \bar{p}_{-i}^\gamma\right)$.[20] The first-order condition with respect to p_i is

$$\alpha_1 - \alpha_2 p_i^\beta + \alpha_3 \bar{p}_{-i}^\gamma - \alpha_2 \beta (p_i - c) p_i^{\beta-1} = 0. \tag{14}$$

This equation needs to hold for all firms. We will show that firms choose the same price in equilibrium.

Consider two arbitrary firms i and j and suppose indirectly that $p_i > p_j$ in equilibrium. Let $y = \sum_{k=1}^{n} p_k - p_i - p_j$. Then the first-order conditions for firms i and j read as

$$\alpha_1 - \alpha_2 p_i^\beta + \alpha_3 \left(\frac{p_j + y}{n-1}\right)^\gamma - \alpha_2 \beta (p_i - c) p_i^{\beta-1} = 0, \tag{15}$$

$$\alpha_1 - \alpha_2 p_j^\beta + \alpha_3 \left(\frac{p_i + y}{n-1}\right)^\gamma - \alpha_2 \beta (p_j - c) p_j^{\beta-1} = 0. \tag{16}$$

Subtracting (15) from (16) yields

$$\alpha_2 \left(p_i^\beta - p_j^\beta\right) + \alpha_3 \left[\left(\frac{p_i + y}{n-1}\right)^\gamma - \left(\frac{p_j + y}{n-1}\right)^\gamma\right] + \alpha_2 \beta \left[(p_i - c) p_i^{\beta-1} - (p_j - c) p_j^{\beta-1}\right] = 0.$$

[20] We assume in this formula that demands are positive in a Nash equilibrium. Later we will see that this is indeed the case.

The first two terms are positive as $p_i > p_j$ and all parameters are positive. We will now show that the last term is also positive. Let $g(x) = (x-c)x^{\beta-1}$. This function is increasing if $x \geq c$: $g'(x) = \beta x^{\beta-1} - c(\beta-1)x^{\beta-2} > 0$ for $x > c(1-\frac{1}{\beta})$.[21] This proves that the last term is also positive as $p_i > p_j$. This, however, leads to a contradiction as positive numbers cannot add up to zero. So we must have $p_i = p_j$: firms charge the same price in a Nash equilibrium. Let p denote the corresponding price. Then (14) gives

$$f(p) \equiv \alpha_1 - \alpha_2 p^\beta + \alpha_3 p^\gamma - \alpha_2\beta(p-c)p^{\beta-1} = 0. \tag{17}$$

We will now show that there is a unique solution to this equation and the corresponding price is larger than the marginal cost. According to Assumption 2.1, $f(c) = \alpha_1 - \alpha_2 c^\beta + \alpha_3 c^\gamma > 0$. Note that $f(p)$ becomes negative for high values of p :

$$f(p) = \alpha_1 - p^\gamma \left[\alpha_2 p^{\beta-\gamma} - \alpha_3 + \alpha_2\beta\left(1-\frac{c}{p}\right)p^{\beta-\gamma}\right],$$

from which it is easy to see that $\lim_{p\to+\infty} f(p) = -\infty$. The derivative of $f(p)$ is $f'(p) = -\alpha_2\beta p^{\beta-1} + \alpha_3\gamma p^{\gamma-1} - \alpha_2\beta p^{\beta-1}\left[1+(\beta-1)\left(1-\frac{c}{p}\right)\right]$. Assumption 2.1 ensures that the sum of the first two terms is negative. The last term is also negative when $p > c$. Thus, $f(p)$ is strictly decreasing in p for $p > c$. Since $f(p)$ is continuous, this proves that there is a unique solution to $f(p) = 0$. Let p^N denote the symmetric Nash equilibrium price. It follows easily from the proof that $p^N > c$ and the demands are positive in the Nash equilibrium.

We will show that the second order condition is satisfied. Differentiating (14) with respect to p_i yields

$$-2\alpha_2\beta p_i^{\beta-1} - \alpha_2\beta(\beta-1)(p_i-c)p_i^{\beta-2} = -\alpha_2\beta p_i^{\beta-1}\left(2+(\beta-1)\frac{p_i-c}{p_i}\right)$$

This is negative for $p = p^N$ since the term in brackets is positive: $\frac{p^N-c}{p^N} \in (0,1)$ as $p^N > c$ and $\beta - 1 > -1$, so $(\beta-1)\frac{p^N-c}{p^N} > -1$. □

The proof of Proposition 3.1

First we derive the coefficients of the perceived demand functions in an SSE in terms of the SSE prices and then we study which prices may constitute an SSE.

From (9) we get $a_i^* = D_i(p^*) + b_i^* p_i^*$. Combining this expression with (8) yields

$$b_i^* = \frac{D_i(p^*)}{p_i^* - c}. \tag{18}$$

Using (18) we can express a_i^* as

$$a_i^* = D_i(p^*)\left(1 + \frac{p_i^*}{p_i^* - c}\right). \tag{19}$$

[21] We will see later that the condition $x \geq c$ holds for the Nash equilibrium price.

