Learning To Trade

(very preliminary and incomplete)

Alan Kirman

Institut Universitaire de France,
EHESS,
Université d'Aix-Marseille lll,
University College London

GREQAM
2 rue de la Charité
13002 Marseille

Tel : 04 91 14 07 51 Fax : 04 91 90 02 27

e-mail kirman@ehess.cnrs-mrs.fr

December 1999

Abstract

In markets it is well known that individuals typically trade with very few partners. Thus markets are characterised by small trading groups. How the choice of trading partners is made is one of the fundamental problems in economics. Either those with whom one trades are given exogenously or one assumes some anonymous mechanism through which trade takes place. The purpose of this paper is to examine the conclusions drawn from different learning models in a simple market framework where individuals can choose their trading partners.

The first model attributes a great deal of rationality to the agents concerned and permits very limited conclusions. The second uses a reinforcement learning rule which can be justified theoretically and which produces stronger conclusions. The last uses Holland's "classifier system" approach and makes very limited demands on the agents. This produces the most striking results as to the type of groups that emerge. All the results are compared with the data concerning the trading relationships on the Marseille wholesale fish market.

Keywords : trading groups, learning, rationality

Introduction

In most markets trade takes place bilaterally, that is, there is no central clearing mechanism. In such markets it is then important for buyers to decide which sellers to choose and there are sellers to try and work out how many buyers were confident …for how much they demand. This paper looks at the problem in which a fixed number of sellers are faced with a fixed number of buyers and the buyers have a demand which is determined by the possibility of reselling outside the market. Sellers can be thought of as either producing the good or as buying the good from some outside source. The problem of which seller to choose is a simple one if the market takes place in one geographical location and all prices are posted. If there is no underlying problem of quality then the buyers will simply move to that seller selling the homogeneous product at the lowest price. As we know the problem is not even as simple as this since markets are full of sellers charging slightly different prices for apparently the same group. Indeed, if the origin of the fruit and the classification is checked on the appropriate container it is frequently felt to be the same and yet the selling price is slightly different. There are many possible explanations for this sort of price dispersion but the problem also arises when market is not in a physical location and sellers are dispersed. This is the case for example when one is looking for some piece of consumer hardware which is sold at various different locations and at various different prices. The traditional literature is that in which search costs a certain amount and then an optimal search … has to be worked out. This paper is more concerned with the problem in which buyers are faced with sellers who have no posted prices and a particular empirical example from which the data at the end of the paper is the wholesale fish market in Marseille. The problem might be thought of as equivalent to one of the many armed or two armed bandit problems that figure in the literature. (See for example Rothchild (1963) and various subsequent authors).

If all prices were known then one could think of the strategy for sellers as being the price they charge and the quantity they choose to supply and the strategies for buyers as which the seller has to choose. Then one could look at the equilibrium of such a problem. However, if all prices are known then the object is homogeneous. The good is homogeneous, then the solution is trivial. The problem is interesting in our case precisely because buyers do not observe prices charged by sellers and sellers have the possibility of charging different prices to different buyers. One possibility would be to think of the transactions happening with no posted prices but then all the information would then be revealed to all the buyers and sellers. However, the drawback with this is that presumably since there is no particular difference between the sellers and buyers they would all choose the same strategy based on the same constraint. In particular, this account of how things happen is not consistent with our market where buyers never observe the prices paid by other buyers and indeed empirical observation reveals that no such prices are communicated in private either. Since prices are not posted and the information is not revealed in our situation a strategy for a seller is then a price to be charged to every buyer and its total quantity to be purchased at the beginning of each period. The strategy of the buyer remains which seller to visit. Once again, if all strategies were known and the product is homogeneous then the choice for each buyer becomes trivial and there would be many equilibrium. Rather than looking for a statuate…. It is of interest to try and examine what sort of procedure agents might use to adapt their behaviour and to see whether it is convergent to some sort stable situation.

