Analysing competition between seed varieties: an application to hybrid corn seed in France

Oscar Alfranca
Universitat Politècnica de Catalunya
Barcelona
Spain

Stéphane Lemarié
INRA - UMR GAEL
Université Pierre Mendès France
BP 47
38040 Grenoble Cedex 9
France
lemarie@grenoble.inra.fr

Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Montreal, Canada, July 27-30, 2003

(graph 1 and 2 can be found at the end of the document)

Introduction

The growing number of products is one of the main features of the actual competition in lot of markets. This evolution is generally linked with the differentiation of products. This aspect has been taken into account in early theoretical contributions like these of Chamberlain (1933) or Hotelling (1929). On the other hand, empirical analyses are more recent, and different problems have been or have to be solved.

The first important problem concerns the estimation of elasticity parameters. Because of the large number of differentiated products, we have generally not enough data to estimate elasticities, when using the traditional theory of demand. Two ways have been proposed to solve this problem (Foncel et al. 1996). The first way is to group some products together and to estimate only elasticity of substitution between groups. Haussman et *al*. (1994) have proposed an application by using sequential budgeting. The second way is in the line of the new theory of demand proposed by Lancaster (1966): the utility of consumers is expressed in function of products characteristics (instead of number of products consumed). Numerous applications have been done using the theory of discrete choice (Anderson et al. 1992). Compared to the traditional theory, the lack of degree of freedom is generally overcome, but some other problems appear (Berry et al. 1995).
The second problem concerns the information on products characteristics. Basically, the differentiation is completely perceived by the demand under the hypothesis of perfect information. But numerous works have shown that the diffusion of innovation is sometimes long because of users learning. Unfortunately, diffusion models do not take into account the level of differentiation[1] as differentiation models do not take into account diffusion process linked with information learning[2].
The third problem is more operational. On a period of time, introduction of innovations leads to an increase of competition, and old products are eliminated. Generally, to predict evolution of market shares, all introductions of innovation have to be know. If no important breakthrough happens, introduction of all new products may be summarized with synthetic variables (number of products, performance of the leading products, etc.). In that case, predictions need only to anticipate evolutions of these synthetic parameters.

The purpose is illustrated here with a study of the French market of hybrid corn seeds (1978-1994). In the section 1, we present the main features of this case, and the set of data used thereafter. Some descriptive statistics are presented to show that this is an exemplary case to illustrate both diffusion and increase of competition effects. The basic model is presented, with

[1] A clear distinction has to be made between the two main features of diffusion process (only the first one will be highlighted in this paper): (i) learning of users which give them more information directly by testing products or indirectly by learning of other users, (ii) incentives on the supply side to improve the products and lower production costs when demand is growing, which accelerate diffusion. The second aspect can be treated with a standard model of differentiation because changes in prices of products and characteristics are explicitly taken into account (see Trajtenberg (1989) for an application to the Scanner).
[2] Recently some authors have mixed the two effects, but these contributions are generally theoretical (Bergemann et al. 1996; de Palma et al. 1998)

its results, in section 2. This model is an application of the theory of discrete choices, and includes diffusion effect as with additional dummies. Two extensions are tested in section 3, and a more general discussion on the combination of diffusion and differentiation effect is addressed in conclusion.

1. The French market of hybrid corn seeds

1.1. Main features

- Supply side

In this sector, the property rights (breeder's right) give a temporary monopoly to the creator of the variety (i.e. the product), with the constraint that genetic contains has to be stable. For that reason, each improvement of product characteristics leads to a new product. Breeders can sale the license to other firm, but this practice is so unusual in France that each product can be consider as sold by only one firm[3]. No important progress has been made on the production process, so production cost can be considered as stable. Note also that sales to farmers are made by local distributors. Their effect will not be considered in the basic model, but will be discuss in the extensions of the models.

- Demand side

Corn is used for silage production to feed animals in the farm, or grain production sold out to an elevator. In each case, the lag between sowing and harvesting is about six month. Climate variation during this time can affect yield, and make the choice of the farmer uncertain. The effect on yield level is such that products ranking can change from year to year. This has two main implications:

- Farmers share risk among different varieties[4]. Sharing is generally more important when the differences of yield between products is low.
- Farmers take more risks when using young products, because these products have not been tested in a large range of climates. The analysis of market share have to take into account the diffusion effect at the beginning of the life cycle. The process first lies on learning of technical value by leader farmers. Then, the choice of leaders spread to followers. These two basic aspect of diffusion has been previously described in numerous contributions (Griliches 1957; Feder et al. 1982; Jensen 1982).