The above described values constitute an SSE only if the inverse demand functions are sensible. That is, the following conditions need to be satisfied for all firms:

$$a_i^* > 0, \tag{20}$$

$$b_i^* > 0, \tag{21}$$

$$P_i^P(0) = \frac{a_i^*}{b_i^*} > c, \tag{22}$$

$$P_i(0, \bar{p}_{-i}^*) = \left[\frac{1}{\alpha_2}\left(\alpha_1 + \alpha_3(\bar{p}_{-i}^*)^\gamma\right)\right]^{\frac{1}{\beta}} > c, \tag{23}$$

$$p_i^* > c. \tag{24}$$

Conditions (20) and (21) ensure that the perceived demand functions are downward-sloping with a positive intercept. Conditions (22) and (23) require that the perceived and the actual inverse demands are larger than the marginal cost at $q_i = 0$. Condition (24) specifies that the SSE prices should be larger than the marginal cost. We will show that some of these constraints are redundant.

Conditions (20) and (21) hold true if and only if $D_i(p^*) > 0$ and $p_i^* > c$. Combining (18) and (19) yields

$$\frac{a_i^*}{b_i^*} = p_i^* - c + p_i^* = 2p_i^* - c.$$

This shows that (22) is equivalent to (24). We can express $D_i(p^*) > 0$ as

$$\left[\frac{1}{\alpha_2}\left(\alpha_1 + \alpha_3(\bar{p}_{-i}^*)^\gamma\right)\right]^{\frac{1}{\beta}} > p_i^*.$$

Combining this with $p_i^* > c$ shows that (23) is satisfied. Thus, the set of SSE prices is given by $p_i^* > c$ and $D_i(p^*) > 0$, or equivalently $c < p_i^* < \left[\frac{1}{\alpha_2}\left(\alpha_1 + \alpha_3(\bar{p}_{-i}^*)^\gamma\right)\right]^{\frac{1}{\beta}}$.

This set is nonempty: the Nash equilibrium price, for example, satisfies the above condition. The maximal SSE price of firm i increases in the price of other firms. Thus, the upper bound of the SSE prices is given by the price $\hat{p}$ for which the demand is 0 if every firm charges this price: $\alpha_1 - \alpha_2\hat{p}^\beta + \alpha_3\hat{p}^\gamma = 0$. The existence and uniqueness of this price can be shown in the same way as for the Nash equilibrium price. □

The proof of Propositions 3.2 and 4.2

Let us consider a heterogeneous setting in which the first n_L firms apply LS learning and the remaining $n - n_L$ firms use gradient learning. The proof of Proposition 3.2 follows from this general case by setting $n_L = 0$.

In the proof we will apply a lemma about the eigenvalues of a matrix that has a special structure. First we will prove this lemma and then we prove Propositions 3.2 and 4.2.

Lemma 6.1 *Consider an $n \times n$ matrix with diagonal entries $d \in \mathbb{R}$ and off-diagonal entries $o \in \mathbb{R}$. In case of $n = 1$ the matrix has one eigenvalue: $\mu = d$. If $n > 1$, then there are two distinct eigenvalues: $\mu_1 = d + (n-1)o$ (with multiplicity 1), and $\mu_2 = d - o$ (with multiplicity $n-1$).*

The proof of Lemma 6.1

The case $n = 1$ is trivial so we focus on $n > 1$. Let $\mathbf{A}$ denote the matrix in question. Due to its special structure, $\mathbf{A}$ can be expressed as $\mathbf{A} = (d-o)\mathbf{I_n} + o\mathbf{1_n}$, where $\mathbf{I_n}$ is the $n-$dimensional identity matrix and $\mathbf{1_n}$ is the n-dimensional matrix of ones.

First note that if λ is an eigenvalue of $o\mathbf{1_n}$ with corresponding eigenvector $\mathbf{x}$, then $\mathbf{x}$ is an eigenvector of $\mathbf{A}$ for the eigenvalue $d - o + \lambda$: if $o\mathbf{1_n}\mathbf{x} = \lambda\mathbf{x}$, then $\mathbf{A}\mathbf{x} = ((d-o)\mathbf{I_n} + o\mathbf{1_n})\,\mathbf{x} = (d-o)\mathbf{x} + \lambda\mathbf{x} = (d - o + \lambda)\,\mathbf{x}$.

