Price competition, Sequential Search and Sellers' rationality

Eric DARMON, CREM-CNRS and University of Rennes 1

7, place Hoche ; F-35000 Rennes (France)

Eric.darmon@univ-rennes1.fr

Roger WALDECK, GET ENST-Bretagne, LUSSI and ICI-UBO (EA-2652)

Technopole Brest Iroise CS 83818, F-29238 Brest Cedex 3, France

Roger.waldeck@enst-bretagne.fr

Abstract

We consider a sequential search model with two types of consumers: ('high cost's) consumers who incur a positive search cost at each visit and informed consumers who visit all the firms at no cost. The objective is to compare Nash market predictions with a market with adaptive sellers using reinforcement learning. Simulation Results show that Reinforcement Learning never converges to Nash equilibrium. However, using Nash equilibrium, one can predict how the two first order statistics vary with respect to market structure. Concerning the average posted price, we find that Nash predictions are well respected, albeit variations are in general less pronounced with reinforcement learners. Concerning price dispersion, only variations with respect to the number of firms follow the same shape as Nash. Yet, increasing the proportion of informed consumers seems to have contradicting effects on price dispersion although Nash predicts in general a decrease of price dispersion in the case of study. The impact of the number of firms and the proportion of informed consumers is in a decrease of the accepted price of informed consumers although Nash predicts such a decrease only for the second parameter within the parameter range considered.

Keywords: search market equilibrium – reinforcement learning process – sequential search strategy – numerical computation

1. Introduction

A large body of literature in economics was concerned with the existence of price dispersion in market for an homogenous good (e.g. (Pratt et al., 1979). (Stigler, 1969) proposed an explanation of price dispersion based on the search costs of buyers searching from a fixed price distribution. Although Stigler considered only the buyer side of the market, it was shown that price dispersion may arise as the sellers' Nash pricing equilibrium in the presence of two types of buyers differing according to the magnitude of their search costs (e.g. (Salop and Stiglitz, 1977). In addition, with firms having constant marginal cost of production, the equilibrium price distribution is in mixed strategy (Rob 1985 ; Varian 1980). (Waldeck, forth.) studies the impact of variations in the market structure – i.e. of the proportion of informed consumers, of the level of the high search cost and of the number of firms – on the average posted price and price dispersion of the mixed strategy Nash equilibrium. The main contribution of Waldeck's paper is to stress the non trivial relationship between information, prices and price dispersion. First, the standard presumption, that increasing the proportion of informed consumers decreases the expected price paid by informed or uninformed consumers, is true for Sequential Search (SS). Second, a higher search intensity of informed consumers, i.e. an increase in the sample size due to an increase in the number of sellers increases the expected market price. Price dispersion is an inverse U-shaped function of the proportion of informed consumers. The objective of the paper is to test whether the Nash predictions are met in the presence of "less rational" sellers. We assume that sellers do not know buyers' search characteristics or have limited computational capacities. Consequently, they cannot formulate a profit-maximizing price strategy and will rely on trial-and-errors processes (adaptive learning) to establish a suitable price strategy. For that, we needed to formulate an agent-based model based on the original analytical model.

The key finding of this paper is that Reinforcement Learning never converges to Nash. However, we can compute either analytically or numerically the two first moments of the price distribution, and study the impact of market structure on these moments. This analysis reveals that Nash predictions are not rejected for first order statistics notably average posted prices albeit variations are in general less pronounced with reinforcement learners. Concerning price dispersion, only variations with respect to the number of firms follow the same shape as Nash. But increasing the proportion of informed consumers seems to have contradicting effects on price dispersion although Nash predicts in general a decrease of price dispersion in the case of study. Besides, the accepted price paid by informed consumers decreases with the number of firms and with the proportion of informed consumers although Nash predicts such a decrease only for the second parameter.

Section 2 presents the simulation model. Section 3 introduces the Nash Search Equilibrium. Grounding on these propositions, we elaborate several hypotheses and present the protocol used to test them. Section 4 derives the main insights from the comparison between NSE and RL outcomes. Section 5 discusses the results and concludes.

2. The Model

There are S sellers (indexed by s) and B buyers (indexed by b). They respectively produce and consume an indivisible and homogeneous item. Each market period t ( ) is divided into three sub-periods: first, sellers post prices (no bargaining, no change in prices inside the period); second, buyers visit sellers and transact; third, sellers evaluate the profit generated by the pricing rule currently used and reward this rule. Using the updated reward, they post a new price for period (t+1), etc.