A number of papers have been written on the subject of modelling a market directly and looking at how individuals within those markets might learn which seller to choose. The closest to our particular are the papers on Bergemann and Valimaki (1996a and 1996b). Earlier papers in this spirit are those by Aghion et al (1993) and Bolton and Harris (1993) together with ……… and Harris (1994). Bergemann and Valimaki consider a situation in which there are two products which are differentiated and they are sold by joblists to either a continuum of agents or a finite number of buyers. In their case the “valuable” product is known and the patio of the other product is unknown ……. Now the problem is that firms or the firm which sells the product under the value will try and charge prices to make people experiment and indeed the firm with the other product may do the same. The buyers on the other and will also want to experiment to try and find out what the quality of the other product is. Now in the Bergemann and Valimaki set up the value of their product is determined by the match to the buyer and the seller. In our case the situation is somewhat different since what determines the value of the match according to the buyer is the price which the seller will charge so in our case the seller controls where as in Bergemann and Valimaki’s case this is exogenous. Their discussion centres around the problem of what will happen as the market evolves and they consider their actors to be Bayesian and to behave accordingly. In the Bergemann and Valimaki model at each period all buyers receive a noisy signal concerning the true value of the product with a indeterminate quality and value and in addition we receive an average or aggregate signal in which it tells them the aggregate assessment of the value of the new product. In the case where there are an infinite number of buyers of course it is sufficient for the buyers who are negligible in size to look only at the aggregate signal. Firm prices in their model are adjusted over time as a result of experience. If the buyers have a strategic effect, that is, if they have presence in the past as buyers at a particular firm has some influence over what the firm should do in the future then the situation is very complicated.

When sellers are passive the result of a model such as that have just been described is as is pointed out by Bergemann and Harris that not enough investment is made in information acquisition. This is because each buyer has a tendency to free ride on the experience of the others. This of course depends crucially on the fact that buyers observe aggregate outcomes. As some buyers become pessimistic about the firm with the indetermined values about the value of the firm with a new product then they will end to do the established firm where they are know the product for value. However, since this firm is behaving optimally it will raise its price as this occurs. Therefore, the other firm has to lower its prices in order to compensate for what is happening. All of this of course will be anticipated by the sellers. Since the buyers themselves have no they fail to take account of the externality they are generating by experimenting. They work through the whole problem. They would not experiment as much so in contrast to the Bolton and Harris model Bergemann and Valimaki find that people do too much experimenting and that firms internalise the benefits from this excess experimentation. The very severe limitations of the case which Bergemann and Valimaki are able to analyse are indicated by the fact that they have two products one of which is of known value and in most of their analysis there is very simple restrictions on it such as that if every consumer chooses one product then its quality turned out to be high where as if fewer than a full measure of consumers choose that product then its value turned out to be low. These sort of extreme restrictions are imposed in order to obtain analytic results. Although their model allows for a high level of rationality ……..individuals it is very restrictive in terms of the cases it is able to consider. That is what we have here in the situation in which a complete model with complete irrational individuals becomes analytically retractable as soon as we have try to reveal some of the restrictions.

How much rationality ?

In the full blown empirical set up individuals are aware of the strategies of the other players and of the relationship between the strategies and the payoffs that they receive. That is, we have a situation in which individuals contemplate their strategies in the light of the strategies of other people and then work out which strategy is best to adopt. One approach to looking at the dynamic set up is to allow for what is called "best response". Learning where one plays its "best response" to peoples last times choices or strategies. Another approach is that of fictitious play in which one plays a best response to the average or to the mixed strategy which is generated by the frequencies of player strategies in the past. All of this rely on some sort of comprehensive knowledge of the model. Easley and Rustichini's (1999) contribution look at a situation in which individuals do not know the full structure of the model and are simply aware of which payoffs they receive from their own actions. They then impose a set of axioms on the procedure which shows that the adaptive learning procedure of this case of this type will lead to the Nash equilibrium which would have occured had all individuals been aware of the full structure of the model. The set up is Easley and Rustichini.is that individuals choose an action and then given the realised value of the state of the system they receive a payoff. The states are identically and independantly distributed but the individual does not known this distribution therefore he observes is the pay off that he receives from his actions. ………Imposes the number of restrictions one gets to the result that I have already mentioned that such a procedure will converge to an equilibrium. This mechanism is in fact closely related to reinforcement learning procedures and in particular to the classifier system developed by John Holland. The major difference it seems to me between game theoretic models and the models which are related to those of Easley and Rustichini. is that game theoretic models are based on the notion that one anticipates what opponents will do in the future given their rational behaviour and therefore one plays as a function of this anticipation a sort of model that we developed in an early paper (Weisbuch et al 2000). We used a so called reinforcement learning process and it was interesting to see how these results compared with those that are given and Easley and Rustichini. What is the basic rule of behaviour in reinforcement learning models ? It is one that assigns weights to each of the possible actions that one might take. Learning then consists in modifying the weights on these actions as one gets more and more experience. The important analytical component to look at is the mapping from weights in one period to weights in the next. Now these weights are then in many cases mapped again into probabilities of choices. For example, the weights might themselves be the probabilities of choices in the future. They might as in other systems be such that an action with the highest weight should be chosen with probability one. Alternatively, there may be some rather more complicated mapping which associates to the weights the probabilities of taking actions. In a fully deterministic model, of course, there would be a mapping from weights into certain choices of the actions. In Weisbuch et al (2000) we choose a very simple and well known learning rule but we could have used a number of other updating rules. What we are settling for in this case is a situation in which individuals have limited rationality. They learn from their experience and the map their experience into their choices. They do this by means of updating rules and although they preserve some rationality in their choices their rationality is far below that which would be used in a fully theoretic model.

