Coexisting stable equilibria under least squares learning*

Dávid Kopányi†

CeNDEF, University of Amsterdam and Tinbergen Institute

July 8, 2014

Abstract

This paper illustrates that least squares learning may lead to sub-optimal outcomes even when firms observe all the variables that affect their demand and they use a locally correct functional form in the estimation.

We consider the Salop model with three firms and two types of consumers that face different transportation costs. Firms do not know the demand structure and they apply least squares learning to learn the demand function. In each period, firms estimate a linear perceived demand function and they play the perceived best response to the previous-period price of the other firms.

This learning rule can lead to three different outcomes: a self-sustaining equilibrium, the Nash equilibrium or an asymmetric learning-equilibrium in which one firm focuses only on the consumers with high transportation costs. The latter equilibria are locally stable therefore the model has coexisting stable equilibria. We analyze the conditions under which the different outcomes are reached.

JEL classification: C62, C63, C72, L13

Keywords: Salop model, least squares learning, heterogeneous consumers

*I am grateful to Jan Tuinstra for his valuable comments that substantially improved the paper.

†email: d.kopanyi@uva.nl

1 Introduction

Firms typically do not know the demand for their product by default, they rather learn demand conditions over time. Learning is especially important when the market is subject to a structural change (e.g. a new product is introduced or a new firm enters the market). A natural way to learn about demand conditions is to gather market data and use it to estimate a demand function. This can be modeled with Least Squares Learning (LSL). LSL consists of two parts: estimation and a decision rule. In a given period, firms use their past observations about prices and demands to estimate the unknown parameters of a so-called perceived demand function. Then, based on the parameter estimates, they choose the action that maximizes their perceived profit. When a new observation arrives, firms update their parameter estimates, leading to (possibly) different prices.

We consider a modified version of the circular road model introduced by Salop (1979). Three firms produce a homogeneous good. Firms are located along a circular road, in equidistant locations. Consumers are uniformly distributed along the circle. When a consumer wants to buy the product, it needs to visit one of the firms. Transportation is costly, consumers face a fixed transportation cost per distance unit. Thus, the total cost of buying the product from a specific firm is given by the sum of the price the firm asks and the transportation cost. Demand is inelastic, each consumer buys exactly one unit of the product. Consumers are assumed to buy the good at the lowest possible total cost. We introduce heterogeneity on the consumer side. There are two types of consumers, one type faces low transportation cost while the other type faces a high one.

Firms do not know the market structure and they use LSL to learn the demand function they face. The true demand function is piecewise linear but firms approximate it with a linear function. Hence the approximation is locally correct but globally incorrect as firm can get a good approximation for only one of the linear parts of the true demand function. In the paper we investigate which outcomes can LSL lead to in this situation. We show that the model has three kinds of equilibria. When firms use all past observations in the estimation, LSL leads to a so-called *self-sustaining equilibrium*. In this equilibrium firms choose the price that maximizes their expected profit subject to their beliefs about demand conditions and their belief is correct at the equilibrium but it is incorrect outside the equilibrium. On the other hand, when only the last few observations are used in the estimation, firms reach either the symmetric *Nash equilibrium* or an *asymmetric learning-equilibrium*. In the asymmetric learning-equilibrium two firms charge a low price and the third one asks a high price. The high-price firm attracts the high-type consumers only whereas the other two firms serve both consumer types. The intuition behind this equilibrium is that the high-price firm does not attract low-type consumers, therefore it underestimates the demand for low prices and it will not perceive it profitable to charge a lower price. In the paper we investigate which conditions determine the outcome of the learning process and we run numerical simulations to evaluate how frequently the different outcomes are reached.

Our paper contributes to the literature on least squares learning. LSL has been widely used in the macroeconomic literature, its convergence properties are well established. Marcet and Sargent (1989) and Evans and Honkapohja (2001) derive conditions under which LSL converges to the rational expectations equilibrium. An important condition is that agents can observe all variables that are relevant for the estimation. Another part of the literature analyzes misspecified LSL, where agents use an incorrect functional form by not taking into account all the relevant variables in the regression. The general result is that agents can reach many different outcomes and the final outcome depends heavily on the initial observations. See for example Gates et al. (1977) and Brousseau and Kirman (1992). In this paper we take an intermediate step between the aforementioned branches of the literature by assuming that agents use all the relevant variables in the regression and that the functional form they use is correctly specified locally but not globally. This situation is of particular interest as the two extreme cases, as we discussed, lead to substantially different outcomes and it is unclear what kind of outcomes could be reached in the intermediate situation. As our results show, in addition to the Nash equilibrium and the self-sustaining equilibrium, a third kind of outcome can be reached in the model: the asymmetric learning-equilibrium. This was not present in previous models.

Least squares learning was applied in market competition in other papers as well. Gates et al. (1977) consider a Cournot oligopoly with differentiated goods where firms do not know the payoff structure. Firms can observe their own production level only and they approximate their average profit with a linear function of their own production. The process can lead to many different outcomes. Kirman (1983) and Brousseau and Kirman (1992) analyze a Bertrand duopoly with differentiated goods in which the true demand function depends linearly on the price set by both firms while the perceived demand function depends on the own price of a firm. Thus, the perceived demand function is misspecified as firms do not take into account the price of the other firm. In this situation firms reach a *self-sustaining* equilibrium. This result generalizes to more firms and to nonlinear demand functions as well, see Anufriev et al. (2013). Firms reach a self-sustaining equilibrium also in our paper when they use all past observations for estimating the perceived demand function. But when they use the last few observations only, they reach either the Nash equilibrium or the asymmetric learning-equilibrium.

The paper is structured as follows. The circular road model is discussed in Section 2. In Section 3 we introduce least squares learning and we derive the equilibria of the model. Simulation results are reported in Section 4. Proofs are presented in the Appendix.

2 The circular road model

The circular road model, one of the baseline models of horizontal product differentiation, was introduced by Salop (1979). In this section we first review a simplified version of the model that is relevant for our analysis and then we introduce heterogeneity on the consumer side.

2.1 Homogeneous consumers

Consider the market for a homogeneous good that is produced by three firms. Firms simultaneously and independently set the price of the good. Firms are located along a circular road, in equal distance from each other. Consumers are uniformly distributed along the circle, their mass (or equivalently the circumference of the circle) is normalized to 1.

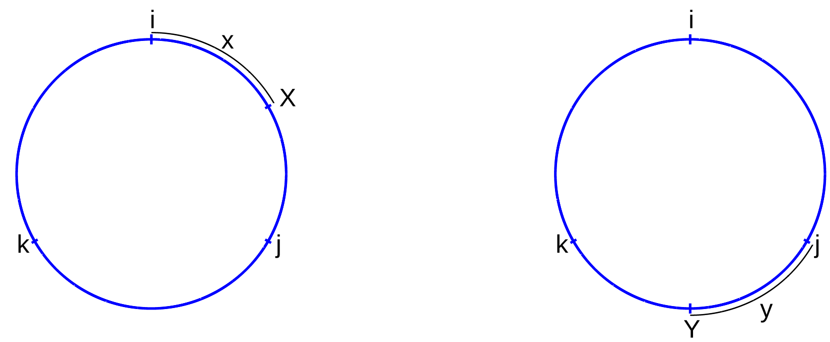
Consumers need to visit one of the firms to purchase the good. They move along the circular road, facing a transportation cost s per distance unit. If the minimal distance between firm i and a given consumer is x, then the consumer's total cost for buying the good from firm i is $p_i + sx$, where p_i is the price charged by firm i and sx is the total transportation cost.[1] Demand is inelastic: each consumer buys exactly one unit of the good. Furthermore, consumers are assumed to buy the good at the lowest possible cost, thus from the firm for which the sum of the price and the total transportation cost is the lowest.

To explain in more detail how demands are determined, we first focus on the competition between firms i and j only and discuss how their demands depend on the prices they charge. Let us consider the consumer that is located on the segment between firms i and j, at distance x from firm i (see Figure 1a). We refer to this consumer as consumer X and to the segment between consumer X and firm i as segment iX.[2] For consumer X, the total cost of buying from firm i is $p_i + sx$ while the total cost of buying from firm j is $p_j + s\left(\frac{1}{3} - x\right)$ since the distance between the two firms is $\frac{1}{3}$. Thus, consumer X buys from firm i rather than from firm j when $p_i + sx < p_j + s\left(\frac{1}{3} - x\right)$. When the two total costs are equal, the consumer is said to be indifferent between the two firms. In this case the location of this *indifferent consumer* can be expressed as $x = \frac{p_3 - p_1}{2s} + \frac{1}{6}$. It is easy to see that when X is the indifferent consumer, the consumers that are located on segment iX prefer firm i to firm j while those on segment Xj prefer firm j.

For determining the demands we distinguish three cases based on the location of the indifferent consumer. First, if the indifferent consumer is located strictly between the two firms, then all the consumers that are closer to firm i than x would buy from firm i rather than for firm j (and vice versa). Thus, firm

[1] It is assumed that firms cannot price discriminate so they cannot charge different prices to consumers from different locations.

[2] We use the following notation for segments of the circle. Segment ab denotes the segment going from point a to point b in clockwise direction. Thus, segments ab and ba form the whole circle together.



(a) Consumer X is located at distance x from i. (b) Consumer Y is located at distance y from j.

Figure 1: Illustrations for the circular road with three firms.

i attracts x consumers while firm j attracts $\frac{1}{3} - x$ consumers. Second, when the indifferent consumer is exactly at the location of firm j, that is for $p_j = p_i + \frac{1}{3}s$, then all the consumers on segment ij prefer firm i to firm j. Let us suppose that the indifferent consumer between firms j and k lies between the two firms, at distance y from firm j (see Figure 1b). We call this consumer Y. In this case the consumers on segment jY are indifferent between firms i and j. To see this note that consumers need to pay the transportation cost for traveling to the location of firm j irrespective of which firm they will choose eventually. And at the location of firm j they are indifferent between the two firms. Indifferent consumers are traditionally assumed to choose one of the firms with equal probability, thus half of of the consumers on segment jY chooses firm i while the other half chooses firm j.[3] Thus, firm j will face a demand of $0.5y$ while firm i attracts $\frac{1}{3} + 0.5y$ consumers. Finally, when the consumer at the location of firm j strictly prefers firm i to firm j, that is when $p_j > p_i + \frac{1}{3}s$, then firm j will not attract any consumer. This small exercise already shows two important features of the model: demand functions are discontinuous and firms can drive each other out of the market.

Taking the above considerations into account, the demand firm i faces can be expressed as a function of prices in the following way. Assume without loss of generality that $p_j < p_k$. Let us first consider the case when firm j does not drive firm k out of the market, i.e. $p_j > p_k - \frac{1}{3}s$. In this case, the demand

[3]We could consider a different share of the indifferent consumers choosing firm i. This would, however, not affect the results of the paper.

function of firm i is given by

$$D_i(p_i,p_j,p_k)=\begin{cases} 1 & \text{if } p_i < p_j-\frac{1}{3}s \\ \frac{11}{12} & \text{if } p_i = p_j-\frac{1}{3}s \\ \frac{1}{2}+\frac{p_j-p_i}{s} & \text{if } p_j-\frac{1}{3}s < p_i < p_k-\frac{1}{3}s \\ \frac{3}{4}+\frac{3p_j-3p_k}{4s} & \text{if } p_i = p_k-\frac{1}{3}s \\ \frac{1}{3}+\frac{p_j+p_k-2p_i}{2s} & \text{if } p_k-\frac{1}{3}s < p_i < p_j+\frac{1}{3}s \\ \frac{1}{12}+\frac{p_k-p_i}{4s} & \text{if } p_i = p_j+\frac{1}{3}s \\ 0 & \text{if } p_j+\frac{1}{3}s < p_i \end{cases} \tag{1}$$

Detailed derivations are presented in the Appendix. We get a different demand function when firm k is driven out of the market by firm j. Analogous calculations as for the previous case yield the following demand function:

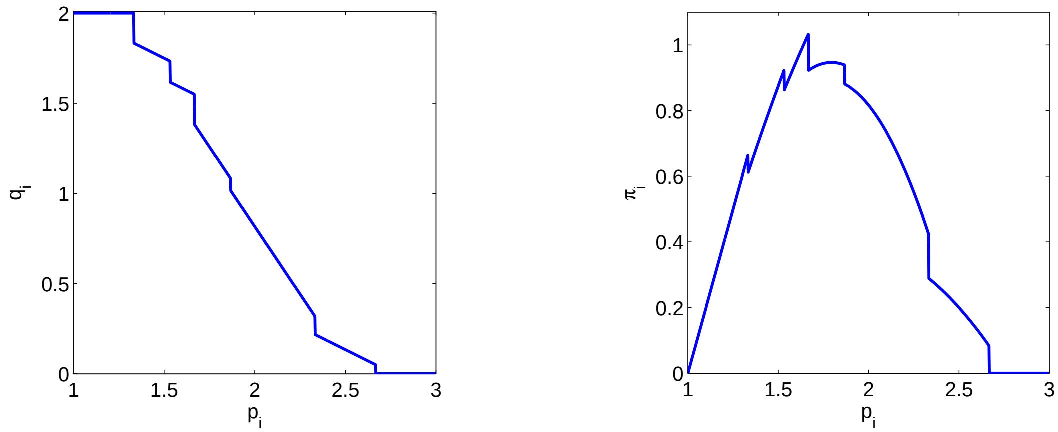
$$D_i(p_i,p_j,p_k)=\begin{cases} 1 & \text{if } p_i < p_j-\frac{1}{3}s \\ \frac{11}{12} & \text{if } p_i = p_j-\frac{1}{3}s \\ \frac{1}{2}+\frac{p_j-p_i}{s} & \text{if } p_j-\frac{1}{3}s < p_i < p_j+\frac{1}{3}s \\ \frac{1}{12} & \text{if } p_i = p_j+\frac{1}{3}s \\ 0 & \text{if } p_j+\frac{1}{3}s < p_i \end{cases}$$

Note that in both cases the demand function is discontinuous and that it consists of piecewise linear parts. This simplified model has a unique symmetric Nash equilibrium, in which each firm charges the price $p = c + \frac{s}{3}$. See Tirole (1988), p. 283 for the proof. Having discussed how the basic model works, let us introduce heterogeneity on the consumer side.