2.1. Buyers

There are $B = 1000$ buyers. Each of them needs to purchase one unit of the good at maximum price v. This value is common to all buyers and will further be set w.l.g. to 100. There are two types of buyers differentiated by their search cost. Buyers of Type *1* are in proportion a of the total population of buyers (with $a \in [0,1]$) and are characterized by a zero search cost. Since sampling an additional seller is at no cost, we will further assume that this category of buyers visit systematically all sellers and shop at the best available price. Therefore, let us call buyers of that type as *informed buyers.* In contrast to informed buyers, buyers of Type 2 are characterized by a positive search cost β incurred at each visit. Let us thus call this second type of buyers as high search cost buyers. Parameter β is defined as a proportion of the valuation of the good v, hence $\beta \in]0,1]$.

Uninformed buyers are characterized by a stopping rule. The first search is costless (hence, all consumers decide to enter into the market). When the last price observation is $\tilde{p}$, a buyer decides to sample an additional seller if the following stopping rule is met :

Stopping rule: $$\sum_{p=0}^{\tilde{p}-1} (\tilde{p} - p) \cdot f(p) > \beta v$$

with f_i the frequency associated to the observation of p_i.

This rule states that if the expected gain from one additional visit is more than the cost βv of the visit, the buyer keeps on searching. The frequencies f_i used to compute the expected benefit, are derived from the posted price distribution[1]. We will suppose that

[1] We discuss this hypothesis in the conclusion.

buyers have a memory of m periods. Hence, this distribution at period t is computed from the prices set by each seller at Periods $(t-m),(t-m+1),...,(t-1)$.

If the stopping rule is not met when all sellers have been visited, the buyer transacts with the best-price seller (thus assuming an intra period recall) with no additional cost. Yet, there is no inter-period recall.

2.2. Sellers

There are S sellers. *Per* unit cost of production (c) is constant and identical for every seller, hence we assume w.l.g. that $c=0$. There is no capacity constraint. At the beginning of each period, sellers independently set the price they will post for this period. Prices are discrete variables ranging from c to v with an exogenous step ϕ, sellers then have $\left(1+(v-c)/\phi\right)$ price rules namely {Price c, Price $c+\phi$, Price $c+2\phi$, …, Price v}. Let us set $\phi=1$, so that possible prices are simply $\{0,1,2,...100\}$[2]. At each period, sellers decide which price (p_t) to post by using a simple RL process[3]: each price corresponds to a rule. In turn, each rule is defined by a triplet {Condition , Action , Fitness}. Since any price rule *can* be applied whatever sellers' environment, there is no condition to fit to use one specific rule. The action part is simply the price sellers decide to post during the period. The fitness of a rule (F) measures its ability to satisfy sellers' objective (i.e. profit). At the first period, sellers anticipate that all posted prices yield the same expected profit. Hence, all pricing rules are initially endowed with the same fitness noted F_0. This initial fitness F_0 can also be interpreted as sellers' expected belief about market-profitability.

$$F^i_{t=0} = F_0 = \delta(Bv) \quad (\forall i) \text{ with } \delta \in [0,1] \tag{1}$$

When coefficient $\delta=0$ (resp. $\delta=1$), sellers are "pessimistic" (resp. optimistic) and expect a null profit (*resp.* maximal profit). During the subsequent periods, sellers revise the expected profit of each rule according to the actual instantaneous profits π_t generated by that rule. Let n_t denote the number of actual transactions implemented at period *t*. For any pricing rule *i*, the fitness of rule *i* (F^i_t) is thus updated as follows:

---

[2] We assume that sellers know the reservation price v rather than learn it. If not such a price would rapidly disappear with RL since it generates zero profit whenever posted.

[3] See (Kirman and Vriend, 2001) for more details about the learning process; see also (Sutton, 1991) for a general description of RL processes.

$$F_{t+1}^i = F_t^i + \alpha\left(\pi_t - F_t^i\right) \text{ with } \pi_t := (p_t - c)(n_t) \qquad \textit{if rule i is} \text{ used during period } t$$
$$F_{t+1}^i = F_t^i \qquad \textit{otherwise} \qquad (2)$$

Coefficient α $\left(\alpha \in \left]0,1\right]\right)$ measures the relative weight attached to current experiences. This updating rule dynamics mimics an adaptive trial-and-error process: if, during the current period, the pricing rule *i* generates a higher profit than what it did in the past, its reward increases and hence, the probability of selecting this rule is reinforced for the next periods (see Equation (3) below).

At each period, each seller sets a price according to a trembling-hand process. In other terms, sellers' behavior is characterized by a trade off between exploration and exploitation: at each period, each rule is selected with a probability drawn from a Boltzmann distribution described by Equation (3):[4]

$$prob\left\{Select\ Rule\ i\right\} = \frac{e^{\frac{F_t^i}{\tau}}}{\sum_j e^{\frac{F_t^j}{\tau}}} \text{ with } \tau > 0 \qquad (3)$$

Parameter τ sets the trade off between exploration and exploitation: as $\tau \to 0$, sellers tend to select only the rules which have generated the highest payoffs in the past (choice of the "greedy action"). As τ increases, sellers tend to explore alternative rules more frequently tending to random choices as $\tau \to \infty$.