- Different forms of differentiation

The differentiation of products in this market is both horizontal and vertical. For a given crop, horizontal differentiation is linked with the adaptation to the geographic areas. Basically, farmers from the north of France must use early varieties, while farmers from the south can use late varieties. Extension services (AGPM in the case of corn) make comparison of varieties within group, each earliness group corresponding to a particular area in France. Seven groups of corn

---

[3] This assumption should not be done for the American market where organization of sales is different. But we will not consider this case here. In France, the only case where license appears frequently is when a firm is a foreign breeder, and sales are done by an exclusive representative in France (e.g.: RAGT is the exclusive representative of Dekalb). Note also that generally the representative does not sale variety from other breeders.

[4] The number of products among which risk is shared is limited for technical reason. Farmers do not use generally more than 5 products for a given crop.

seeds are distinguished in France by the AGPM, and they are numbered from 10 (early, adapted to the North) to 16 (late, adapted to the South). Generally, firms try to sale a minimum of one variety by earliness group. Note also that in the circumstance of the French market, early varieties are used for silage production while late varieties are used for grain production.

As horizontal differentiation is based on the earliness characteristic, vertical differentiation is based on other technical characteristics which correspond to the yield or to the resistance to diseases. Each new product is registered on an official catalogue if its yield is better than the reference level estimated from a small sample of current most sold varieties. This official catalogue is a powerful tool to promote new better products and the market share of off-catalogue products is negligible[5]. For this study, the technical value of seeds can be reasonably summarized by the yield level because: (i) it is a synthetic variable which capture a part of the effect of other characteristics as resistance to diseases, (ii) the productivist model promoted in France incited farmers to choose seeds with high yield level. The use of corn production do not affect the main characteristic: yield is measured on the all plant for the silage, while it is measure only on grain production in the other case. Vertical differentiation can easily be checked by the fact that, within an earliness group, varieties with high yield are more expensive.

Basically, the farmers choose between all the varieties adapted to his geographic area. Competition between products occurs first of all within earliness groups. Nevertheless, at the frontier between two areas of adaptation, farmer may choose between products of different groups. Finally, the substitution within earliness groups is higher than the substitution between earliness groups, but this last one is not negligible.

1.2. Data used in this study

For each variety registered in the French and the European catalogues, four types of stable data are available for each variety: the date of release, the earliness group (from 10 to 16), the breeder(s) which owns the property right, and the repesentant(s). This data comes from the bulletin of varieties jointly distributed by the CTPS and AGPM.

For all these varieties, annual sales in France from 1976 to 1994 were provided by the GNIS. The year 1988 was eliminated because of too much missing data. Unfortunately, no data on prices were available. Data on technical characteristics are freely available to farmers. For each variety, the CTPS provides yield estimation (form trial tests) during two years before release and AGPM provides estimation for the two or three following years. This set of technical characteristics is very interesting because it reflects the general assessment of farmers on product. Moreover, as CTPS and AGPM are extension services whose objective is to provide the best informations to farmers, we can reasonnably consider this information as good indicator of farmers' choices.

---

[5] This catalogue does not exist in United States, and this is the main reason why organisation of sales is quite different.

1.3. Evidences on the diffusion and competition effects

We can start with a simple descriptive model where market share of the product *i* at the *t th* year after its release is decompose as follow:

$$S_{it} = A_i \cdot B_t \qquad (1)$$

Despite its very rough form, this model gives interesting results[6]. Estimates of *B* are reported in graph 1, and other results in table 1. Estimations of *B* follow a very general pattern, first increasing and then decreasing.

Table 1. Results from the regression on the diffusion effect

Groupe	10	11	12	13	14	15	16
Nb obs.	986	1194	575	378	397	480	228
Nb. par.	193	242	144	74	94	98	54
R^2	0.751	0.772	0.810	0.809	0.801	0.763	0.784

The first years correspond to a diffusion period. Basically, the yield is the major technical characteristic of crops, because it influences profit of farmers. However, because of climate and farming changes from one year to another, the yield of one variety can fluctuate. For that reason, it takes several years for farmers to have a good knowledge of products' performance[7].

The decrease of *B* reveals the increase of competition. Two figures are very illustrative of this evolution (see table 2):

- For the whole French market, there were 175 products in the late 70's and 550 products in the middle 90's. The number of products has been multiplied by three for all the groups, with some differences between early groups where this number has increased tenfold, and late groups where it has doubled. The multitude of product is both the result of an increase of the number of product proposed by each firm, and an increase of the number of firm on the market.
- Between 1976 and 1994, yield growth varies between 20% and 35% depending on the segment, which is very large compare to the 1-3% mean progress that each product brings. It means that during this two decades, the list of products has been completely renewed one or several times.

[6] Estimation has been made on the log of the market share.

[7] This aspect was taken into account in Griliches (1957) seminal study on diffusion of hybrid corn. As he explained after (Griliches 1980), the diffusion of hybrid corn was not only due to revelation of informations, but also to hybrid improvment between the 30's and the 50's. In our study, diffusion is only due to revelation of informations because data are studied at the variety level, for which the characteristics are stables.