Updating Rules

Although in Weisbuch et al (2000) we choose a particular updating rule which is referred to elsewhere for example as the …….response rule and is derived from early Bush and ….. Mostella learning rules. Many other updating rules could be considered. We discuss this in some detail in Weisbuch et al (1997) and essentially the idea is that at any period the value of an action that we deduced from the payoffs to that action for the time and then its new value is determined by taking some weighted sum of the pervious values and of the latest values. Once that is done then one wants to map this rule into the probabilities of taking the appropriate actions. In our case action corresponds to choice of shop and payoff corresponds to the profit realised from choosing that ….. Of course, there are two possible outcomes one in which one obtains some profit because has been able to purchase at the shop and in the other case there is no profit since the shop had already sold out.

INSERT***

A Specific Learning Model of the Market

In the first and simplest version of the market we consider a situation in which all sellers charge the same price and in which there is no progress … that they will be sold out since they all choose to sell enough to satisfy all potential buyers. This clearly implies no rationality whatsoever on the part of the sellers but makes it basically a simple case to study and to compare with what Easley and Rustichini.would forecast for this particular set up. We therefore consider two cases in which the profit for the buyers is determined simply by the difference between the purchasing prices which is fixed and the selling price which they can obtain on their own markets. But in this case we can also think of the sellers price in some optimal way that this would be very simple to calculate if we write down the simple demand function for the buyer for his home market and then calculate from the seller the optimal price to charge for that quality. So we now have a situation in which the profit of each transaction is determined as the same model or the sellers and we can work out analytically what will happen in this case which is our learning rule. The first step in this procedure is to replace the discrete time system by a continuous time system and then to …………….. states of that. What we also do and this is very important is to replace the stochastic variable by using its expected value so this is what is called the …….. approach and simply looks at the deterministic equivalent of the underlying stochastic process.

INSERT**** WEISBUCH KIR & HERR

We seem to see in the results from our model a contradiction between what we have found and what would have been forecast by Easley and Rustichini. Why is this ? The first observation to make is that the forecast is that individuals will continue to have positive probability of going to sellers even though those sellers may not have the best price. However, there is an important feature to observe here which is that what we have done is use a deterministic approximation of a stochastic process. Therefore, we have to look at what will happen in the stochastic process itself and see whether it is consistent with the result that we found in the deterministic case. The problem here is that in the stochastic process the limit of the deterministic process lets say is .9.1 probabilities of visiting two different stores will eventually switch because with the .1 probability of visiting the other store eventually a sufficient consecutive visits to that store will occur even though this is with a low probability and then the indivdual will switch and spend his time at that store. In the very long distribution nevertheless of the process will leave of us with a positive weight on even the less advantageous store. This contrast will be Easley and Rustichini.result which suggest that in the long run the distribution will be concentrated on the better option. One is this, if one examines an example given by Easley and Rustichini the answer is clear. They assume what they call exchangeability, that is that the order in which observations occur is of no importance. This is in contrast to our use of a discount factor. In a stationary world the discounting of earlier observations makes no sense. However, the aim of our model is to provide a basis for analysing more complicated situations and in particular non stationary worlds. In such circumstances discounting earlier observations corresponds to taking account of the fact that these observations were generated by a process which is not identical to the current one.

In addition, the states of the world are being chosen by the sellers. There are thus far from being the identically distributed variables that are envisaged in the Easley and Rustichini framework. Another feature of our model is that at no time do individuals observe that prices that others have been trading at. In addition, they attribute zero value to profits from sellers whom they do not visit. This means that the value attributed to alternatives diminishes over time independently of the value that would have been realised from those alternatives had they been chosen. In our particular empirical case this corresponds to reality but there are many circumstances in which indivudals do learn something about the terms at which others trade. The important feature that emerges from our model is the bimodal distribution of loyalties. This corresponds precisely to what is observed in the empirical evidence.

Notice that in our model the level of cumulated profit will depend on the profit per transaction and the discount rate. The crucial value of β is determined by these. Thus, one would expect in the real market those individuals who have a lower discount rate and to have higher profit levels to be most loyal. In the fish market the proxy for the discount rate is the number of visits per month, (at a constant discount rate less frequent observations get discounted more). Thus, we should expect those who make large transactions and who come to the market frequently to be the most loyal. This is exactly what shows up in the empirical evidence.