3. Nash predictions and Hypothesis testing

3.1. Theoretical predictions and hypotheses

The Nash Symmetric Equilibrium (NSE) is a solution concept to study this system in the context of rational choices. In this context, NSE leads to several theoretical propositions summed by *Propositions 0* to *3.* All proofs are detailed in Waldeck [forth.]. The first proposition describes the equilibrium price distribution. The remaining propositions depict the variations of the three main statistics on this market (average posted price, average price accepted by informed buyers and standard deviation of posted prices) with respect to buyers' search behaviors (i.e. parameters a, S and β).

---

[4] F_0^i are the fitness initialized by $\delta(Bv)$ $(\forall i)$

Proposition 0 (price distribution). *The (symmetric) Nash equilibrium price cumulative distribution is as follows:*

- *General case* $\left(0<a<1\right)$ *:* $F(p;a,S)=1-\left(\frac{(1-a)\left(r(\beta,a,S)-p\right)}{Sap}\right)^{\frac{1}{S-1}}$ *with the lower (resp. upper) bound of the distribution defined by* $b\left(a,S,r(\beta,a,S)\right)=(1-a)r(\beta,a,S)/\left((S-1)a+1\right)$ *(resp.* $r(\beta,a,S)=\min\left\{v,\frac{\beta v}{1-\overline{p}_{FSS}(a,s)}\right\}$*). Variable* $\overline{p}_{FSS}(a,S)$ *is the mean of the price distribution is the case of Fixed Sample Size Search (FSS) with cumulative price distribution defined by* $\left[F(p;a,S)\right]_{b(a,S,v=1))}^{r=v=1}$.

Proposition 1 (Reservation price of uninformed buyers). *At the Nash Symmetric Equilibrium, the expected number of visits of type 2 buyers is equal to 1. Type 2 buyers will therefore be called uninformed buyers. The reservation price of uninformed buyers is defined by* $r(\beta,a,S)=\min\left\{v,\frac{\beta}{1-\overline{p}_{FSS}(a,S)}\right\}$*). The reservation price of uninformed consumers is an increasing function of* S *and* β*. It is a decreasing function of* a*.*

Proposition 2. (expected posted price) *i) The NSE expected posted price decreases from the monopoly price v to the marginal cost 0 as a varies from 0 to 1; ii) The NSE expected posted price increases with s* $\left(i.e.\ s_1<s_2\Rightarrow E(p_1;a,s_1)<E(p_2;a,s_2)\right)$ *and with* β *as long as* $r(\beta,a,S)<v$. [5]

Proposition 3. (standard deviation of posted prices). *Price dispersion is i) an increasing function of the magnitude of the search cost* β *as long as* $r(\beta,a,S)<v$*, ii) an inverse U-shaped function of parameter a* $\left(0<a<1\right)$*, and iii) an inverse U-shaped function of* s *decreasing to zero when* s *tends to infinity*[6].

Proposition 4. (expected price paid by informed buyers) The NSE expected price paid by informed buyers decreases with a $\left(i.e.\ s_1<s_2\Rightarrow E(p_1;a,s_1)<E(p_2;a,s_2)\right)$ and increases with β as long as $r(\beta,a,S)<v$.

---

[5] Note that since type 2 buyers visit exactly one seller with probability one , expected price paid by uninformed consumers are identified with expected posted price.

[6] The shape of price dispersion with respect to S is only a numerical result.

For each of these propositions, we formulate a related hypothesis to test whether the RL-outcomes fit the NSE propositions. These hypotheses are related to the case where β =10%[7].

- **Hypothesis 0.** The RL price distribution converges to the Nash distribution.
- **Hypothesis 1.1.** The RL mean price is decreasing in a .
- **Hypothesis 1.2.** The RL mean price is increasing in S .
- **Hypothesis 1.2 bis (weaker form).** The RL mean price is non decreasing in S.
- **Hypothesis 2.1.** The RL standard deviation is a decreasing function of a for a greater than 0.2 or for all a greater than or equal to 0.1 when S=2.
- **Hypothesis 2.1 bis (weaker form).** The RL does not show an opposite sign of variation to Nash with respect to a .
- **Hypothesis 2.2.** The RL standard deviation is a increasing function of S .
- **Hypothesis 2.2 bis (weaker form).** The RL standard deviation is non decreasing in S.
- **Hypothesis 3.1.** The RL mean price paid by informed consumers is decreasing in a .

3.2. Test implementation

To compare NSE- with RL- outcomes, we ran 25 simulations for each parameters configuration (a, S). The process has been simulated in Java: the appendix reports an abridged pseudo code of the model. Source code, classes, pseudo code and an executable version of the program are available on request. We define each simulation as a session. For each session, we computed the posted price distribution, the aggregate mean and the standard deviation of the price distribution over the last 100 periods.