2.2 Heterogeneous consumers

Let us consider the same market structure as before. Suppose that there are two types of consumers. The types differ in the transportation cost they face: low-type consumers face a unit cost of s while high-type consumers pay a unit cost of S, where $s < S$. The amount of consumers of each type is normalized to 1, both types are uniformly distributed along the circular road. Firms cannot distinguish the two types, they cannot price discriminate between different consumers.

Similarly to the case with homogeneous consumers, firms can drive each other out of the market by choosing a sufficiently low price. Moreover, firms can be driven out one part of the market only: it can occur that a firm attracts high-type consumers but not low-type ones. Consider for example the situation $p_j + \frac{1}{3}s < p_i < p_j + \frac{1}{3}S$. In this case the low-type consumer that is located at the position of firm i buys from firm j rather than from firm i. Consequently, firm i does not attract low-type consumers. On the



(a) Demand function of firm i

(b) Profit function of firm i

Figure 2: Demand function of firm i for $p_j = 2$ and $p_k = 2.2$. Other parameters: $s = 1$, $S = 2$ and $c = 1$.

other hand, the high-type consumer at the location of firm i prefers firm i to firm j. Thus, in the given situation, firm j drives firm i out of the market for low-type consumers but not for high-type consumers.

Demand functions are discontinuous and consist of piecewise linear parts again. There are more parts than under homogeneous consumers since firms can be driven out of multiple markets in this case. We do not report here the exact formula for the demand function as it is not important to know it for understanding the results of the paper. The relevant linear parts of the demand function are derived in Section A.2. Figure 2 illustrates the demand and profit functions of firm i when the other two firms do not drive each other out of the market for either consumer type. We can see that the demand function indeed consists of linear parts. We can also see that the profit-maximizing decision of firm i in the given situation is to drive firm k out of the market.

The model with heterogeneous consumers has a unique Nash equilibrium in pure strategies. Proposition 2.1 specifies the equilibrium price. The proof of the proposition is presented in Section A.2.

Proposition 2.1 *The Salop model with three firms and two types of consumers has a unique symmetric Nash equilibrium in pure strategies. In this equilibrium all three firms charge the price* $p_N = \frac{2Ss}{3(S+s)} + c$, *serving both consumer types. The model does not have asymmetric pure-strategy Nash equilibria.*

Note that the Nash equilibrium price is increasing in both s and S. The intuition behind this result is the following. When transportation costs are higher, it is harder for firms to attract consumers that are located farther away from them (or equivalently, it is more costly for consumers to visit firms that are farther away from them). This reduces competition, firms gain more market power and the equilibrium price increases consequently.

It is easy to see that s has a larger impact on the equilibrium price than S does. To understand this result, note the following. When a transportation cost increases, firms have an incentive to increase their price since they get more market power in the given market segment. When a firm increases its price, it will lose some low-type as well as high-type consumers. Since low-type consumers are more mobile, the firm will lose more low-type consumers. Thus, it is more favorable for firms when the transportation cost of low-type consumers increases since this makes low-type consumers less mobile, resulting in a lower decrease in demand after a price increase. Thus, the equilibrium price increases more when s increases.

After analyzing the static model under full information, we now turn to a dynamic model in which firms do not know the market specification and they try to learn the demand condition based on the information they receive about the market.

3 The model with unknown demand

When firms do not know the structure of the game, they need to learn the demand function to find the optimal action. When firms apply least squares learning, they approximate the demand function with a linear function and they estimate the unknown parameters using past observations about price and production levels.

We assume that the only information the firms have about the market is that there are three firms on the market. Thus, they do not know either about the circular-road structure of the market or about facing different consumer types. Firms are competing with each other on the same market over time and they can observe the price charged by their competitors and the corresponding demand for their own good (but not those of their competitors). Thus, firms gather information about the market over time and they can use this information to learn about the demand for their product.

In the following subsection we specify the learning method the firms use and then we discuss the equilibria of the model under learning.

3.1 Least squares learning

Firms approximate the demand for their product with a linear function. To distinguish between the true demand function and the approximation, we will refer to the latter as *perceived* demand function. The perceived demand function of firm i is given by

$$\hat{D}_i(p) = a_i - b_{ii}p_i + b_{ij}p_j + b_{ik}p_k, \tag{2}$$

where a_i denotes the demand intercept and b_{ix} denotes the effect of firm x's price on the demand for firm i's product ($x = i, j, k$). Parameters a_i and b_{ix} are estimated with OLS regression using observations about past prices and own-production levels.

Firms might not want to use all past observations for the estimation therefore we need to make a distinction between a firm's observations and information set. *Observations* of firm i consist of the prices of all three firms and the demand firm i faces for all past periods whereas the *information set* contains only those observations that are used in the regression.[4] The rationale behind not using all observations in the regression is that older observations might carry less information about current demand conditions than more recent ones, especially when there is a structural break in the data. Even though demand conditions are fixed in the model we consider, not using all past observations, as we will see, has important consequences for the properties of LSL.

Let us suppose that firms use the last τ observations for the regression. Then parameter estimates for firm i are given by the standard OLS formula

$$\beta_i = \left(X'_{i,\tau} X_{i,\tau}\right)^{-1} X'_{i,\tau} y_{i,\tau}, \tag{3}$$

where $\beta_i = (a_i, b_{ii}, b_{ij}, b_{ik})'$ is the 4×1 vector of parameter estimates, $X_{i,\tau}$ is the $\tau \times 4$ matrix containing the price observations for the last τ periods (explanatory variables) and $y_{i,\tau}$ is the $\tau \times 1$ vector of the last τ demand observations of firm i (dependent variable).[5]

Given the parameter estimates of the perceived demand function, firm i maximizes its perceived profit $\hat{\pi}_i(p) = (p_i - c)\hat{D}_i(p)$. This gives the following best-response price:

$$p_i^{\mathrm{BR}} = \frac{a_i + b_{ij}p_j + b_{ik}p_k}{2b_{ii}} + \frac{c}{2}. \tag{4}$$

Let us now discuss timing. At the end of period t firms have observations about all t periods. Parameter estimates are obtained by (3). In order to stress that parameter estimates are changing over time, we will denote the parameter estimates at the end of period t as $a_{i,t}$, $b_{ii,t}$, $b_{ij,t}$ and $b_{ik,t}$. Since firms are determining their price simultaneously, they can play the best response only against the *expectations* they have about the prices of other firms. Thus, we have to replace p_j with $p^e_{j,t+1}$ and p_k with $p^e_{k,t+1}$ in (4), stressing again the dependence on time. We assume that firms form naive expectations, meaning that they expect other firms to charge the same price as in the previous period: $p^e_{j,t+1} = p_{j,t}$ and $p^e_{k,t+1} = p_{k,t}$. This leads to the following pricing formula for period $t + 1$:

[4]Note that we use the term information set in its econometrical sense and not in its game theoretical sense.

[5]Similar formulas apply when firms use all past observations. The only difference is that X and y contain the prices and the corresponding demand for all past periods.

$$p_{i,t+1} = \frac{a_{i,t} + b_{ij,t}p_{j,t} + b_{ik,t}p_{k,t}}{2b_{ii,t}} + \frac{c}{2}. \tag{5}$$

Note that profit maximization requires $b_{ii,t} > 0$, that is the perceived own-price effect must be negative. Since the perceived demand functions the firms use are not correctly specified, it might occur that the parameter estimate for $b_{ii,t}$ becomes negative. In this case, (5) does not give the perceived profit-maximizing price. Also note that when (5) yields a price that is lower than the marginal cost, the firm makes a negative profit (provided that it faces a positive demand). Thus, (5) is not applicable in this case either. In order to overcome these problems with LSL, we augment the method with the following rule.

Random price rule: When $b_{ii,t} \leq 0$ or (5) yields $p_{i,t+1} < c$, then firm i chooses a price randomly from the uniform distribution on a predefined interval I.

Interval I is specified in Section 4. We need to impose additional rules to overcome some numerical issues that may occur when firms do not use all observations in the regression. When prices start to settle down at a given value, there is not enough dispersion in the observations and matrix $X_{i,\tau}$ is close to being singular, resulting in imprecise parameter estimates. This can lead to extremely high prices for some periods. Since it should be clear for firms that large unexpected price changes result from the aforementioned issue, it is reasonable to assume that firms do not follow pricing rule (5) in this case, they rather keep their price unchanged. This leads to the following rule.

No jump rule: If (5) yields a price that is K times higher than the price of firm i in the previous period, then the firm will keep its price unchanged and charge the same price as in the previous period.[6]

When there is not enough dispersion in the price observations, matrix $X_{i,\tau}$ can become singular, making the estimation impossible. We assume that firms keep their price unchanged when estimation is not possible.

Impossible estimation rule: When (3) is not applicable due to the singularity of $X_{i,\tau}$, then firm i will keep its price unchanged and charge the same price as in the previous period.

We will elucidate the effect of these rules in the Discussion in Section 5. Let us now turn to the steady states of the process.

[6] Alternatively, we could use impose an upper bound on price changes as Weddepohl (1995). In that case firms choose the highest possible price when (5) would result in a too large price jump. Since large price jumps are associated with imprecise parameter estimates in the model we consider, it makes more sense not to change the price at all.

3.2 Equilibria under least squares learning

The system is in a steady state when neither the parameter estimates of the perceived demand functions nor the prices change. It must hold for any steady state that the true and perceived demand functions coincide for each firm at the given price vector p^*, that is $D_i(p^*) = \hat{D}_i(p^*)$ for $i = 1, 2, 3$. To see this, note the following. When $D_i(p^*) = \hat{D}_i(p^*)$, the perceived demand function perfectly approximates the true demand function for the given price vector as the corresponding estimation error is 0. Since the parameter estimates of the perceived demand function are obtained by minimizing the sum of squared errors, this implies that the parameter estimates do not change in this case.

The same condition characterizes the self-sustaining equilibria in Brousseau and Kirman (1992). Thus, the steady states of the model with least squares learning are self-sustaining equilibria: firms play the best response subject to their beliefs about demand conditions (i.e. the perceived demand functions) and about the actions of other firms and their beliefs are correct at the equilibrium price vector. Self-sustaining equilibria can be formally defined as follows.

Definition 3.1 *Price vector $p^* = (p_i^*, p_j^*, p_k^*)$ and the parameter estimates $\{a_i, b_{ii}, b_{ij}, b_{ik}\}$ $(i, j, k = 1, 2, 3; i \neq j \neq k)$ constitute a self-sustaining equilibrium if the following conditions hold for each firm i :*

$$p_i^* = \frac{a_i + b_{ij}p_j^* + b_{ik}p_k^*}{2b_{ii}} + \frac{c}{2}, \tag{6}$$

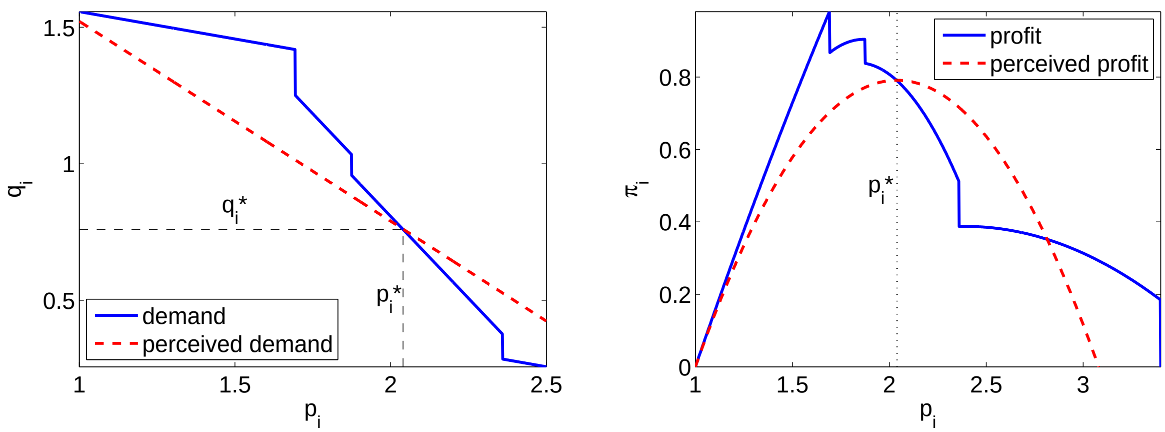
$$\hat{D}_i(p^*) = D_i(p^*). \tag{7}$$

Condition (6) shows that firms play the best response subject to their beliefs and (7) means that beliefs are confirmed in equilibrium as the actual demand is the same as the demand the firm expects to get.

It can be seen from the definition that there are many different self-sustaining equilibria, thus the model has multiple steady states. Proposition 3.2 specifies which price vectors can form a self-sustaining equilibrium. The proof is presented in Section A.3.

Proposition 3.2 *For any price vector $p = (p_i, p_j, p_k)$ satisfying the conditions $p_i > c$ and $D_i(p) > 0$ for $i = 1, 2, 3$, there exist parameter estimates $\{a_i, b_{ii}, b_{ij}, b_{ik}\}$ $(i, j, k = 1, 2, 3; i \neq j \neq k)$ such that the model is in a self-sustaining equilibrium.*

Thus, prices exceed the marginal cost and each firm faces a positive demand in a self-sustaining equilibrium. Note that the condition $D_i(p) > 0$ implies that none of the firms can be driven out of the market for *both* types of consumers. But it is not required that each firm should attract both consumer types. In the above result, we did not take into account that $\{a_i, b_{ii}, b_{ij}, b_{ik}\}$ are not freely chosen but they result from estimation. Therefore not all the price vectors that satisfy the conditions of Proposition 3.2



(a) Demand functions in a self-sustaining equilibrium (b) Profit functions in a self-sustaining equilibrium

Figure 3: Demand and profit functions of firm i in a self-sustaining equilibrium. Parameters: $s = 1$, $S = 5$ and $c = 1$. Equilibrium prices: $p_i^* = 2.0394$, $p_j^* = 2.0267$ and $p_k^* = 2.2083$.

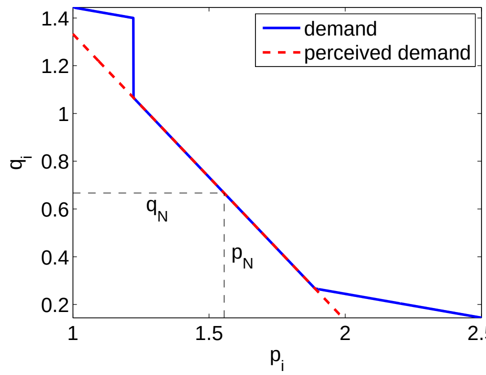
can necessarily be *reached*, despite the fact that we can find parameter values for which they constitute a self-sustaining equilibrium.