It is easy to see that $o\mathbf{1_n}$ has two distinct eigenvalues: $o \cdot n$ with multiplicity 1 and 0 with multiplicity $n-1$. Thus, $\mathbf{A}$ has two distinct eigenvalues: $\mu_1 = d - o + o \cdot n = d + (n-1)o$ with multiplicity 1 and $\mu_2 = d - o$ with multiplicity $n-1$. □

Now the dynamical properties of the heterogeneous oligopoly can be studied in the following way. Suppose that LS prices have settled down at some level and let p_i^* denote the price of LS learner i $(i = 1, \ldots, n_L)$. Since LS prices have settled down, the law of motion of the prices set by LS learners can be approximated by $p_{i,t+1} = p_{i,t}$ for $i = 1, \ldots, n_L$ as price changes become smaller as the number of observations increases. The law of motion of the price set by gradient learners is given by $p_{j,t+1} = p_{j,t} + \lambda \frac{\partial \pi_j(p_t)}{\partial p_{j,t}}$ for $j = n_L + 1, \ldots, n$. Then the Jacobian (evaluated at the steady state) is of the following form:

$$\mathbf{J} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{B} & \mathbf{A} \end{pmatrix},$$

where $\mathbf{I}$ is the $n_L \times n_L$ identity matrix, $\mathbf{0}$ is an $n_L \times (n - n_L)$ matrix of zeros, $\mathbf{B}$ is an $(n - n_L) \times n_L$ matrix with all entries equal to

$$o = \lambda \frac{\partial^2 \pi_j(p)}{\partial p_i \partial p_j} = \lambda \alpha_3 \gamma \frac{1}{n-1} \left(\frac{\sum_{s=1}^{n_L} p_s^* + (n - n_L - 1)p_G}{n-1} \right)^{\gamma - 1},$$

and $\mathbf{A}$ is an $(n - n_L) \times (n - n_L)$ matrix with diagonal entries

$$d = 1 + \lambda \frac{\partial^2 \pi_j(p)}{\partial p_j^2} = 1 - \alpha_2 \beta \lambda p_G^{\beta - 1} \left(2 + (\beta - 1) \frac{p_G - c}{p_G} \right)$$

and off-diagonal entries equal to o.

Due to its special structure, the eigenvalues of J are given by the eigenvalues of I and the eigenvalues of A. The stability properties of gradient learning are determined fully by the eigenvalues of A. Applying Lemma 6.1, the eigenvalues that determine the stability of gradient learning are $\mu_1 = d + (n - n_L - 1)o$ with multiplicity 1 and $\mu_2 = d - o$ with multiplicity $n - n_L - 1$. If $n - n_L = 1$, then the unique eigenvalue is $\mu = d$.

When $n - n_L = 1$, the stability condition becomes

$$\lambda\alpha_2\beta p_G^{\beta-1}\left(2 + (\beta-1)\frac{p_G - c}{p_G}\right) < 2.$$

The primary bifurcation is a period-doubling bifurcation: the eigenvalue becomes -1. When $n - n_L > 1$, the stability conditions $-1 < \mu_i < 1$ simplify to $0 < \lambda M_1 < 2$ and $0 < \lambda M_2 < 2$ where

$$M_1 = \alpha_2\beta p_G^{\beta-1}\left[2 + (\beta-1)\frac{p_G - c}{p_G}\right] - \alpha_3\gamma\left(\frac{n_L\bar{p}^* + (n - n_L - 1)p_G}{n-1}\right)^{\gamma-1}\frac{n - n_L - 1}{n-1},$$
$$M_2 = \alpha_2\beta p_G^{\beta-1}\left[2 + (\beta-1)\frac{p_G - c}{p_G}\right] + \alpha_3\gamma\left(\frac{n_L\bar{p}^* + (n - n_L - 1)p_G}{n-1}\right)^{\gamma-1}\frac{1}{n-1},$$

and $\bar{p}^* = \frac{1}{n_L}\sum_{s=1}^{n_L} p_s^*$ is the average LS price.

By setting $n_L = 0$ it is easy to see that the above expressions simplify to

$$M_1 = \alpha_2\beta p_N^{\beta-1}\left[2 + (\beta-1)\frac{p_N - c}{p_N}\right] - \alpha_3\gamma p_N^{\gamma-1},$$
$$M_2 = \alpha_2\beta p_N^{\beta-1}\left[2 + (\beta-1)\frac{p_N - c}{p_N}\right] + \alpha_3\gamma p_N^{\gamma-1}\frac{1}{n-1}$$

for the case of a homogeneous gradient-learning oligopoly, where p_N is the symmetric Nash equilibrium price. Since $\alpha_3 > 0$, $M_2 > M_1$. It follows from Assumption 2.1 that $\alpha_2\beta p^{\beta-1} > \alpha_3\gamma p^{\gamma-1}$ for all $p \geq c$. This ensures that M_1 is always positive:

$$M_1 = \alpha_2\beta p_N^{\beta-1}\left[2 + (\beta-1)\frac{p_N - c}{p_N}\right] - \alpha_3\gamma p_N^{\gamma-1} > \alpha_2\beta p_N^{\beta-1}\left[1 + (\beta-1)\frac{p_N - c}{p_N}\right] > 0$$

since $1 + (\beta-1)\frac{p_N - c}{p_N} > 0$. Thus, the relevant stability condition in the homogeneous case is $\lambda M_2 < 2$. At the bifurcation value of λ, $n-1$ eigenvalues become -1 while the remaining eigenvalue is positive and smaller than one. □