We present the results of simulations on (a, S) using a model where $\tau = 0.1$, α = 0.8, δ = 1 and the number of rounds is T = 2000. The number of rounds (T) has been chosen so as to ensure that the process converged to a stationary position within this time horizon. As a benchmark, we consider a situation where *i)* individuals' initial expectations are optimistic $(\delta = 1)$[8], *ii)* price volatility is not too high or random (as it might be with higher temperature) and *iii)* sellers' memory is not too long $(\alpha = 0.8)$. Typically, as α = 0.8, only 3 to 4 periods of time remain significant in individual decisions.

---

[7] To conduct this preliminary analysis, we chose an intermediate value of this coefficient. When $\beta \to 0$, consumers tend to visit the whole set of sellers (hence all uninformed consumers become informed), while when $\beta \to 1$, consumers only visit one seller per period.

[8] The impact of δ seems only to be on the speed of learning but does not affect the results otherwise.

4. RL Predictions

We first test the previous hypotheses in the reference case: to test HO, we implemented a Kolmogorov-Smirnoff adequation test (testing that the RL and the NSE distribution are equal). This test holds only for continuous distribution i.e. for any value of (a, S) such that $0 < a < 1 \;\; S \neq 1$.

Result 1: The stationary distribution with reinforcement sellers does **not** *converge to the Nash distribution for parameters 0 < a < 1 . Hypothesis 0 does not hold.*

Figure 1 shows the discrepancy between the Nash (F theor) and the RL (Fobs) cumulative distributions.