Since perceived demand functions are linear while the true demand functions are piecewise linear, firms cannot fully learn the true demand conditions: they can correctly learn the parameters of at most one linear part. Note that condition (7) is required to hold at the equilibrium point only, thus firms need not learn in general any linear part correctly. Panel (a) of Figure 3 illustrates the true and the perceived demand functions of a firm in a typical self-sustaining equilibrium. The two functions cross each other in a single point thus the firm does not learn any linear part of the true demand function correctly. Panel (b) depicts the true and the perceived profit functions. The figure shows that in the SSE firm i maximizes its perceived profit but the price it chooses does yield the true profit maximum.

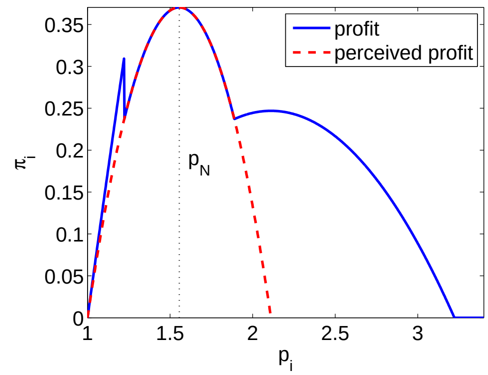
Even though it is not the case typically, there are self-sustaining equilibria in which firms correctly learn the part of the true demand function on which they operate. Proposition (3.3) specifies these equilibria. The proof is presented in Section A.4.

Proposition 3.3 *The model with least squares learning has two self-sustaining equilibria in which firms correctly learn the linear part of the true demand function on which they operate. The Nash equilibrium of the game is always such an equilibrium of the learning process. When* $\frac{S}{s} \geq \Sigma_1 = \frac{7+\sqrt{89}}{4} \approx 4.1085$, *there also exists an asymmetric equilibrium in which two firms charge* $p_L = \frac{11Ss}{12S+15s} + c$ *and the third firm chooses* $p_H = \frac{2S^2+8Ss}{12S+15s} + c$. *We refer to this equilibrium as asymmetric learning-equilibrium (ALE).*

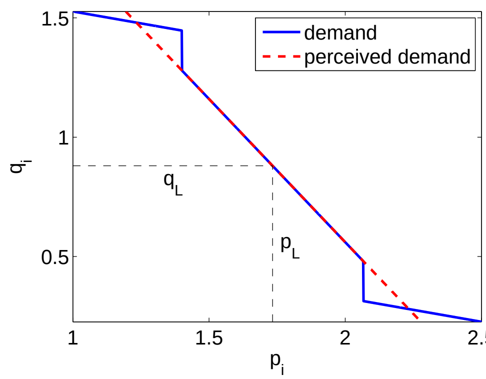
Figure 4 illustrates the demand and profit functions in the Nash equilibrium and in the asymmetric



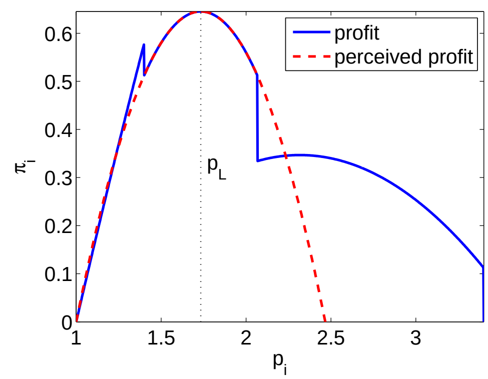
(a) Demand functions (Nash equilibrium)



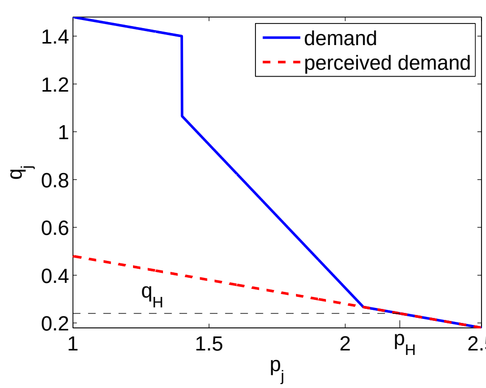
(b) Profit functions (Nash equilibrium)



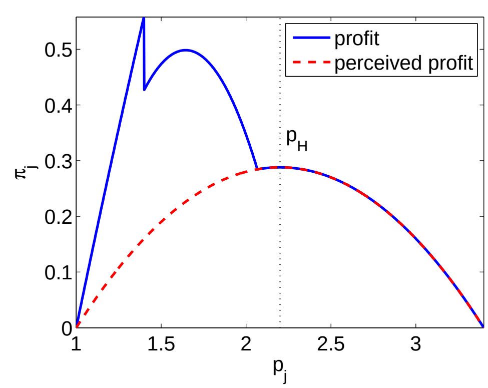
(c) Demand functions of low-price firms (ALE)



(d) Profit functions of low-price firms (ALE)



(e) Demand functions of the high-price firm (ALE)



(f) Profit functions of the high-price firm (ALE)

Figure 4: Demand and profit functions in the Nash equilibrium and in the asymmetric learning-equilibrium. Parameters: $s = 1$, $S = 5$ and $c = 1$.

learning-equilibrium. Panels (a), (c) and (e) confirm that in both equilibria firms correctly approximate the linear part of the true demand function on which they operate. Panel (b) shows that the true profit maximum coincides with the maximum of the perceived profit function of firms in the Nash equilibrium. The same holds for the low-price firms in the ALE (see panel (d)). Note, however, that the perceived profit maximum does not correspond to the true profit maximum for the high-price firm (panel (f)). This is why the ALE is not a Nash equilibrium of the game under known demand. As panel (e) shows, the high-price firm underestimates the demand for lower prices and thus it does not perceive it more profitable to charge a lower price, even though it would yield a higher profit.

We show in Section A.5 that low-price firms earn a higher profit than the high-price firm in the ALE only when $\frac{S}{s}$ is sufficiently low. When $\frac{S}{s} > \frac{89+11\sqrt{73}}{8} \approx 22.87$, the high-price firm earns a higher profit. In this case the high-price firm still underestimates the demand for low prices but the perceived profit maximum coincides with the true profit maximum. On the other hand, low-price firms perceive a relatively high slope and they underestimate the demand for high prices. Their perceived profit maximum does not coincide with the true profit maximum as it would be more profitable to charge a higher price.

As we have seen, the model with least squares learning has three types of equilibria: a general self-sustaining equilibrium, the Nash equilibrium and the asymmetric learning-equilibrium. Next we will investigate which equilibria can be reached and which factors determine which of the equilibria is reached. It turns out that a special property of the information set plays a crucial role in this. Before defining this property, note that different price vectors may correspond to different demand conditions. For example, firm i may serve both types of consumers for one price vector whereas it might serve high-type consumers only for another price vector. These price observations carry information about different structural parameters as they lie on different linear parts of the true demand function. We call price vectors in the information set of firms *aligned* when each firm serves the same consumer type(s) for each price vector. We distinguish two kinds of aligned price vectors. When all three firms serve both consumer types, we speak about *symmetrically aligned* prices. When two of the firms serve both consumer types while the third one attracts high-type consumers only, we speak about *asymmetrically aligned* prices.[7] We define these concepts formally as follows.

Definition 3.4 *A set of price vectors $P \subseteq \mathbb{R}^3$ is called* symmetrically aligned *when all three firms attract both types of consumers for all $p \in P$:*

$$|p_i - p_j| < \frac{s}{3} \quad \forall i, j = 1, 2, 3.$$

A set of price vectors $P \subseteq \mathbb{R}^3$ is called asymmetrically aligned *when firms i and j attract both types of*

[7]Note that prices could be aligned in other ways as well. For example, we could consider the case when exactly one firm attracts both types of consumers while the other two firms attract high-type consumers only. We do not consider other possibilities because they are not relevant for the equilibria of the learning process, as we have seen.

consumers while firm k attracts only the high-type consumers for all $p \in P$:

$$|p_i - p_j| < \frac{s}{3}$$
$$\min\{p_i, p_j\} + \frac{s}{3} < p_k < \min\{p_i, p_j\} + \frac{S}{3}.$$

A set of price vectors $P \subseteq \mathbb{R}^3$ is called not aligned *when it is neither symmetrically, nor asymmetrically aligned.*

The condition $|p_i - p_j| < \frac{s}{3}$ ensures firms i and j do not drive each other out of the market for either consumer type. The condition $\min\{p_i, p_j\} + \frac{s}{3} < p_k < \min\{p_i, p_j\} + \frac{S}{3}$ means that firm k is driven out of the market for low-type consumers but nor for the high-type ones.

When prices are aligned, then the corresponding demand observations are consistent in the sense that they lie on the same linear part of the demand function. That is, observations carry information about the same linear demand parameters and consequently firms correctly learn the parameters that characterize the linear part of the true demand function on which they operate.

Since firms play the best response to the prices of the other firms, subject to their perceived demand function, it is important to analyze the conditions under which a set of aligned price observations remains aligned after updating the set with the best-response prices. Lemma 3.5 summarizes these conditions. The proof can be found in Section A.6.

Lemma 3.5 *When price observations are symmetrically aligned, then updating the information set with the best-response prices always results in symmetrically aligned price observations again.*
When price observations are asymmetrically aligned, there are three possibilities.

1. *For $\frac{S}{s} < \Sigma_1$ price observations will not be asymmetrically aligned after updating the information set with the best-response prices sufficiently many times.*

2. *For $\frac{S}{s} \in [\Sigma_1, \Sigma_2)$ with $\Sigma_2 = 2 + \sqrt{6} \approx 4.4495$, the updated price observations will be asymmetrically aligned if the following condition holds for the most recent price observation p :*

$$\frac{s}{3}\left[2\left(\frac{S}{s}\right)^2 - 7\frac{S}{s} - 4\right] + \left(1 + \frac{S}{s}\right)|p_i - p_j| + \min\{p_i, p_j\} \geq p_k.$$

3. *For $\frac{S}{s} \geq \Sigma_2$, price observations always remain asymmetrically aligned after updating the information set with the best-response prices.*

According to this lemma, when the information set is symmetrically aligned, then it always remains symmetrically aligned. Thus, firms will learn the true parameters of the corresponding linear part. As

Proposition 3.3 shows, the only equilibrium that firms may reach in this situation is the Nash equilibrium. Concerning asymmetrically aligned observations, Lemma 3.5 says that when $\frac{S}{s}$ is not high enough, the information set will not be asymmetrically aligned eventually even if firms start with an asymmetrically aligned information set. Thus, if we keep on updating an asymmetrically aligned information set with the best response prices, the information set will not be asymmetrically aligned from some point on. So in this case the possible steady states of the model are a general SSE and the Nash equilibrium. For intermediate values of $\frac{S}{s}$, an extra condition is needed for ensuring that the updated information set remains asymmetrically aligned. Thus, all three steady states may exist for these values of $\frac{S}{s}$. On the other hand, an asymmetrically aligned information set always remains asymmetrically aligned by updating it with the best response prices when $\frac{S}{s}$ is high enough. Thus, the only equilibrium in this case is the asymmetric learning-equilibrium. When price observations are not aligned, then firms cannot learn the true parameters of the linear part on which they operate, consequently the only kind of steady state in the given situation is a general self-sustaining equilibrium.

Note that these results concern existence only, under specific conditions. We have not yet analyzed the stability of these equilibria. Proposition 3.6 summarizes the dynamical properties of the steady states. The proof is presented in Section A.7.

Proposition 3.6 *Both the Nash equilibrium and the asymmetric learning-equilibrium are locally stable equilibria of the model with least squares learning.*

According to the proposition, firms will reach the Nash equilibrium when initial prices are close to the Nash equilibrium price. A similar result hold for the asymmetric learning-equilibrium. Combining these considerations with Lemma 3.5, we can conclude that the model has coexisting locally stable steady states when $\frac{S}{s}$ is sufficiently high. Note that Proposition 3.6 does not mention the stability of the general self-sustaining equilibria. Brousseau and Kirman (1992) shows that firms do not converge to a self-sustaining equilibrium in general. To process slows down only because the weight of a new observation decreases when firms use all observations in the estimation.

Taking into account the above results, we summarize the long-run outcome of the model in Table 1. When firms use all observations in the estimation, then all three equilibria can occur. More specifically, when initial observations are symmetrically aligned, firms converge to the Nash equilibrium. When initial observations are asymmetrically aligned, firms reach the asymmetric learning-equilibrium when $\frac{S}{s}$ is sufficiently high. When initial observations are not aligned or if they are asymmetrically aligned but $\frac{S}{s}$ is not high enough, then firms move towards a self-sustaining equilibrium.

When only the last τ observations are used in the regression, then firms reach the Nash equilibrium

initial observations		information set: all observations	information set: last τ observations
symm. aligned		Nash	Nash
asymm. aligned	$\frac{S}{s} \leq \Sigma_1$	SSE	Nash / SSE
	$\Sigma_1 < \frac{S}{s} < \Sigma_2$	ALE / SSE	ALE / Nash / SSE
	$\Sigma_2 \leq \frac{S}{s}$	ALE	ALE
not aligned		SSE	ALE / Nash / SSE

Table 1: The possible outcomes of the model with least squares learning for different types of initial observations and different number of observations in the information set.

for symmetrically aligned initial prices. When initial prices are asymmetrically aligned and $\frac{S}{s}$ is high enough, then firms converge to the asymmetric learning-equilibrium. For the other cases we cannot predict which equilibrium is reached. Our conjecture is that the information set will eventually become either symmetrically or asymmetrically aligned, so firms reach either the Nash equilibrium or the asymmetric learning-equilibrium in the end. This conjecture is based on the local stability of both the Nash equilibrium and the ALE. Since both equilibria are locally stable, we expect that observations will not jump between the different linear parts of the demand function. In this case the proportion of either the symmetrically or the asymmetrically aligned observations will increase in the information set and the information set becomes either symmetrically or asymmetrically aligned eventually. If this conjecture does not hold, then the process does not converge at all as observations keep on jumping between the different linear parts of the true demand function.

In the next section we will run computer simulations to check whether our conjecture is correct. Also we will investigate how often the different outcomes are reached.