More Sophisticated Sellers

In the model just described sellers are able to charge the best price to those with whom they are faced since they are aware of the buyers demand at their own selling point. However, the sellers do not make a sophisticated choice of the amount that they should sell. The simplest analysis was based on the idea that sellers provide an adequate amount to serve all their clients. In fact, sellers should choose the optimal amount given the number of buyers they expect to present themselves. A naïve rule would be to supply that quantity which would correspond to the expected number of buyers. This is not in general optimal. If the distribution of numbers of buyers is known then the optimal quantity can easily be calculated and it is given by

$$P = \frac{\exp(\beta J_{ij})}{_{j} \exp(\beta J_{ij})}$$

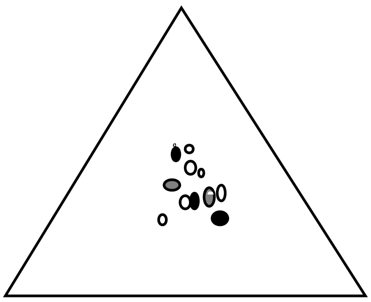


Figure 5 a

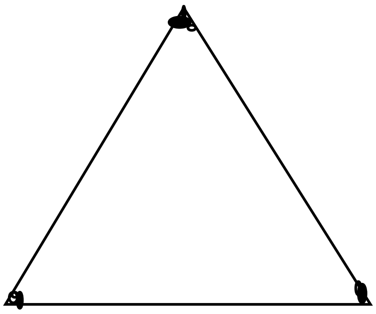


Figure 5 b

References

Aghion P and C Harris, and B Jullien (1991) :, " Optimal Learning by Experimentation ", *Review of Economic Studies,* 58, 621-654

Bergemann D and J Valimaki (1995) : " Learning and Strategic Pricing : Further Results, " Mimeo, Northwestern University and Yale University

Bergemann D and J Valimaki (1996a), " Learning and Stragegic Pricing ", *Econometrica*, Vol 64, pp 1125-1149

Bolton P and C Harris (1993) :, " Strategic Experimentation, " Discussion Paper TE/93/261, LES, London

Bush R and F Mosteller (1955): *Stochastic Models for Learning.* New York: Wiley

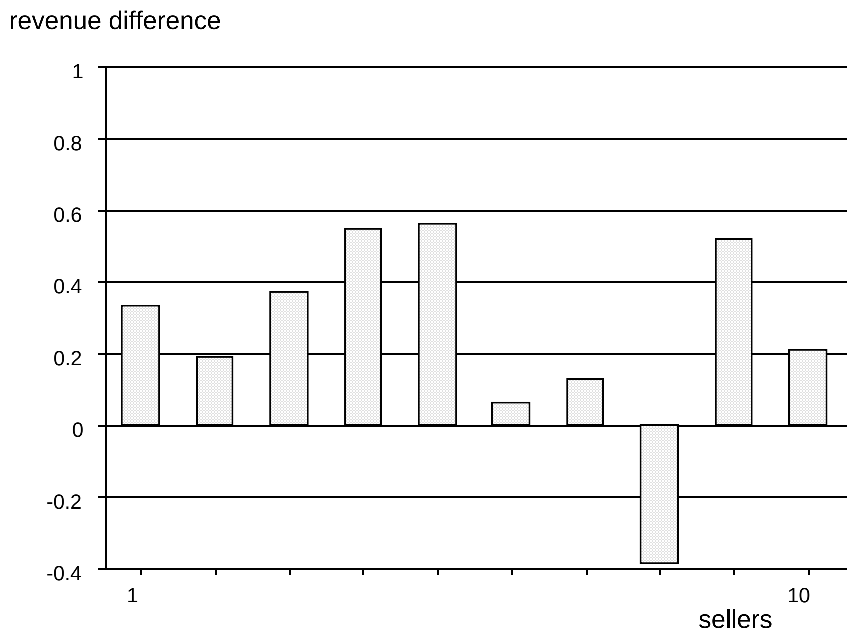
Easley D and A Rustichini (1999) , "Choice Without Beliefs" *Econometrica* Vol. pp 1157-1184

Erev I and A E Roth (1997): "On the Need for Low Rationality, Cognitive Game Theory: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," typescript

Felli L and C Harris (1994): "Job Matching, Learning and the Distribution of Surplus," Mimeo, LSE London

Kirman A and N J Vriend, (1999), "Evolving Market Structure: An ACE Model of Price Dispersion and Loyalty", Mimeo Queen Mary and Westfield College, Univ of London

Weisbuch G., A.P. Kirman, and D. Herreiner (2000), "Market Organisation and Trading Relationships" *Economic Journal* (forthcoming)


revenue difference
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
1
10
sellers