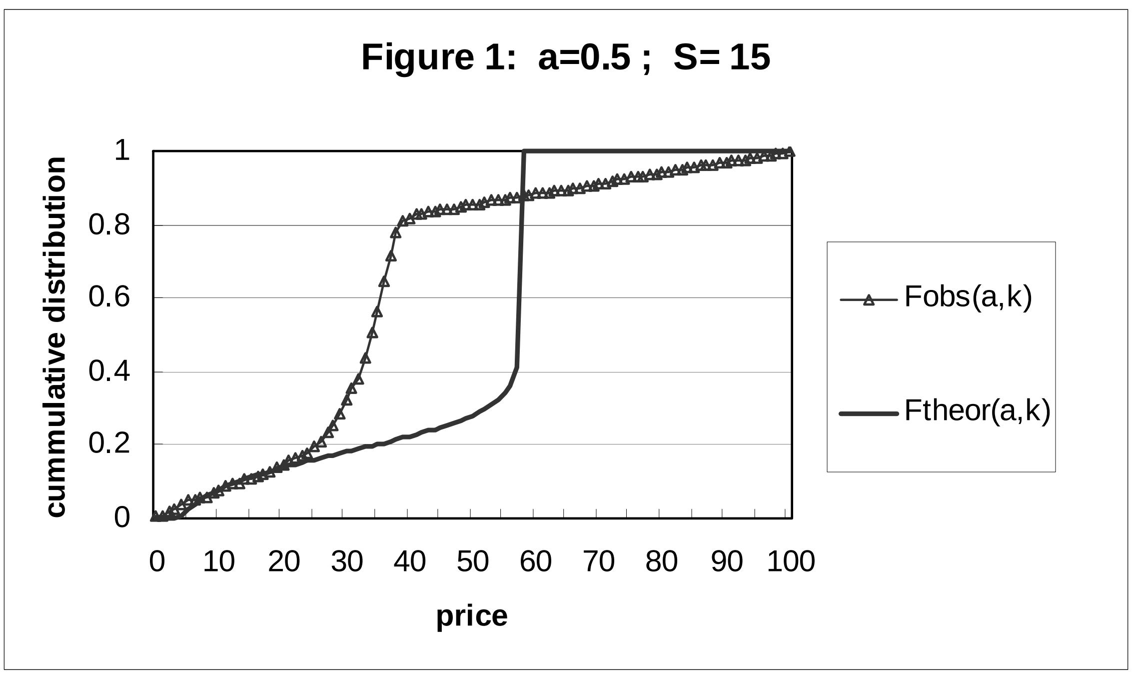


Figure 1. Nash (F theor) and the RL (Fobs) cumulative distributions

4.1. Posted Price

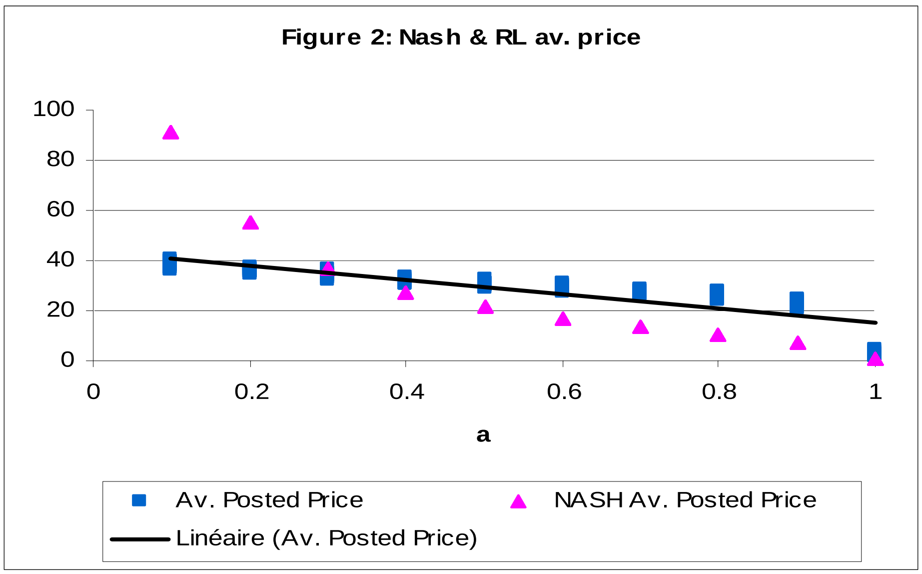
We expect from the Nash analysis, that the average posted price is decreasing with respect to *a* and increasing with respect to *S*. To test this gross relationship, we ran a simple econometric analysis and tested the following equation:

$$\text{AV\_POSTED\_PRICE} = \alpha_0 + \alpha_1 a + \alpha_2 S$$

Table 1 (see Annex) reports the results of this test using the previously described dataset. It evidences that a has a significantly negative impact on the average price posted by sellers[9]. Besides, coefficient estimate associated to S are both positive and highly significant.

Result 2-1: Hypotheses 1.1 and 1. 2 are checked from an econometric point of view, that is the general shape is such that average posted price is decreasing with a and increasing with S.

One can graphically illustrate these relationship by making cross sections of the dataset for specific values of a or S. For instance, Figure 2 displays such plot for $S = 5$ (right figure) and $a = 0.5$ (left figure).



[9] We used several other specifications (log-log, semi-log, before and after correcting for heteroskedasticity) and both the sign and the significance of coefficient estimates are preserved when considering these alternative specifications.

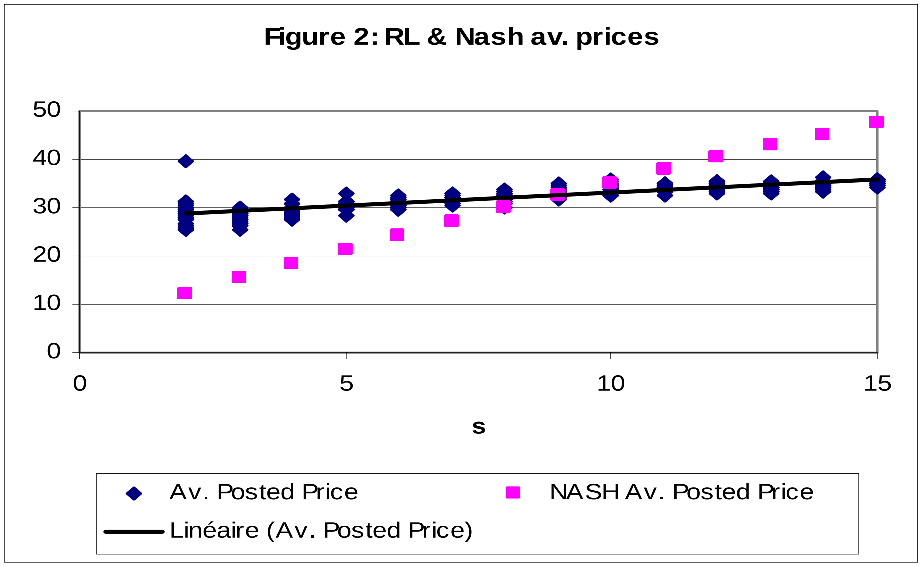


Figure 2 (2-1 and 2-2). Nash and RL mean price as a function of *a* (left) and *s* (right)

We then split the dataset and considered each value of *a* and *S* independently. For all parameter configurations *S*, the slope of the regression line is negative (cf. figure 2, right) . However, the results for S are more contrasted for the cases a = 0.1, 0.2, 0.3 for which the increase in average price with respect to S are less pronounced.

This result is confirmed by a one-sided Wilcoxon signed rank test. For this, we test whether over the 25 simulations there are a significant number of cases where the average price decreases with an increase of a by a step of 0.1 and a significant number of cases where the average price increases with an increase of S by step of 1.