4 Simulation results

We run simulations with 1000 different initializations. Each initialization runs until the maximal price change is smaller than the threshold value of 10^{-5}, i.e. $\max_i |p_{i,t} - p_{i,t-1}| \leq 10^{-5}$, or until period 1000 is reached. We fix the market parameters $c = 1$ and $s = 1$, and we vary the value of S. Based on the theoretical results, we consider 6 different values for S, for which the equilibria have different dynamical properties. Table 2 summarizes the values of S we consider and the corresponding prices in the Nash equilibrium and in the asymmetric learning-equilibrium. For $S = 2$ and $S = 4$ the ALE does not exist. For

	S					
	2	4	4.2	4.35	4.5	10
p_N	1.4444	1.5333	1.5385	1.5421	1.5455	1.6061
p_L	-	-	1.7064	1.7121	1.7174	1.8148
p_H	-	-	2.0532	2.0810	2.1087	3.0741

Table 2: The Nash equilibrium price and prices in the asymmetric learning-equilibrium for different values of S. Other parameters: $s = 1$ and $c = 1$.

$S = 4.2$ and $S = 4.35$ the ALE exists but an asymmetrically aligned information set not always remains asymmetrically aligned after updating it with the best response prices. For the last two values of S an asymmetrically aligned information set always remains asymmetrically aligned.

Concerning the parameters in the learning method, we fix $K = 5$ in the random price rule. Whenever firms need to pick a price randomly, they use the interval $I = [c, p_H + c]$. We believe that these choices are appropriate since all the prices that are relevant for the long-run outcome of the model are between c and p_H and they are always smaller than 3.1 for the model parameters we use.[8] We consider different values for τ (the number observations used in the estimation). Since there are 4 parameters to be estimated, we need at least 4 observations in the information set. We will investigate how the size of the information set affects the outcome of the model.

Since we conjectured to observe substantially different outcomes when firms use all observations compared to the case when they use the last τ observations only, we discuss the simulation results for these cases in separate sections.

4.1 All observations

First we investigate the outcome of the model when firms use all observations for estimating the perceived demand function. In this case, firms can move towards a general SSE, they case reach the Nash equilibrium or the ALE (provided it exists). As we have shown, the latter two equilibria are reached only when the initial observations are aligned. Since initial prices are drawn randomly, information sets are typically not aligned, therefore a general SSE is reached, in which firm do not approximate correctly the linear part on which they operate.

Figure 5 illustrates the time series of prices in one simulation for $S = 2$. The figure shows that prices settle down fast and that firms charge different prices. The given simulation stopped in period 1000,

[8] Also note that our theoretical results do not depend on the rules that augment least squares learning.

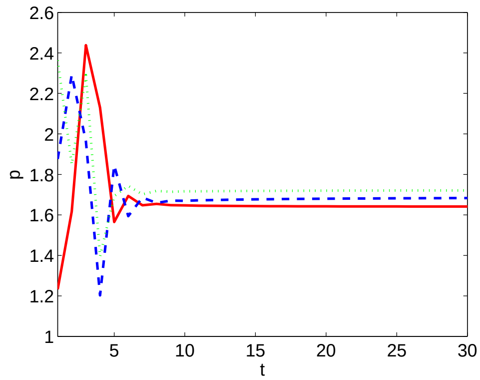


Figure 5: Time series of prices for $S = 2$. Other parameters: $s = 1$ and $c = 1$.

the maximal difference between the true and perceived demands at the final price vector is $0.3 \cdot 10^{-3}$, confirming that firms move towards a self-sustaining equilibrium.

As Proposition 3.2 shows, many price vectors can be part of an SSE. Therefore it is worthwhile to investigate the distribution of final prices. Figure 6 shows histograms of the final prices over the 1000 different initializations, for different values of S. The histograms show that there is substantial price dispersion and that neither the Nash-equilibrium nor the ALE provides a benchmark outcome when all observations are used. As S increases, the distribution seems to become flatter.

	S = 2	S = 4	S = 4.2	S = 4.35	S = 4.5	S = 10
mean	1.6060	1.8647	1.8836	1.9045	1.9269	2.6018
median	1.5781	1.8307	1.8582	1.8787	1.9004	2.5203
stdev	0.1625	0.2372	0.2315	0.2362	0.2451	0.5694

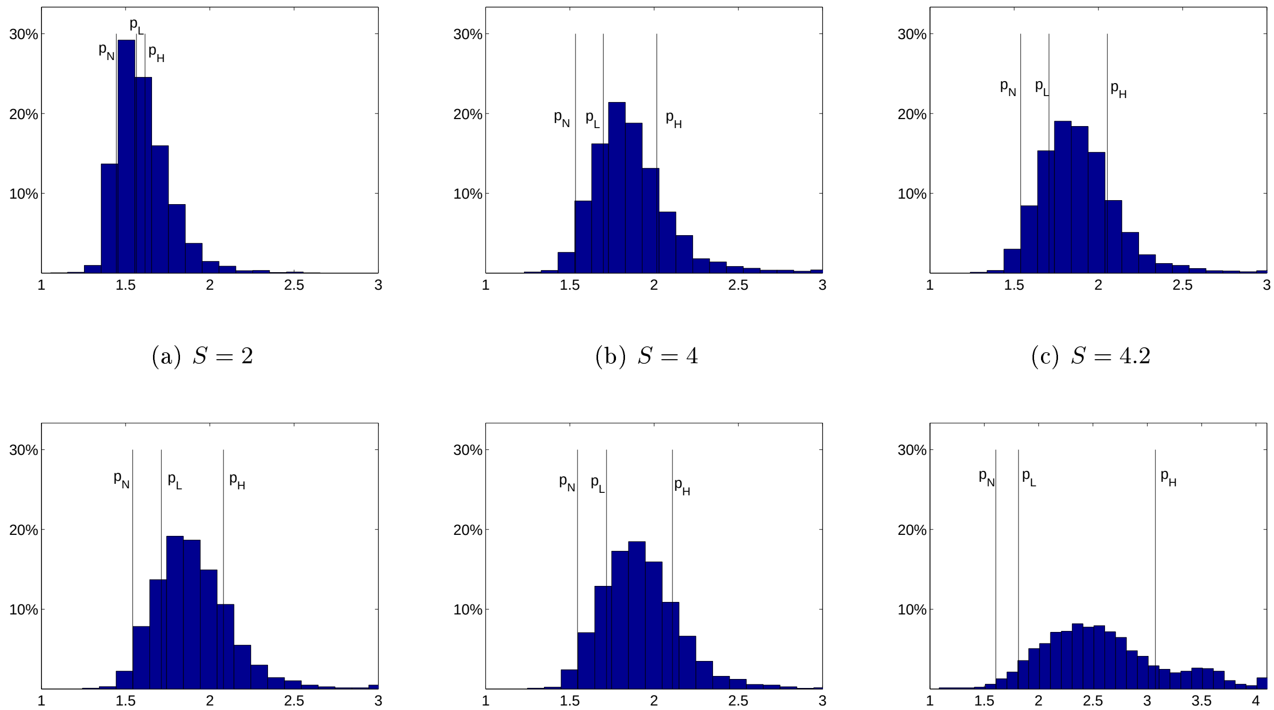
Table 3: Descriptive statistics of final prices for different values of S. Other parameters: $s = 1$ and $c = 1$.

Table 3 shows descriptive statistics of the final prices for different values of S. As S increases, both the average and the median prices increase.[9] There is not much difference in the standard deviations.

In order to measure how close firms get to a self-sustaining equilibrium, we calculate the absolute difference between the actual and perceived demands at the final price vectors. The difference is 0 in an SSE. Table 4 shows descriptive statistics of these differences for different values of S. The first three rows show the mean, minimal and maximal absolute difference whereas the last three rows report the number of initializations for which the difference is smaller than 10^{-2}, 10^{-3} and 10^{-4} for all three firms.

We can conclude from the table that differences are rather small in all cases. In almost all cases,

[9]Note that the upper bound of the interval for initial prices also increases.



(a) $S = 2$ (b) $S = 4$ (c) $S = 4.2$

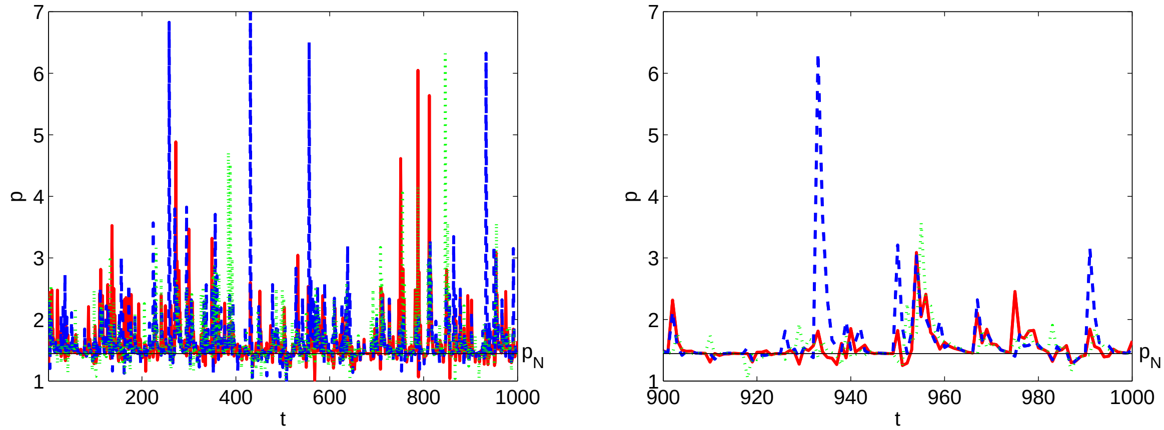
(d) $S = 4.35$ (e) $S = 4.5$ (f) $S = 10$

Figure 6: Histogram of final prices for different values of S. Other parameters: $s = 1$ and $c = 1$.

	S = 2	S = 4	S = 4.2	S = 4.35	S = 4.5	S = 10
mean	$5.9 \cdot 10^{-4}$	$8.9 \cdot 10^{-4}$	$8.0 \cdot 10^{-4}$	$8.2 \cdot 10^{-4}$	$7.9 \cdot 10^{-4}$	$1.0 \cdot 10^{-3}$
min	$2.6 \cdot 10^{-8}$	$2.2 \cdot 10^{-16}$	$3.9 \cdot 10^{-16}$	$1.1 \cdot 10^{-16}$	0	0
max	$4.7 \cdot 10^{-2}$	$1.4 \cdot 10^{-1}$	$8.2 \cdot 10^{-2}$	$9.0 \cdot 10^{-2}$	$7.4 \cdot 10^{-2}$	$9.4 \cdot 10^{-2}$
diff $\leq 10^{-2}$	978	971	973	974	977	962
diff $\leq 10^{-3}$	906	895	887	889	887	869
diff $\leq 10^{-4}$	371	442	454	455	465	615

Table 4: Descriptive statistics of the absolute difference between the true and perceived demands at final prices, for different values of S. Other parameters: $s = 1$ and $c = 1$.

maximal difference is at most 10^{-2}. This confirms that firms get close to a self-sustaining equilibrium when all observations are used in the regression.



(a) all periods

(b) last periods

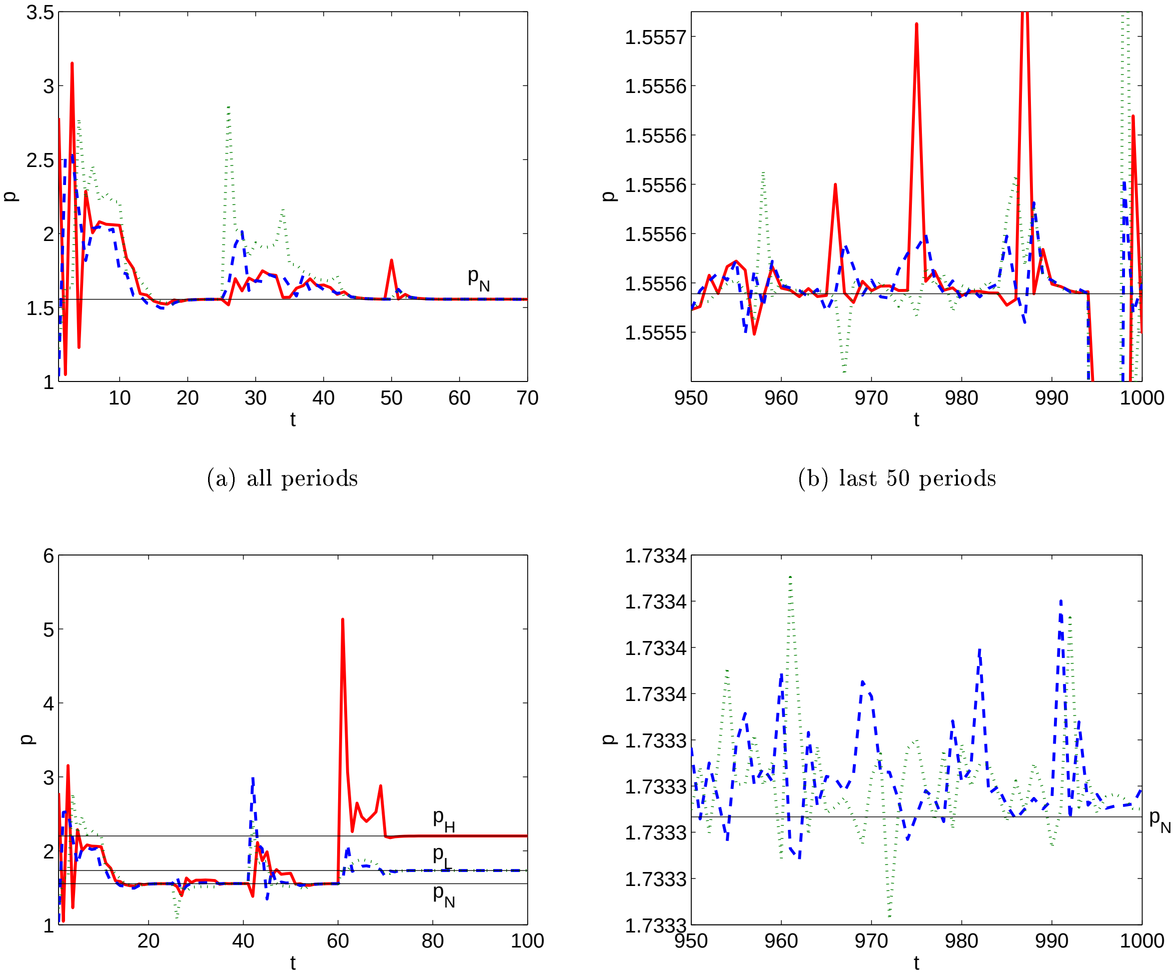
Figure 7: Time series of prices for $\tau = 4$. Parameters: $s = 1$, $S = 2$ and $c = 1$.