Result 2-2:

- *Average price is decreasing for all figures a, except 1 over 112 (alpha = 0.05). Hypothesis 1.1 is true.* [10]
- *Only 53.8 % of figures showed a significant increase with S and 5.1% show a significant decrease with S. a= 0.1 , 0.2 and 0.3 show the lowest rate of increase (less than 40% of the figures increases with S). The weaker hypotheses 1.2 (bis) that the RL mean price is non decreasing in S may nevertheless be accepted in a large majority of cases.*

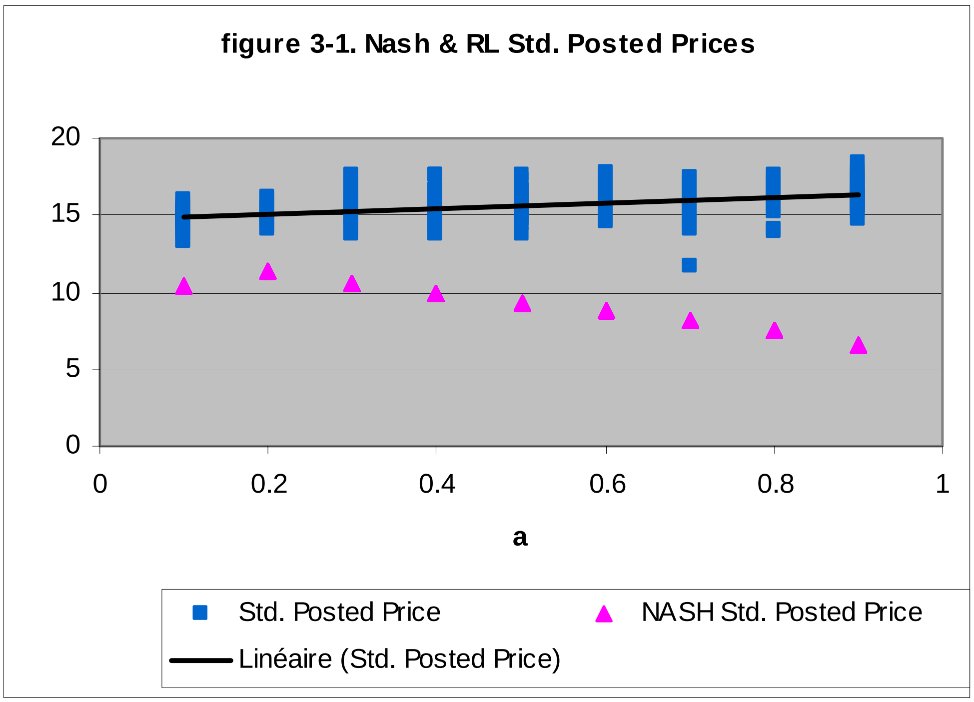
In conclusion, one can say that Nash and RL lead to the same conclusion with respect to *a* whereas there only a gross relationship with respect to S. However, in general, Nash is not false with respect to *S*, that is only a few figures exhibit a significantly opposite sense of variation. However, for low values of *a* the relationship between posted prices and S is rather flat.

[10] All tests will be presented for α = 0.05. There is no significant change by taking α = 0.01 or α = 0.1.

4.2. Price Dispersion

From the Nash analysis, we expect when $\beta = 0.1$ that the dispersion of posted price is an inverse-function. The maximal price dispersion is reached when a=0.2[11]. Besides, we expect price dispersion to be decreasing with respect to *S*.

Figure 3-1 and 3-2 show a plot for $S = 5$ (fig 3.1) and $a = 0.5$ (fig 3.2). As shown in these plots, Nash seems to be met for S whereas this might not be the case for a.



[11] In fact, price dispersion is a inverse U-shaped function of a and in the case S=2 the maximum is attained for a value below or at 0.1. Since a varies in step of 0.1 , this maximum of the Nash distribution is only grossly approximated.

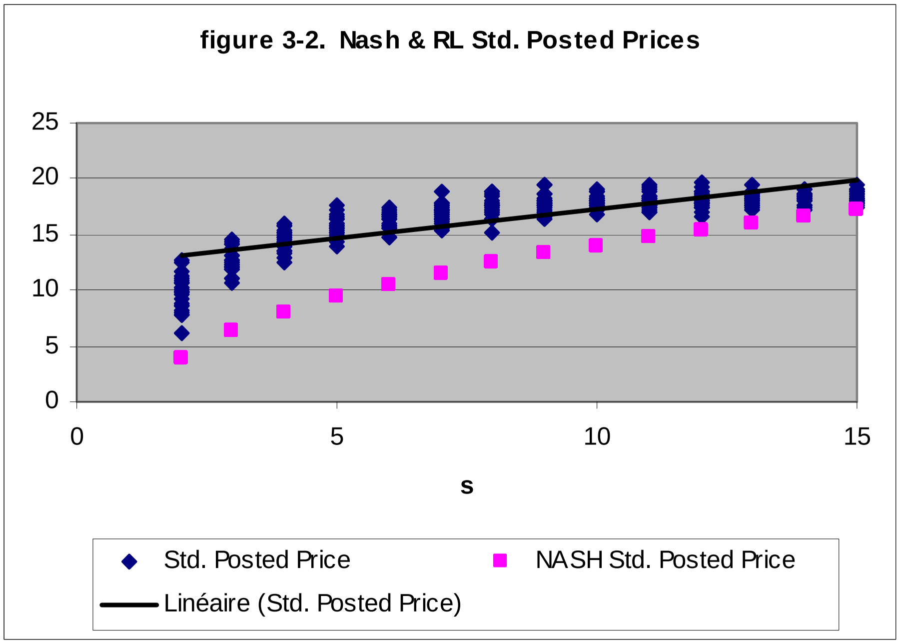


Figure 3 (31 and 32). Nash and RL standard deviation of price as a function of *a* (left) and *s* (right)

Result 3-1:

- ***Results for a:***
  - *we divide the test in two: for a≥0.2 we perform a linear econometric test. For a < 0.2, the Wilcoxon test shows that out of 14 figures, 2 do not validate the Nash assumption whereas none invalidates Nash (i.e. none show a sign in a opposite direction to Nash).*
  - *Only 25.9 % of the cases are consistent with Nash whereas 11.6 % show a opposite variation to Nash. These values are for an increase in a from 0.2 to 0.3 up to 0.5 to 0.6 where the predicted decrease is in fact a significant increase in some cases. (Wilcoxon ranked sign test).*
  - *Nash is not met. Neither hypothesis 2.1 or its weaker form may be accepted.*

- ***Results for** S:*
  - *Simple econometric analysis evidences a significant increase of standard deviation with respect to S in line with Nash predictions (table 2).*
  - *However, only 43.6% of figures show a significant increase whereas none shows a contrary sense of variation (Wilcoxon ranked sign test).*
  - *One can conclude that also Nash predictions concerning price dispersion are not false with respect to S. Its weaker form 2.2bis is also true.*

4.3. Price paid by informed Buyers

We expect from the Nash analysis that the proportion of informed buyers has a negative impact on the price paid by the informed consumers. We also expect that the number of firms has a non significant effect on this average price[12].

Results 4.1: Table 3 shows that all coefficients of the linear regression are significantly decreasing meaning that the general shape show that prices accepted by informed consumers decrease both with a and with S. This result is additionally confirmed by the Wilcoxon test which shows that only 4 figures out 112 show a non significant decrease with a. 68% of the figures show a significant decrease with respect to S whereas none shows a significant increase with respect to S (all test one sided at 5%).

5. Conclusion

Our purpose in that paper was to know whether sellers endowed with simple learning rules (reinforcement) may learn to play sophisticated price strategies such as those described by Nash Search Equilibrium. For that, we designed an agent-base model close to a standard search model (two categories of buyers, sequential search strategies), and compared the strategies played by adaptive sellers to those played at Nash Equilibrium. A general result is that the two price distributions never coincide, meaning that adaptive sellers can never guess the exact Nash price strategy in such context. However, one can use a weaker criterion before rejecting Nash. We thus compared how the first-order statistics (mean price, price dispersion) vary with respect to buyers' search behavior. We can conclude that with respect to a change of a (proportion of informed buyers), Nash predictions are met for both the posted price set by sellers and the accepted price paid by informed buyers. Simulation Results for price dispersion are more mitigated. Only 25% of figures confirm Nash prediction but in the majority of cases, a has no impact on price dispersion when considering a step by step increase. With respect to a change in S, one can conclude that Nash is not false as to the prediction concerning posted price and price dispersion. That is RL does not predict opposite results.

This study has several implications regarding the comparisons between formal and agent-based models. One needs to separate the problem raised by the validation of the agent-based model to those raised by the formulation of the analytical model.

The first issue is essentially statistical. In this paper, we used two types of validation processes: regression analysis and Wilcoxon tests. However, the results derived from both procedures may be disturbed by the trembling-hand process introduced by the rule selection. This kind of 'noise' may then explain why in some cases some relationships are only weakly (e.g. non increase rather than increase) checked. The difficulty is then to

---

[12] This is not a theoretical but only a numerical result from the Nash distribution which states that accepted price do not vary whenever $r(\beta) < v$ which is the case when $\beta = 0.1$ and $S < 16$.

compare a noise-free price distribution defined at Nash Equilibrium, to a noised distribution inherited from simulation results.