4.2 Simulations with last τ observations

Next we turn to the case when information sets contain the last τ observations only. Our conjecture was that information sets become either symmetrically or asymmetrically aligned in this case and firms converge either to the Nash equilibrium or to the asymmetric learning-equilibrium. From Proposition 3.3 we know that the ALE does not exist for $S = 2$ and $S = 4$, thus the Nash equilibrium should always be reached for these values of S.

As we discussed, at least 4 observations are needed for the regression. It turns out that the process does not converge typically when firms use exactly $\tau = 4$ observations. Figure 7 illustrates the time series of prices in a simulation with $\tau = 4$. The figure shows that prices do not settle down at the Nash equilibrium price. They start converging towards the Nash equilibrium (already indicating that the Nash equilibrium is locally stable) but at some point they diverge away from it. The reason behind this is that when there is not enough dispersion in the observations, parameter estimates become imprecise and one of the firms will charge a relatively large price. When firms use 4 observations only, then the weight of a single observations is apparently large enough and the outlier observation can drive the prices far from the equilibrium.

In contrast, when firms use more observations, the weight of a single observation decreases, thus a single outlier does not drive away prices from the equilibrium that much. We indeed find convergence when the size of the information set increases. Figure 8 shows typical time series for $\tau = 8$. Panel (a) shows an example where prices converge to the Nash equilibrium. Prices seem to settled down at the Nash-equilibrium price after some initial oscillations. Panel (b) shows the same time series but for the last 50 periods of the simulation. It turns out that we do not find exact convergence but small oscillations



(a) all periods

(b) last 50 periods

(c) all periods

(d) last 50 periods

Figure 8: Time series of prices for $\tau = 8$. Parameters: $s = 1$, $S = 5$ and $c = 1$.

around the Nash equilibrium. This is caused by the same numerical problem as we have for $\tau = 4$: parameter estimates become imprecise when there is not enough variation in the observations.[10]

In order to investigate whether firms always converge either to the Nash equilibrium or to the ALE, we run 1000 simulations for each (S, τ) combination that we consider and we calculate which proportion of the final price vectors lies in a small neighborhood of the Nash equilibrium and the ALE respectively. Table 5 summarizes the results. The table shows 4 numbers for each (S, τ) combination. The upper values refer to the Nash equilibrium whereas the lower ones to the ALE. The numbers that are not in brackets correspond to the 0.001-neighborhood of the given equilibrium while the numbers in brackets show the

[10]To confirm that these oscillations are due to numerical problems we run the same simulations with using the true demand coefficients when observations in the information set are aligned. In this case we always find exact convergence to one of the equilibria. These simulations serve as a theoretical benchmark only since the true coefficients are not available for firms.

		τ 6	τ 8	τ 10	τ 20
S	**2**	96.9% (90.4%)	99.0% (97.1%)	100% (99.8%)	99.9% (99.9%)
		0.0% (0.0%)	0.0% (0.0%)	0.0% (0.0%)	0.0% (0.0%)
	4	95.4% (88.5%)	99.5% (98.1%)	99.9% (99.3%)	100% (100%)
		0.0% (0.0%)	0.0% (0.0%)	0.0% (0.0%)	0.0% (0.0%)
	4.2	60.1% (55.6%)	79.6% (78.6%)	83.0% (82.5%)	88.2% (88.2%)
		37.7% (33.6%)	20.2% (19.6%)	16.8% (16.8%)	11.7% (11.7%)
	4.35	52.9% (47.6%)	71.2% (70.2%)	77.6% (77.2%)	84.3% (84.3%)
		45.1% (41.8%)	28.4% (28.1%)	22.3% (21.8%)	15.3% (15.3%)
	4.5	48.5% (46.1%)	67.0% (65.9%)	72.5% (72.0%)	81.4% (81.4%)
		49.9% (46.5%)	33.0% (32.1%)	27.3% (26.8%)	18.5% (18.5%)
	10	7.7% (7.2%)	13.5% (13.1%)	15.4% (15.4%)	20.5% (20.5%)
		91.3% (79.2%)	86.2% (81.6%)	84.5% (82.1%)	78.5% (78.5%)

Table 5: Proportion of outcomes in the 0.001 and the 0.0001-neighborhoods (in brackets) of the Nash equilibrium (upper numbers) and the asymmetric learning-equilibrium (lower numbers) over 1000 simulations, for different values of S and τ. Other parameters: $s = 1$ and $c = 1$.

proportion of final price vectors in the 0.0001-neighborhoods.[11]

The table confirms that firms almost always converge either to the Nash equilibrium or to the asymmetric learning-equilibrium. As we have discussed before, there is not exact convergence in the model, that is why not all the outcomes lie in the small neighborhood of the equilibria. Note that as τ increases, a higher proportion of the final price vectors lies in the neighborhoods that we consider. This is due to the fact that when firms use more observations in the estimation, a single outlier does not drive prices away from the equilibrium that much. The table also shows that the Nash equilibrium is reached more often as τ increases. On the other hand, the ALE becomes more focal as S increases.

To exclude the effect of the numerical problem, we run the same simulations with firms using the true parameters of the given part of the demand function when the information set is aligned. Table 6 summarizes the results. Since now there is exact convergence, we show the values that correspond to the 0.0001-neighborhoods only. Note that the numbers for each (S, τ) combination add up to 100%, confirming

[11] We say that a vector (x_1, x_2, x_3) lies in the ε neighborhood of another vector (y_1, y_2, y_3) if their Euclidean distance is smaller than or equal to ε : $\sqrt{\sum_{i=1}^{3} (x_i - y_i)^2} \leq \varepsilon$.

		τ			
		6	**8**	**10**	**20**
S	**2**	100% 0%	100% 0%	100% 0%	100% 0%
	4	100% 0%	100% 0%	100% 0%	100% 0%
	4.2	94.4% 5.6%	97.0% 3.0%	96.9% 3.1%	97.1% 2.9%
	4.35	92.1% 7.9%	95.3% 4.7%	96.0% 4.0%	97.0% 3.0%
	4.5	90.9% 9.1%	93.2% 6.8%	94.1% 5.9%	95.5% 4.5%
	10	50.9% 49.1%	56.4% 43.6%	56.3% 43.7%	58.5% 40.7%

Table 6: Percentage of outcomes in the 0.0001-neighborhood of the Nash equilibrium and the asymmetric learning-equilibrium when the true coefficients are used for aligned price observations. Number of simulations: 1000, other parameters: $s = 1$ and $c = 1$.

that firms always reach either the Nash equilibrium or the ALE. We again find that the Nash equilibrium is reached more often as τ increases and that firms converge to the ALE more often as S increases. Note, however, that there are substantial differences in the numbers compared to Table 5: the Nash equilibrium is reached much more often than before. This shows that the numerical problem that occurs when there is not enough variation in the observations, has an important effect on which equilibrium will eventually be reached. The results suggest that the Nash equilibrium is less stable than the ALE in the sense that the numerical problem can drive prices from the Nash equilibrium to the ALE more often than the other way around. In fact, panel (c) of Figure 8 shows a situation where prices settle down around the Nash equilibrium initially but after a high price realization the firms converge to the ALE. This finding can explain why firms converge more frequently to the Nash equilibrium when τ increases. As we discussed, when the size of the information set increases, a single observation has a smaller effect on the parameter estimates. Therefore a high price realization that may occur after prices have settled down around the Nash equilibrium, has a smaller impact on the parameter estimates, therefore the best-response prices stay in the basin of attraction of the Nash equilibrium instead of reaching the basin of attraction of the ALE.

We checked the robustness of our results with respect to the number of periods in the simulations and the number of different initializations. We focused on the $\tau = 8$ case and we ran simulations with the previously used values of S. We ran one set of simulations with 10000 periods instead of 1000, while we considered 10000 different initialization in the other set of simulations instead of 1000. The outcome of these simulations is shown in Table 7 in Section A.8. The results are in line with the previous ones, the Nash equilibrium and the ALE are reached in about the same proportion of the cases as before. Therefore we conclude that our results are robust.

5 Discussion

We have shown that least squares learning can result in a sub-optimal outcome for some firms even when the perceived demand function is correctly specified locally and firms can observe the actions of each other.

We considered the Salop model with 3 firms in equidistant locations and with two types of consumers, differing in their transportation costs. Firms do not know the market structure and they apply least squares learning to learn about demand conditions. They approximate the (piecewise linear) demand function with a linear perceived demand function and they maximize their profit subject to their perceived demand function. The model has three kinds of equilibria: a general self-sustaining equilibrium, the Nash equilibrium and an asymmetric learning-equilibrium. In a self-sustaining equilibrium firms approximate the true demand function correctly only in the equilibrium point but the approximation is incorrect for any other point. In the Nash equilibrium and the ALE firms correctly learn the linear part of the true demand function on which they operate. In the ALE the high-price firm attracts high-type consumers only and thus it underestimates the demand for low prices. Therefore it does not perceive it profitable to charge a lower price. We prove that both the Nash equilibrium and the ALE are locally stable, thus the model can have coexisting stable equilibria. As firms use more observations in the regression, the Nash equilibrium is reached more often. In contrast, the ALE is reached more often as the transportation cost of high-type consumers increases.

In the model we made some assumptions whose effect should be discussed. First of all, we introduced heterogeneity on the demand side of the market. This is not an unrealistic assumption as consumers could easily differ in their transportation costs, moreover it makes the model more general. With homogeneous consumers, the true demand function is still piecewise linear so least squares learning can lead to an SSE or to the Nash equilibrium. The ALE, however, does not exist since if a firm does not attract any consumers, then it will charge a lower price eventually since the observations with zero demand will move the parameter estimates in a direction that yields a lower price. Thus, consumer heterogeneity is essential for having an asymmetric outcome.

We augmented least squares learning with seemingly ad hoc rules. These rules are, however, quite reasonable and they improve the learning process. According to the random price rule, firms choose a random price when the estimation leads to an upward-sloping demand function or when the pricing rule gives a price that is lower than the marginal cost. In either case, the normal pricing rule does not give the profit maximum so it should not be used. When the own price effect is positive, the perceived demand function is not sensible economically. Firms should charge an infinitely high price in this case to maximize their profit. It should be clear for firms that the positive own price effect comes from some estimation problem, therefore they should not follow the normal pricing rule. Note that it would not affect the final outcome much if firms charged a very high price in this situation. They would face zero demand in the given period, leading to new parameter estimates that give a lower price. This would affect the distribution of final prices in an SSE but we would still find convergence for the Nash equilibrium or to the ALE when firms do not use all observations in the regression. Concerning the case where the pricing formula gives a price that is lower than the marginal cost, firms are always better off by choosing a random price that may yield a positive profit than following the pricing formula that results in a negative profit.

We introduced the no jump rule and the impossible estimation rule to overcome a problem when the process starts to converge. In that case there is not enough variation in the observations, leading to imprecise parameter estimates. This drives away the process from the equilibrium. It might seem as if these rules lead to an artificial stability in the model as we require firms to use the same price as in the previous period but actually these rules rule out an artificial instability. Note that this problem occurs only when the process started to converge. Thus, firms observe that prices have settled down around some values and then the new parameter estimates lead to an unexpectedly large price. First of all firms might be reluctant to make such a big price change, secondly after observing the time series of prices it should be clear that this sudden price change comes from a numerical problem, therefore it is better not to change the price. As we discussed before, it would not change our results much if firms asked a high price (until some upper bound is reached) instead of keeping the price fixed. Concerning the impossible estimation rule, when parameter estimates cannot be obtained, then firms either choose a price randomly or they fix the price as we suggest. Note that a rule like the no jump rule or the impossible estimation rule is essential for having convergence in a model that is not subject to noise (e.g. demand shock) when firms do not use all past observations in the regression. If the process were to converge to a certain value, then estimation is not possible when each observation perfectly correspond to the steady state. Thus, it needs to be specified what happens when parameter estimates cannot be calculated. Keeping the pice unchanged is a reasonable solution to this problem.

We have seen that in the Nash equilibrium and the asymmetric learning-equilibrium, perceived demand functions are correctly specified in the neighborhood of the equilibrium price. This makes these outcomes more robust than the self-sustaining equilibria in Brousseau and Kirman (1992) in the sense that in case of an SSE a firm would discover that its perceived demand function is misspecified by choosing a slightly different price. This is not the case for the Nash equilibrium and the asymmetric learning-equilibrium.[12]

Let us now elaborate on whether our results could still hold in more general models. If we increase the number of firms or the number of consumer types in the Salop framework, the the true demand function remains piecewise linear and thus firms may learn only one linear part correctly. Therefore firms can reach an SSE or the Nash equilibrium again. Our conjecture is that firms can reach more than one asymmetric outcomes in this situation. There are more possibilities for having asymmetric outcomes as firms may attract different consumer groups or a given consumer group can be served by different number of firms. We expect that the same kind of outcomes can occur in different market structures as well. Consider for example a situation where firms are competing on different markets and demand functions are nonlinear on both submarkets. In this case aggregate demand functions will not be smooth, there will be a kink at the price for which the demand on the smaller market becomes 0. If the functional form of the perceived demand functions correspond to the functional form of the true demand function on one (or both) of the submarkets, then the same kind of asymmetric equilibria can occur. One firm might be focusing on the bigger market only, charging a higher price, while other firms might be active on both markets.

The aim of this paper was to analyze the properties of least squares learning in a situation when firms can observe the prices of all other firms and the functional form they use in the regression is locally correct. We found that the same kind of outcome can be reached as under an incorrect functional form or with not observing all the relevant variables (SSE), and firms can also reach the outcome that corresponds to correct functional form and full observation (the Nash equilibrium). Additionally, we found another kind of outcome, the asymmetric learning-equilibrium, that was not found in other settings. Based on our findings, we can make the following recommendation for applying least squares learning. Not all observations should be used in the regression as observations might correspond to different demand regions and therefore the estimation may not yield a good approximation. Firms should experiment with their price (by charging a lower sales price) every now and then to make sure that they do not get locked up in a suboptimal situation. Finally, the learning process should be augmented with extra rules in order to overcome the numerical problem with estimation when there is not enough dispersion in the observations.

[12] Note, however, that this difference is due to the different informational structure of the two models. In Brousseau and Kirman (1992) firms can observe their own actions only, whereas they have full information about the actions in our model.

References

Anufriev, M., Kopányi, D., and Tuinstra, J. (2013). Learning cycles in Bertrand competition with differentiated commodities and competing learning rules. *Journal of Economic Dynamics and Control*, 37(12):2562–2581.

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.

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

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.

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.

Salop, S. C. (1979). Monopolistic competition with outside goods. *The Bell Journal of Economics*, 10(1):141–156.

Tirole, J. (1988). *The theory of industrial organization*. MIT press.

Weddepohl, C. (1995). A cautious price adjustment mechanism: chaotic behavior. *Journal of Economic Behavior & Organization*, 27(2):293–300.

A Appendix

A.1 Derivation of demand function (1)

PROOF.

When $p_i < p_j - \frac{1}{3}s$, then firms j and k are driven out of the market since firm i attracts each consumer. So $q_i = 1$ in this case.

When $p_i = p_j - \frac{1}{3}s$, then firm k is driven out of the market, only firms i and j attract consumers. For the given prices the consumer at the location of firm j is indifferent between the two firms. There exists another indifferent consumer, say Y, on segment ji. Let us suppose that consumer Y is located at distance y from firm i. Then it holds for this consumer that $p_i + sy = p_j + s\left(\frac{2}{3} - y\right)$. Using that $p_i = p_j - \frac{1}{3}s$, we get $y = 0.5$. Since the consumer at the location of firm j is indifferent between the two firms, all the consumers between firm j and consumer Y are indifferent between firms i and j. The amount of these consumers is $1 - \frac{1}{3} - \frac{1}{2} = \frac{1}{6}$. Half of them chooses firm j, firm i attracts all other consumers. Thus, $q_i = \frac{11}{12}$.

When $p_j - \frac{1}{3}s < p_i < p_k - \frac{1}{3}$, then firm k is driven out of the market and there is one indifferent consumer on segment ij and another one on segment ji. Suppose that the indifferent consumer on segment ji lies at distance y from firm i. For this consumer it holds that $p_i + sy = p_j + s\left(\frac{2}{3} - y\right)$, from which $y = \frac{p_j - p_i}{2s} + \frac{1}{3}$. Suppose that the other indifferent consumer (on segment ij) lies at distance x from firm i. A similar calculation as before yields $x = \frac{p_j - p_i}{2s} + \frac{1}{6}$. Then firm i gets a demand of $q_i = x + y = \frac{p_j - p_i}{s} + \frac{1}{2}$.

When $p_i = p_k - \frac{1}{3}$, then all three firms are active and the consumer at the location of firm k is indifferent between firms i and k. There is an indifferent consumer on segment ij. Let is suppose that he is located at distance x from firm i. In this case, $x = \frac{p_j - p_i}{2s} + \frac{1}{6}$. There is another indifferent consumer on segment jk. Let us suppose that he is located at distance y from firm k. Similar calculations as before yield $y = \frac{p_j - p_k}{2s} + \frac{1}{6}$. Since the consumers between the location of firm k and the location of the indifferent consumer on segment jk are indifferent between firms i and k, firm i attracts half of them. In this case, the demand of firm i is given by $q_i = x + \frac{1}{3} + 0.5y = \frac{3p_j - p_k - 2p_i}{4s} + \frac{7}{12}$. Using that $p_i = p_k - \frac{1}{3}s$, the previous expression simplifies to $q_i = \frac{3}{4} + \frac{3p_j - 3p_k}{4s}$.

When $p_k - \frac{1}{3}s < p_i < p_j + \frac{1}{3}s$, then all three firms are active and there is an indifferent consumer between each 2 firms. Suppose that the indifferent consumer between firms i and j is located at distance x from firm i. Then, as we calculated before, $x = \frac{p_j - p_i}{2s} + \frac{1}{6}$. Similarly, if the indifferent consumer between firms i and k lies at distance y from firm i, then $y = \frac{p_k - p_i}{2s} + \frac{1}{6}$. Thus, $q_i = x + y = \frac{p_j + p_k - 2p_i}{2s} + \frac{1}{3}$.

For $p_i = p_j + \frac{1}{3}$, the consumer at the location of firm i is indifferent between firms i and j. Let the indifferent consumer between firms i and k be located at distance y from firm i. Then $y = \frac{p_k - p_i}{2s} + \frac{1}{6}$, as before. Since the consumers that are between this indifferent consumer and firm i, are indifferent between firms i and j, firm i attracts only half of them. Thus, $q_i = 0.5y = \frac{p_k - p_i}{4s} + \frac{1}{12}$.

When $p_i > p_j + \frac{1}{3}$, even the consumer at the location of firm i chooses firm j. Thus, firm i is driven out of the market: $q_i = 0$.

□

A.2 The proof of Proposition 2.1

PROOF.

First note that $p_i \geq c$ must hold for each firm in equilibrium. Otherwise the firm with the lowest price, say firm j, would always face a positive demand and would make a certain loss on each product. The firm could increase its profit by choosing a higher price for which its profit is at least 0. This can be achieved by any $p_j \geq c$.

Now we will show that each firm must face a positive demand in equilibrium. To see this suppose that firm i is driven out of the whole market by firm j, that is $p_j < p_i - \frac{1}{3}S$. Let $c \leq p_j \leq p_k$ without loss of generality. In this case, firm i can increase its profit by choosing the price $p_i = c + \varepsilon$ for a sufficiently small but positive ε. For this price firm i will not be driven out of the market since $p_i - \frac{1}{3}S = c + \varepsilon - \frac{1}{3}S < c$ for a sufficiently small ε, meaning that firm i can only be driven out of the market with a price that is smaller than the marginal cost. This, as we have seen, cannot occur in equilibrium.

The condition that each firm must have a positive demand in equilibrium implies that all three firms must attract high-type consumers. Thus, equilibria can differ only in the number of firms attracting low-type consumers. There might be three possibilities: 3, 2 or 1 firm attracts low-type consumers. We investigate these cases separately.

Case 1: symmetric Nash equilibrium

When all three firms attract low-type consumers, then firm i faces the following demand function: $D_i(p) = \frac{2}{3} + (p_j + p_k - 2p_i)\left(\frac{1}{2S} + \frac{1}{2s}\right)$. To see this note that there is one low-type and one high-type indifferent consumer between any two firms. The low-type indifferent consumer between firms i and j is at the distance $x = \frac{p_j - p_i}{2s} + \frac{1}{6}$ from firm i. A similar formula applies for the high-type indifferent consumer and for the indifferent consumers between firms i and k.

Firm i maximizes its profit with respect to its price: $\max_{p_i} (p_i - c)D_i(p)$. The first-order conditions for firms 1, 2 and 3 respectively are

$$\frac{2}{3} + (p_2 + p_3 - 2p_1)\left(\frac{1}{2S} + \frac{1}{2s}\right) - \left(\frac{1}{S} + \frac{1}{s}\right)(p_1 - c) = 0, \tag{A.1}$$

$$\frac{2}{3} + (p_1 + p_3 - 2p_2)\left(\frac{1}{2S} + \frac{1}{2s}\right) - \left(\frac{1}{S} + \frac{1}{s}\right)(p_2 - c) = 0, \tag{A.2}$$

$$\frac{2}{3} + (p_1 + p_2 - 2p_3)\left(\frac{1}{2S} + \frac{1}{2s}\right) - \left(\frac{1}{S} + \frac{1}{s}\right)(p_3 - c) = 0. \tag{A.3}$$

Subtracting (A.2) from (A.1) yield $\frac{5}{2}(p_2 - p_1)\left(\frac{1}{S} + \frac{1}{s}\right)$, from which $p_1 = p_2$. Similarly, subtracting (A.3) from (A.2) gives $p_1 = p_3$. Let p^N denote this common price. Then the first-order conditions simplify to $\frac{2}{3} - \left(\frac{1}{S} + \frac{1}{s}\right)(p^N - c) = 0$, from which

$$p^N = \frac{2Ss}{3(S+s)} + c.$$

The corresponding profits are $\pi^N = \frac{2}{3}(p^N - c) = \frac{4}{9}\frac{Ss}{S+s}$.

The price vector $p = (p^N, p^N, p^N)$ constitutes a Nash equilibrium only if none of the firms has an incentive to deviate from this price unilaterally. A firm can deviate in two possible ways. It can drive out the two other firms from the market of the low-type consumers or it can drive out the other firms from the whole market.[13]

Let us first consider the case when firm 1 chooses $p^N - \frac{1}{3}S \leq p_1 \leq p^N - \frac{1}{3}s$. In this case firm 1 attracts the low-type consumers and the three firms share the high-type consumers. Thus, firm 1 faces the following demand function: $D_1(p) = \frac{4}{3} + \frac{p^N - p_1}{S}$. To find the optimal price, the following constrained optimization problem needs to be solved:

$$\max_{p_1 \leq p^N - \frac{s}{3}} (p_1 - c)\left(\frac{4}{3} + \frac{p^N - p_1}{S}\right).$$

The Karush-Kuhn-Tucker conditions yield

$$\frac{4}{3} + \frac{p^N - p_1}{S} - \frac{1}{S}(p_1 - c) \geq 0$$
$$\left(\frac{4}{3} + \frac{p^N - p_1}{S} - \frac{1}{S}(p_1 - c)\right)\left(p^N - \frac{s}{3} - p_1\right) = 0.$$

Let us suppose that $\frac{4}{3} + \frac{p^N - p_1}{S} - \frac{1}{S}(p_1 - c) = 0$. This gives $p_1 = \frac{2S}{3} + \frac{p^N + c}{2}$. We need to check whether the

[13] We do not have to consider marginal deviations from p_N since the first-order conditions imply that the local profit maximum is reached at p_N.

condition $p_1 \leq p^N - \frac{s}{3}$ is satisfied.

$$\begin{aligned}
\frac{2S}{3} + \frac{p^N + c}{2} &\leq p^N - \frac{s}{3} \\
0 &\leq \frac{p^N - c}{2} - \frac{2S}{3} - \frac{s}{3} \\
0 &\leq \frac{Ss}{3(S+s)} - \frac{2S}{3} - \frac{s}{3} \\
0 &\leq Ss - 2S(S+s) - s(S+s) \\
0 &\leq -2S^2 - 2Ss - s^2,
\end{aligned}$$

where we used the formula for p^N. The last condition is never satisfied so we can conclude that $p_1^D = p^N - \frac{1}{3}s$ is the optimal deviation in this case. The corresponding demand and profit are $q_1^D = \frac{1}{3}\left(4 + \frac{s}{S}\right)$ and $\pi_1^D = \left(p^N - \frac{1}{3}s - c\right)\frac{1}{3}\left(4 + \frac{s}{S}\right)$, which simplifies to $\pi_1^D = \frac{s}{9}\frac{S-s}{S+s}\left(4 + \frac{s}{S}\right)$. Firm 1 does not have an incentive to deviate if $\pi^N \geq \pi_1^D$, which gives

$$\begin{aligned}
\frac{4}{9}\frac{Ss}{S+s} &\geq \frac{s}{9}\frac{S-s}{S+s}\left(4 + \frac{s}{S}\right) \\
4S &\geq (S - s)\left(4 + \frac{s}{S}\right) \\
0 &\geq -3s - \frac{s^2}{S}.
\end{aligned}$$

The last inequality is always satisfied as $S, s > 0$. Thus, this deviation is never profitable.

Now let us consider the other deviation when firm 1 drives the other firms out from the whole market. In this case $p_1 < p^N - \frac{1}{3}S$ should hold. Note, however, that $p^N - \frac{1}{3}S = \frac{2Ss}{3(S+s)} + c - \frac{1}{3}S = \frac{1}{3}S\frac{s-S}{S+s} + c < c$ as $s - S < 0$. This means that firm 1 would have to charge a price below the marginal cost to attract every consumer, leading to negative profits.

Thus, firms do not have an incentive to deviate unilaterally from the price p^N. The price vector $p = (p^N, p^N, p^N)$ is the unique symmetric Nash equilibrium.

Case 2: asymmetric situation with 2 firms serving low-type consumers

Now we will show that the situation in which exactly one firm focuses only on high-type consumers, cannot constitute a Nash equilibrium. Assume without loss of generality that firm 3 charges a high price such that only high-type consumers buy from firm 3: $\min\{p_1, p_2\} + \frac{S}{3} \geq p_3 \geq \min\{p_1, p_2\} + \frac{s}{3}$. In this situation the demand functions are as follows:

$$D_1(p) = \frac{5}{6} + \frac{p_2 - p_1}{s} + \frac{p_2 + p_3 - 2p_1}{2S}, \tag{A.4}$$
$$D_2(p) = \frac{5}{6} + \frac{p_1 - p_2}{s} + \frac{p_1 + p_3 - 2p_2}{2S}, \tag{A.5}$$
$$D_3(p) = \frac{1}{3} + \frac{p_1 + p_2 - 2p_3}{2S}. \tag{A.6}$$

Profit maximization yields the following first-order conditions (for firms 1, 2 and 3 respectively)

$$\frac{5}{6} + \frac{p_2 - p_1}{s} + \frac{p_2 + p_3 - 2p_1}{2S} - \left(\frac{1}{s} + \frac{1}{S}\right)(p_1 - c) = 0, \tag{A.7}$$
$$\frac{5}{6} + \frac{p_1 - p_2}{s} + \frac{p_1 + p_3 - 2p_2}{2S} - \left(\frac{1}{s} + \frac{1}{S}\right)(p_2 - c) = 0, \tag{A.8}$$
$$\frac{1}{3} + \frac{p_1 + p_2 - 2p_3}{2S} - \frac{1}{S}(p_3 - c) = 0. \tag{A.9}$$

If a Nash equilibrium exists in the given situation, it must be the solution of these first-order conditions. By subtracting (A.8) from (A.7), it can be seen that $p_1 = p_2$ must hold. Therefore, let $p_1 = p_2 = p_L$ and $p_3 = p_H$. The first-order conditions then simplify to

$$\frac{5}{6} + \frac{p_H - p_L}{2S} - \left(\frac{1}{s} + \frac{1}{S}\right)(p_L - c) = 0, \tag{A.10}$$
$$\frac{1}{3} + \frac{p_L - p_H}{S} - \frac{1}{S}(p_H - c) = 0. \tag{A.11}$$

Subtracting (A.11) from (A.10) yields $\frac{1}{2} + \frac{3}{2}\frac{p_H - p_L}{S} - \frac{1}{s}(p_L - c) + \frac{1}{S}(p_H - p_L) = 0$, from which

$$\frac{p_H - p_L}{2S} = \frac{p_L - c}{5s} - \frac{1}{10}. \tag{A.12}$$

Combining (A.10) with (A.12) gives $\frac{11}{15} + \frac{p_L - c}{5s} - \left(\frac{1}{s} + \frac{1}{S}\right)(p_L - c) = 0$. Solving this equation for p_L yields

$$p_L = \frac{11Ss}{12S + 15s} + c.$$

Plugging this expression for p_L in (A.12) yields an equation that can be solved for p_H. The solution simplifies to

$$p_H = \frac{2S^2 + 8Ss}{12S + 15s} + c.$$

Note that the previous calculations yield admissible prices only when $p_L + \frac{1}{3}S \geq p_H \geq p_L + \frac{1}{3}s$, or equivalently $\frac{1}{3}s \leq p_H - p_L \leq \frac{1}{3}S$. Using that

$$p_H - p_L = \frac{2S^2 - 3Ss}{12S + 15s},$$

the condition $\frac{1}{3}s \leq p_H - p_L$ leads to $4Ss + 5s^2 \leq 2S^2 - 3Ss$, or equivalently $2\left(\frac{S}{s}\right)^2 - 7\frac{S}{s} - 5 \geq 0$. Solving this quadratic equation gives that $\frac{S}{s} \geq \frac{7+\sqrt{89}}{4}$ must hold.