The other issues relates to the formulation of the agent-based models and specifically to the search behaviour of buyers. The standard notion of Nash equilibrium used in analytical models necessarily implies some simplifications. In our case, the game is a one-period game and the simplifying assumption is that the price distribution is common knowledge. More specifically, it is assumed that sellers can endogenize the search behaviour of buyers when computing their price strategy. When done, buyers become “aware” of this distribution when starting their search process. It is thus shown that informed buyers visit the whole set of sellers and that uninformed ones visit one seller only. Yet, this setting may have diverse agent-based “translations”. First, one may ask where the price distribution that buyers use to compute their search, comes from. Do they know the current price distribution? Do they the previous price distribution (t-1) rather? Do they know the previous price distribution**s** (t-H, t-H-1, … t-1 where H depicts the size of buyers’ memory) instead? Or do they only posses some private information depending on the past search experience of each buyer? Second, we assumed in that paper that informed buyers systematically visit the whole set of sellers, and buy at the lowest available price. An alternative assumption would have been to suppose that informed consumers visit all sellers until they reach the minimum price in their memory. Moreover if buyers have different memories the stopping rules could well be different across informed consumers and it is not clear what the outcome of such a process will be These final remarks open some stimulating directions for further research in Agent based - and Equilibrium based - models comparison.

6. Acknowledgements

Table 1: Results of linear regressions for average posted price

	Variations with respect to *a*	**Variations with respect to** *s*	**Significativité globale de la régression**	**Heteroskedasticity test**	**Test RAMSEY**
Linear regression	- (-17.5)***	+ (.49)***	R2 aj. = 0.85***	Heteroskedasticity	Omitted Variables
Linear regression with adjusted heteroskedasticity	Id.	Id.	Id.	Id.	Id.

Table 2: Results of linear regressions for standard deviation paid by informed consumers

	Variations with respect to a	**Variations with respect to s**	**Significativité globale de la régression**	**Heteroskedasticity test**	**Test RAMSEY**
Linear regression $std = \alpha_1 s + \alpha_2 a + const$		+ (.53)***	R2 aj. = 0.48***	Heteroskedasticity	Omitted Variables
Linear regression with adjusted heteroskedasticity		Id.	Id.		Id.

Table 3: Results of linear regressions for accepted price paid by informed consumers

	Variations with respect to a	**Variations with respect to s**	**Significativité globale de la régression**	**Heteroskedasticity test**	**Test RAMSEY**
Linear regression	- (-17.89)***	- (.79)***	R2 aj. = 0.86***	Heteroskedasticity	Omitted Variables
Linear regression with adjusted heteroskedasticity	Id.	Id.	Id.		Id.

7. Acknowledgements

We would like to thank the participants of the WEHIA (Workshop on Heterogeneous Interacting Agents) 2006 for their helpful comments.

8. References

BEGGS A. W. (2005), On the convergence of reinforcement learning, *Journal of Economic Theory*, 122:1, 1-36.

CAMERER C. and Ho T.-H. (1999), Experience-Weighted Attraction Learning in Normal Form Games, *Econometrica*, 67:4, 827-874.

DARMON E. and Waldeck R. (2005a), Comparing Nash to Adaptive Sellers, *in Advances in Cognitive Economics, NBU Series in Cognitive Science.*

DARMON E. and Waldeck R. (2005b), Convergence of Reinforcement Learning to Nash Equilibrium: a search-market experiment, *Physica A*, 355:1, EREV I. and Roth A. E. (1998), Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria, *American Economic Review*, 88:4, 848-881.

HOPKINS E. and Seymour R. (2002), The stability of price dispersion under seller and consumer learning, *International Economic Review*, 19:1, 46-76.

KIRMAN A. P. and Vriend N. J. (2001), Evolving Market Structure: An ACE Model of Price Dispersion and Loyalty, *Journal of Economic Dynamics and Control*, 25:3-4, 459-502.

POSCH M. (1997), Cycling in a Stochastic Learning Algorithm for Normal Form Games, *Journal of Evolutionary Economics*, 7:2, 193-207.

PRATT, J.W., Wise, D. and Zeckhauser,R. (1979). Price differences in almost competitive markets. Quarterly journal of economics 93, 189-211.

ROB, Raphael (1985). Equilibrium Price Distributions. Review of Economic Studies 52, 487-504.

SALOP, S and Stiglitz, J.E. (1977). Bargains and Ripoffs: A model of monopolistically competitive price dispersion. Review of Economic Studies 44, 493-510.

STIGLER G. J. (1969), The economics of Information, *Journal of Political Economy*, 69,

SUTTON R. G. (1991), *Reinforcement Learning,* Cambridge University Press.

VARIAN H. (1980), A Model of Sales, *American Economic Review*, 70, 651-9.

WALDECK R. (forth.), Search and Price Competition, *Journal of Economic Behavior and Organization*,

WALDECK R. (forth.), Prices in a shopbot Market , *Electronic Commerce Research*.