The condition $p_H - p_L \leq \frac{1}{3}S$ leads to $12S + 15s \geq 6S - 9s$, from which $6S + 24s \geq 0$. This condition is satisfied as $S, s > 0$. Thus, this type of asymmetric Nash equilibrium may exist only when $\frac{S}{s} \geq \frac{7+\sqrt{89}}{4}$.

The price vector $p = (p_L, p_L, p_H)$ constitutes a Nash equilibrium only if none of the firms has an incentive to deviate unilaterally. Now we will show that either the high-price firm or the low-price firms can earn a higher profit by charging a different price. First, let us calculate the profits under $p = (p_L, p_L, p_H)$. Plugging the prices in demand functions (A.4)-(A.6) yields $q_1 = q_2 = q_L = \frac{11S+11s}{12S+15s}$ and $q_3 = q_H = \frac{2S+8s}{12S+15s}$. The corresponding profits are $\pi_1 = \pi_2 = \pi_L = \frac{121Ss(S+s)}{(12S+15s)^2}$ and $\pi_3 = \pi_H = S\left(\frac{2S+8s}{12S+15s}\right)^2$.

First let us suppose that the high-price firm deviates and charges $p_3^D = p_L$, where superscript D refers to deviation. In that case $q_3^D = \frac{2}{3}$ since all three firms charge the same price. The corresponding profit is $\pi_3^D = \frac{2}{3}\frac{11Ss}{12S+15s}$. This deviation leads to a higher profit for firm 3 when

$$\begin{aligned} \frac{2}{3}\frac{11Ss}{12S+15s} &> S\left(\frac{2S+8s}{12S+15s}\right)^2 \\ 11s(12S+15s) &> 6S^2 + 48Ss + 96s^2 \\ 0 &> 6S^2 - 84Ss - 69s^2. \end{aligned} \tag{A.13}$$

Now let us suppose that firm 1 deviates by charging $p_1^D = p_H$. In that case firm 1 serves the high-type consumers only so it faces a similar demand function as (A.6). Thus, its demand equals $q_1^D = \frac{1}{3} + \frac{p_L - p_H}{2S} = \frac{6S+13s}{2(12S+15s)}$ and the corresponding profit is $\pi_1^D = \frac{6S+13s}{12S+15s}\frac{S(S+4s)}{12S+15s}$. This deviation leads to a higher profit for firm 1 when

$$\begin{aligned} \frac{6S+13s}{12S+15s}\frac{S(S+4s)}{12S+15s} &> \frac{121Ss(S+s)}{(12S+15s)^2} \\ (6S+13s)(S+4s) &> 121s(S+s) \\ 6S^2 - 84Ss - 69s^2 &> 0. \end{aligned} \tag{A.14}$$

Comparing conditions (A.13) and (A.14), we find that one of the firms always has an incentive to deviate whenever $6S^2 - 84Ss - 69s^2 \neq 0$. Now we will show that the high-price firm has an incentive to deviate even if the previous equation holds with equality. Note that we did not consider the *optimal* deviation in the previous calculations. We only showed that there exists a deviation that is more profitable under certain conditions. When $6S^2 - 84Ss - 69s^2 = 0$ holds, firm 3 is indifferent between charging p_L and p_H (keeping the price of the other two firms fixed):

$$\pi_3(p_L, p_L, p_L) = \pi_3(p_L, p_L, p_H). \tag{A.15}$$

We will now show that the marginal profit of firm 3 is not equal to 0 at $p = (p_L, p_L, p_L)$. This implies that a marginal deviation from $p_3 = p_L$ (in the appropriate direction) yields a strictly higher profit, thus $p = (p_L, p_L, p_H)$ cannot be a Nash equilibrium.

The marginal profit of firm 3 at $p = (p_L, p_L, p_L)$ can be calculated using (A.3):

$$\left.\frac{\partial \pi_3}{\partial p_3}\right|_{p=(p_L,p_L,p_L)} = \frac{2}{3} - \left(\frac{1}{S} + \frac{1}{s}\right)(p_L - c).$$

Plugging in the formula for p_L yields $\frac{2}{3} - \frac{S+s}{Ss}\frac{11Ss}{12S+15s}$, which simplifies to $\frac{7S+9s}{12S+15s}$. This expression is always positive since $S, s > 0$. Thus, firm 3 can get a strictly higher profit by marginally increasing its price: $\pi_3(p_L, p_L, p_L + \varepsilon) > \pi_3(p_L, p_L, p_L)$ for a small enough $\varepsilon > 0$. Combining the last inequality with (A.15) shows that $p = (p_L, p_L, p_H)$ cannot be a Nash equilibrium.

Thus, we have shown that one of the firms can always get a higher profit by unilaterally changing its price. We can conlude that there does not exist an asymmetric Nash equilibrium in pure strategies where exactly two firms attract low-type consumers.

Case 3: asymmetric situation with 1 firm serving low-type consumers

Now we will show that the situation in which two firms focus only on the high-type consumers, cannot constitute a Nash equilibrium. Assume without loss of generality that firm 1 charges a low price such that it attracts every low-type consumer: $p_1 + \frac{S}{3} \geq \{p_2, p_3\} \geq p_1 + \frac{s}{3}$. In this situation the demand functions are as follows:

$$D_1(p) = \frac{4}{3} + \frac{p_2 + p_3 - 2p_1}{2S},$$
$$D_2(p) = \frac{1}{3} + \frac{p_1 + p_3 - 2p_2}{2S},$$
$$D_3(p) = \frac{1}{3} + \frac{p_1 + p_2 - 2p_3}{2S},$$

with the corresponding first-order conditions for profit maximization

$$\frac{4}{3} + \frac{p_2 + p_3 - 2p_1}{2S} - \frac{1}{S}(p_1 - c) = 0, \tag{A.16}$$
$$\frac{1}{3} + \frac{p_1 + p_3 - 2p_2}{2S} - \frac{1}{S}(p_2 - c) = 0, \tag{A.17}$$
$$\frac{1}{3} + \frac{p_1 + p_2 - 2p_3}{2S} - \frac{1}{S}(p_3 - c) = 0. \tag{A.18}$$

By subtracting (A.18) from (A.17), it can be seen that $p_2 = p_3$ must hold. Let $p_1 = p_L$ and $p_2 = p_3 = p_H$. Then the first-order conditions simplify to

$$\frac{4}{3} + \frac{p_H - p_L}{S} - \frac{1}{S}(p_L - c) = 0, \tag{A.19}$$
$$\frac{1}{3} + \frac{p_L - p_H}{2S} - \frac{1}{S}(p_H - c) = 0. \tag{A.20}$$

Subtracting (A.20) from (A.19) yields $1+\frac{3}{2S}(p_H-p_L)+\frac{1}{S}(p_H-p_L)=0$. This equation, however, does not give an admissible solution. Since every coefficient is positive and the right hand side is 0, $p_H < p_L$ must hold, which contradicts the assumption $p_H \geq p_L+\frac{s}{3}$. Thus, there exists no asymmetric pure-strategy Nash equilibrium in which exactly one firm serves the low-type consumers. □

A.3 The proof of Proposition 3.2

PROOF.
To simplify notation, let $\hat{D}_i(p)=A_i-b_{ii}p_i$, where $A_i=a_i+b_{ij}p_j+b_{ik}p_k$. Then using (4) the best response price is given by

$$p_i^{\mathrm{BR}}=\frac{A_i}{2b_{ii}}+\frac{c}{2}. \tag{A.21}$$

Since $p_i=p_i^{\mathrm{BR}}$ in an SSE, the perceived demand is given by

$$\hat{D}_i(p)=\frac{A_i-b_{ii}c}{2}. \tag{A.22}$$

Note that 9 variables characterize an SSE under the simplified notation: 1 price and the 2 parameters of the perceived demand function for each firm. On the other hand, there are 6 conditions (best response price and equality of actual and perceived demands for each firm). Thus, the system of equations that characterizes an SSE might be solved, with 3 free variables. We will now show that for a given price vector $p=(p_i,p_j,p_k)$ we can find values of $\{A_i,b_{ii}\}_{i=1}^3$ such that the system is in an SSE.

From (A.21) we get $A_i=b_{ii}\,(2p_i-c)$. Combining this with (A.22), the perceived demand simplifies to $\hat{D}_i(p)=b_{ii}\,(p_i-c)$. Since the actual and the perceived demands must coincide at price vector $p=(p_i,p_j,p_k)$, it must hold that $D_i(p)=b_{ii}\,(p_i-c)$, from which

$$b_{ii}=\frac{D_i(p)}{p_i-c}. \tag{A.23}$$

Combining this with the previous formula for A_i yields

$$A_i=\frac{D_i(p)}{p_i-c}\,(2p_i-c)\,. \tag{A.24}$$

Thus, for a given price vector $p=(p_i,p_j,p_k)$, formulas (A.23) and (A.24) specify the values of b_{ii} and A_i under which the system is in an SSE.

Let us investigate which price vectors lead to an economically sensible perceived demand function. That is, we want to characterize the set of prices for which $b_{ii}>0$ and $A_i>0$ (i.e. the perceived demand function is downward-sloping and the "intercept" is positive)[14].

It follows from (A.23) that $b_{ii}>0$ if and only if $D_i(p)>0$ and $p_i>c$. Under these conditions, $A_i>0$ is satisfied as well. □

[14] For having an economically sensible perceived demand function, one might consider introducing the conditions $a_i>0$,

A.4 The proof of Proposition 3.3

PROOF.

We know from Proposition 3.2 that $D_i(p) > 0$ must hold for each firm. Since each firm must face a positive demand in an SSE, each firm must attract high-type consumers. This implies that there are three possible SSE in which firms correctly learn one linear part of the true demand function, depending on whether 1, 2 or 3 firms serve low-type consumers.[15]

When all 3 firms attract low-type consumers, then demand conditions are characterized by $D_i(p) = \frac{2}{3} + (p_j + p_k - 2p_i)\left(\frac{1}{2S} + \frac{1}{2s}\right)$ (see Case 1 in Section A.2). The best response function can be derived from first-order conditions (A.1)-(A.3). As we have seen, these first-order conditions have a unique solution, which corresponds to the Nash equilibrium of the model with known demand. Thus, when all 3 firms serve both consumer types and firms correctly learn the corresponding linear part of the true demand function, then the Nash equilibrium is the unique steady state of the learning process.

When only 2 firms attract low-type consumers, then demand conditions are characterized by (A.4)-(A.6) (see Case 2 in Section A.2). The corresponding best response functions can be derived from first-order conditions (A.7)-(A.9). These first-order conditions have a unique solution, in which the low-price firms charge $p_L = \frac{11Ss}{12S+15s} + c$ and the high-price firm asks the price $p_H = \frac{2S^2+8Ss}{12S+15s}$. We have also seen that this outcome exists only when $\frac{S}{s} \geq \Sigma_1$. Even though this outcome is not a Nash equilibrium of the model with known demand, it is a steady state of the learning process. The reason behind this is that firms do not know that it would be profitable to change their price unilaterally since they approximate the demand function with a linear function, implying that they do not know that they would get a much higher demand by undercutting other firms. Thus, when only 2 firms serve both consumer types and firms correctly learn the corresponding linear part of the true demand function, then the unique steady state is given by 2 firms charging p_L and 1 firm charging p_H. We refer to this outcome as asymmetric learning-equilibrium.

We have seen that when only 1 firm serves the low-type consumer, then first-order conditions (A.16)-(A.18) do not yield an admissible solution. Therefore the learning process does not have a steady state in this situation.

This shows that the Nash equilibrium and the asymmetric learning-equilibrium are the only steady states in which all three firms correctly learn the linear part of the true demand function on which they operate.

$b_{ij} > 0$ and $b_{ik} > 0$ in addition to the condition $A_i > 0$. Note, however, that these extra conditions do not restrict the set of admissible prices further as for a given positive A_i one can always find values for a_i, b_{ij} and b_{ik} such that $A_i = a_i + b_{ij}p_j + b_{ik}p_k$ holds and the conditions on the signs are satisfied.

[15] Note that these are exactly the same cases that we analyzed in Section A.2.

$\square$

A.5 Comparison of profits in the ALE

We have seen in Section A.2, Case 2 that profits in the ALE are given by $\pi_L = \frac{121Ss(S+s)}{(12S+15s)^2}$ and $\pi_H = S\left(\frac{2S+8s}{12S+15s}\right)^2$. The low-price firms make a higher profit only if $121s(S+s) > 4(S+4s)^2$, from which $0 > 4\left(\frac{S}{s}\right)^2 - 89\frac{S}{s} - 57$. This gives $\frac{S}{s} < \frac{89+11\sqrt{73}}{8} \approx 22.87$.

A.6 The proof of Lemma 3.5

Proof.

When price observations are aligned, estimation yields the true parameters that characterize the given linear part of the demand function. Under symmetrically aligned price observations the parameter estimates are given by $a_i = \frac{2}{3}$, $b_{ii} = \frac{1}{S} + \frac{1}{s}$ and $b_{ij} = b_{ik} = \frac{1}{2S} + \frac{1}{2s}$ (see Section A.2, Case 1 for the corresponding demand function). Thus, using (4), the best response of firm i is

$$p_i^{\mathrm{BR}} = \frac{\frac{2}{3} + \left(\frac{1}{2S} + \frac{1}{2s}\right)(p_j + p_k)}{2\left(\frac{1}{S} + \frac{1}{s}\right)} + \frac{c}{2}.$$

Then $|p_i^{\mathrm{BR}} - p_j^{\mathrm{BR}}| = \frac{1}{4}|p_j - p_i|$ for any two firms. Since price observations were symmetrically aligned, $|p_j - p_i| < \frac{s}{3}$ holds and therefore $|p_i^{\mathrm{BR}} - p_j^{\mathrm{BR}}| < \frac{s}{3}$ is also satisfied. Thus, adding the best-response prices to the price observations gives a symmetrically aligned set again.[16]

When firms i and j attract both types of consumers while firm k attracts high-type consumers only, then firms learn the following demand parameters: $a_i = a_j = \frac{5}{6}$, $b_{ii} = b_{jj} = \frac{1}{S} + \frac{1}{s}$, $b_{ij} = b_{ji} = \frac{1}{2S} + \frac{1}{s}$, $b_{ik} = b_{jk} = \frac{1}{2S}$, $a_k = \frac{1}{3}$, $b_{kk} = \frac{1}{S}$ and $b_{ki} = b_{kj} = \frac{1}{2S}$ (see Section A.2, Case 2 for the corresponding demand functions). Using (4), the best-response prices are given by

$$p_i^{\mathrm{BR}} = \frac{\frac{5}{6} + \left(\frac{1}{2S} + \frac{1}{s}\right)p_j + \frac{1}{2S}p_k}{2\left(\frac{1}{S} + \frac{1}{s}\right)} + \frac{c}{2},$$
$$p_j^{\mathrm{BR}} = \frac{\frac{5}{6} + \left(\frac{1}{2S} + \frac{1}{s}\right)p_i + \frac{1}{2S}p_k}{2\left(\frac{1}{S} + \frac{1}{s}\right)} + \frac{c}{2},$$
$$p_k^{\mathrm{BR}} = \frac{\frac{1}{3} + \frac{1}{2S}(p_i + p_j)}{2\frac{1}{S}} + \frac{c}{2}.$$

Then $|p_i^{\mathrm{BR}} - p_j^{\mathrm{BR}}| = \frac{\left(\frac{1}{2S}+\frac{1}{s}\right)}{2\left(\frac{1}{S}+\frac{1}{s}\right)}|p_j - p_i| = \frac{2S+s}{4S+4s}|p_j - p_i| < \frac{s}{3}$ since $\frac{2S+s}{4S+4s} < 1$ and $|p_j - p_i| < \frac{s}{3}$ because

[16] Notice the contraction mapping feature of playing the best-response price. This implies that symmetrically aligned prices converge to the same value. Since prices are best response to each other, firms will reach the Nash equilibrium in this case.

price observations were asymmetrically aligned. Thus, the first condition in the definition is satisfied.[17]

Let us suppose that $p_i \leq p_j$ in the most recent price observation. In that case, $p_j^{\mathrm{BR}} \leq p_i^{\mathrm{BR}}$ and it must hold for having asymmetrically aligned price observations that $p_j^{\mathrm{BR}} + \frac{s}{3} < p_k^{\mathrm{BR}} < p_j^{\mathrm{BR}} + \frac{S}{3}$. Using the formulas above, it can be shown that

$$p_k^{\mathrm{BR}} - p_j^{\mathrm{BR}} = \frac{1}{12(S+s)}\left[2S^2 - 3Ss - 3Sp_i + (3S+3s)p_j - 3sp_k\right].$$

We will first show that the condition $p_k^{\mathrm{BR}} - p_j^{\mathrm{BR}} < \frac{S}{3}$ is always satisfied. Using the formula for $p_k^{\mathrm{BR}} - p_j^{\mathrm{BR}}$, the condition simplifies to

$$-\frac{S}{3}\left(2\frac{S}{s} + 7\right) + \left(1 + \frac{S}{s}\right)(p_j - p_i) < p_k - p_i.$$

The left-hand side is smaller than $-\frac{S}{3}\left(2\frac{S}{s} + 7\right) + \left(1 + \frac{S}{s}\right)\frac{s}{3}$ since $p_j - p_i < \frac{s}{3}$. It is easy to see that this expression is always negative. On the other hand, the right-hand side is positive since $p_k - p_i > \frac{s}{3}$. Thus, $p_k^{\mathrm{BR}} - p_j^{\mathrm{BR}} < \frac{S}{3}$ is always satisfied.

Next let us consider the condition $p_k^{\mathrm{BR}} - p_j^{\mathrm{BR}} > \frac{s}{3}$. Using the formula for $p_k^{\mathrm{BR}} - p_j^{\mathrm{BR}}$, the condition simplifies to

$$\frac{s}{3}\left[2\left(\frac{S}{s}\right)^2 - 7\frac{S}{s} - 4\right] + \left(1 + \frac{S}{s}\right)(p_j - p_i) > p_k - p_i. \tag{A.25}$$

The left-hand side of the inequality is greater than or equal to $\frac{s}{3}\left[2\left(\frac{S}{s}\right)^2 - 7\frac{S}{s} - 4\right]$ as $p_j - p_i \geq 0$. The right-hand side is smaller than $\frac{S}{3}$ since price observations are asymmetrically aligned. Thus, a sufficient condition for (A.25) to hold is that

$$\frac{s}{3}\left[2\left(\frac{S}{s}\right)^2 - 7\frac{S}{s} - 4\right] \geq \frac{S}{3}.$$

This leads to $2\left(\frac{S}{s}\right)^2 - 8\frac{S}{s} - 4 \geq 0$, for which $\frac{S}{s} \geq 2 + \sqrt{6}$ must hold. Thus, when the latter condition holds, asymmetrically aligned price observations always remain asymmetrically aligned, irrespective of the exact values in the last price observation. On the other hand, when $\frac{S}{s} < 2 + \sqrt{6}$, condition (A.25) has to hold for the most recent price observation in order to have asymmetrically aligned price observations again.

Since price observations were asymmetrically aligned, $\frac{s}{3} < p_k - p_i$. Thus, the following condition must hold

$$\frac{s}{3} < \frac{s}{3}\left[2\left(\frac{S}{s}\right)^2 - 7\frac{S}{s} - 4\right] + \left(1 + \frac{S}{s}\right)(p_j - p_i).$$

As we have seen, playing the best response works as a contraction mapping for the low-price firms, therefore $p_j - p_i \to 0$ if the information set always remains asymmetrically aligned. Thus, $\frac{s}{3} \leq \frac{s}{3}\left[2\left(\frac{S}{s}\right)^2 - 7\frac{S}{s} - 4\right]$

[17]Note the contraction mapping feature again, which implies that the low-price firms will reach the same price if the information set always remains asymmetrically aligned.

must hold. This leads to $1 \leq 2\left(\frac{S}{s}\right)^2 - 7\frac{S}{s} - 4$, from which $\frac{S}{s} \geq \frac{7+\sqrt{89}}{4}$. Thus, when the latter condition does not hold, then an asymmetrically aligned information set cannot stay asymmetrically aligned.

□

A.7 The proof of Proposition 3.6

PROOF.

We will now show that both the Nash equilibrium and the ALE are locally stable equilibria. First we will describe the system in the neighborhood of the equilibria and then we show that the eigenvalues of the Jacobian are always less than 1 in absolute value. First we focus on the Nash equilibrium.

Part 1: Stability of the Nash equilibrium

When prices in the information set are symmetrically aligned, then firms learn the correct demand parameters of the linear part on which they operate. Moreover, as we have seen in Lemma 3.5, updating the information set with the best-response prices results in a symmetrically aligned information set again. Thus, the parameters of the perceived demand functions do not change in this case. Then the perceived demand function of firm i is always given by $\hat{D}_i(p) = \frac{2}{3} - \left(\frac{1}{S} + \frac{1}{s}\right) p_i + \left(\frac{1}{2S} + \frac{1}{2s}\right) p_j + \left(\frac{1}{2S} + \frac{1}{2s}\right) p_k$ (see Section A.2, Case 1 for the demand parameters of the relevant linear part). Then the next-period price of firm i is given by

$$p_{i,t+1} = \frac{1}{3}\frac{Ss}{S+s} + \frac{1}{4}\left(p_{j,t} + p_{k,t}\right) + \frac{1}{2}c.$$

This holds for every firm i, therefore the Jacobian of the system is given by

$$J = \begin{pmatrix} 0 & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & 0 \end{pmatrix}.$$

The characteristic equation is given by

$$k(\lambda) = -\lambda^3 + \frac{3}{16}\lambda + \frac{1}{32} = 0.$$

It is easy to see that $k(\lambda)$ can also be expressed as $k(\lambda) = -\left(\lambda - \frac{1}{2}\right)\left(\lambda + \frac{1}{4}\right)^2$. Thus, the eigenvalues of the Jacobian are $\lambda_1 = \frac{1}{2}$ and $\lambda_2 = -\frac{1}{4}$. Both eigenvalues are smaller than 1 in absolute value, therefore the Nash equilibrium is locally stable.

Part 2: Stability of the ALE

When prices in the information set are asymmetrically aligned, then firms learn the correct demand parameters of the linear part on which they operate. Moreover, as we have seen in Lemma 3.5, updating the information set with the best-response prices results in an asymmetrically aligned information set again when $\frac{S}{s} > \Sigma_2$.[18] Thus, the parameters of the perceived demand functions do not change in this case. Suppose that firms i and j are the low-price firms and firm k is the high-price firm. Then the perceived demand function of firm i is always given by $\hat{D}_i(p) = \frac{5}{6} - \left(\frac{1}{S} + \frac{1}{s}\right) p_i + \left(\frac{1}{2S} + \frac{1}{s}\right) p_j + \frac{1}{2S} p_k$ while that of firm k is $\hat{D}_k(p) = \frac{1}{3} - \frac{1}{S} p_k + \frac{1}{2S} p_i + \frac{1}{2S} p_j$ (see Section A.2, Case 2 for the demand parameters of the relevant linear part). Then the next-period price of firms i and k are given by

$$\begin{aligned} p_{i,t+1} &= \frac{5}{12}\frac{Ss}{S+s} + \frac{1}{4}\frac{2S+s}{S+s} p_{j,t} + \frac{1}{4}\frac{s}{S+s} p_{k,t} + \frac{1}{2}c, \\ p_{k,t+1} &= \frac{1}{6}S + \frac{1}{4}\left(p_{i,t} + p_{j,t}\right) + \frac{1}{2}c. \end{aligned}$$

The next-period price of firm j is given by a similar formula as for firm i, we just need to switch i and j. Then the Jacobian of the system is given by

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

where $A = \frac{1}{4}\frac{2S+s}{S+s}$, $B = \frac{1}{4}\frac{s}{S+s}$ and $C = \frac{1}{4}$. The characteristic equation is given by

$$k(\lambda) = -\lambda^3 + \left(A^2 + 2BC\right)\lambda + 2ABC = 0.$$

It is easy to see that $k(\lambda)$ can also be expressed as $k(\lambda) = -(\lambda + A)\left(\lambda^2 - A\lambda - 2BC\right)$. Thus, one eigenvalue is $\lambda_1 = -A$. This eigenvalue is always smaller than 1 in absolute value. $A > 0$ since $S, s > 0$. $A < 1$ if and only if $2S + s < 4(S + s)$, which is always satisfied.

The other two eigenvalues are the solutions of the equation $\lambda^2 - A\lambda - 2BC = 0$. The discriminant is $D = A^2 + 8BC > 0$, so there are two real roots: $\lambda_{2,3} = \frac{A \pm \sqrt{A^2+8BC}}{2}$. Root $\lambda_2 = \frac{A+\sqrt{A^2+8BC}}{2}$ has the larger absolute value. Its absolute value is smaller than 1 if and only if $\sqrt{A^2 + 8BC} < 2 - A$, from which - using that $A < 1$ - $A^2 + 8BC < 4 - 4A + A^2$. This simplifies to the condition $A + 2BC < 1$.

Plugging in the values for A, B and C yields $A + 2BC = \frac{1}{4}\frac{2S+s}{S+s} + 2\frac{1}{4}\frac{s}{S+s}\frac{1}{4} = \frac{1}{8}\frac{4S+3s}{S+s}$. This is smaller than 1 in absolute value if and only if $4S + 3s < 8(S + s)$, which is satisfied for any $S, s > 0$.

[18] Even if $\frac{S}{s} > \Sigma_2$ does not hold, we can consider a sufficiently small neighborhood of the ALE for which the updated information set is asymmetrically aligned. This can be done as (A.25) holds for (p_L, p_L, p_H) whenever the ALE exists.

Thus, all three eigenvalues are smaller than 1 in absolute value, implying that the ALE is locally stable.

□

A.8 Robustness check with respect to the number of periods and the number of different initializations

		estimated coefficients		true coefficients	
		10000 periods	**10000 runs**	**10000 periods**	**10000 runs**
S	**2**	99.7% (98.4%)	99.6% (97.9%)	100%	100%
		0.0% (0.0%)	0.0% (0.0%)	0.0%	0.0%
	4	99.5% (97.6%)	99.5% (97.6%)	100%	100%
		0.0% (0.0%)	0.0% (0.0%)	0.0%	0.0%
	4.2	74.7% (73.7%)	79.0% (77.6%)	97.0%	96.3%
		25.1% (24.4%)	20.8% (20.1%)	3.0%	3.7%
	4.35	67.2% (66.0%)	73.1% (71.7%)	95.3%	94.6%
		32.7% (30.9%)	26.6% (25.7%)	4.7%	5.4%
	4.5	63.1% (62.1%)	69.1% (68.0%)	93.2%	93.2%
		36.3% (35.0%)	30.6% (29.6%)	6.8%	6.8%
	10	11.0% (10.8%)	13.0% (12.7%)	56.4%	53.2%
		88.8% (82.4%)	86.6% (81.1%)	43.6%	46.8%

Table 7: Outcome of simulations with 10000 periods and 10000 different initializations. The proportion of outcomes in the 0.001 and the 0.0001-neighborhoods (in brackets) of the Nash equilibrium (upper numbers) and the asymmetric learning-equilibrium (lower numbers) for different values of S and $\tau = 8$. Other parameters: $s = 1$ and $c = 1$.