Table 2. Innovation in the different earliness groups

Group	Main use	Number of seed varieties				Annual yield increase
		80	85	90	95	
10	E	27	55	108	148	24%
11	E	53	74	113	150	21%
12	E/G	17	27	54	107	32%
13	E/G	16	28	36	29	18%
14	G	25	26	33	46	38%
15	G	29	36	35	41	25%
16	G	7	20	26	24	33%
		174	267	405	543	

2. The basic model

2.1. The basic model

As described in the previous section, sales are determined by two main factors: the yield (x) which reflects the differentiation level, and the level of knowledge on product characteristics (b). More formally, the utility of the variety i for the farmer j at the year t is decomposed as follow:

$$u_{ijt} = x_{it-1} + b_{it} + e_{ijt} \qquad (2)$$

Farmer's utility is a direct function of his profit. Because payoff is the product of yield by crop price, it is logical to consider the utility as a linear function of yield. To represent learning on product characteristics, x is defined as an estimation of the yield, and it can change from year to year. We introduce a lag of one period, because the choice at t depends on all the information revealed until the previous period t-1. The level of information is supposed to be identical for all the farmers, and the variation of estimation and/or information level among farmers is represented by the term e_{ijt} which expresses the specificity of choices.

We consider that the knowledge acquisition depends more on the context (extension system, farmers network, etc.) than on the variety. For that reason, we suppose that b is identical for all the varieties with the same age. In other words, if T_i is the date of release of the variety i, the equation (2) can be modified in the following way:

$$u_{ijt} = x_{it-1} + b_{it-T_i} + e_{ijt} \qquad (3)$$

After comparing of the different varieties within a given earliness group, we suppose that the farmer chooses the best one. Like most of the traditional models, we make the hypothesis that the specificity term is independently and identically distributed (iid) following a double exponential distribution. The market share of the product i within the earliness group I follows a multinomial logit[8]:

[8] See Anderson et *al.* (1992) for a complete treatment.

$$s_{it} = \frac{\exp\left(a \cdot x_{it-1} + b_{it-T_i}\right)}{\sum_{i \in I} \exp\left(a \cdot x_{i't-1} + b_{i't-T_{i'}}\right)} \quad (4)$$

The parameter a_t is defined in such a way that the denominator of equation (4) is equal to exp(a_t). It leads to the final form of the basic model:

$$s_{it} = \frac{\exp\left(a \cdot x_{it-1} + b_{it-T_i}\right)}{\exp\left(a_t\right)} \Leftrightarrow \ln\left(s_{it}\right) = a \cdot x_{it-1} + b_{it-T_i} - a_t \quad (5)$$

a is a parameter to be estimated which reflects the sensitivity of farmers to the difference of yield (a is expected to be positive). This sensitivity can change depending on three factors:

1. For any farmer in the area of an earliness group, the difference of yield is more or less easy to estimate. For the case studied here, it corresponds to the difference between areas oriented toward silage production, and areas oriented toward grain production. Grain production is sold, and for that reason the farmer knows precisely the quantity he produced. Conversely, silage production is used inside the farm for livestock feeding and the quantity can only be roughly estimated. For that reason, we expect a to be higher in grain production oriented areas (group 12 to 16).
2. The aggregate sensitiveness of farmers to differences in yield is higher when the choices of farmers inside the area are homogeneous. The ranking of variety can change from place to place inside an area of adaptation because of micro-adaptations not captured by the differentiation of earliness groups. Moreover, farming habits may also change from place to place, and this can influence the ranking. Such a variation is captured in the model by e_{ijt} in equation (2) and (3). High variance of e expresses very heterogeneous choice and, by construction, it leads to small level of a.
3. As discussed before, no reliable data on prices are available. Generally, models of differentiated products consider price as an endogenous variable: the demand is first expressed as a function of price and technical characteristics; then, Nash equilibrium on price is established on the supply side. Our model which does not take prices into account is equivalent to a model where price, at the equilibrium, is a linear function of yield. In other words, the parameter a captures both the sensitivity to yield differences and the sensitivity to price differences. Because the effect of price on sales is generally negative, the higher the sensitivity of farmers to prices differences is, the lower is a.

a_t is a synthetic estimated indicator of competition level. The higher a_t is, the higher the competition level is and the lower the market share is (*ceteris paribus*). a_t can be explained as a function of descriptive statistics, in order to have more simple forecast. exp(a_t) has first been defined as a the sum of utility of the products, which is also the number of product times the average utility. If we suppose that the distribution of product yield and product ages are stable, then we can eliminate the effect of b and replace it by a constant m.

$$\sum_{i \in I} \exp\left(a \cdot x_{i't-1} + b_{i't-T_{i'}}\right) \simeq m + \Omega\left(I_t\right) \cdot \exp\left(a \cdot x_{0t}\right) \quad (6)$$

After more development, we obtain:

$$a_t = \ln\left(\sum_{i'\in I} \dots\right) = c_0 + c_1 \cdot \ln\left(\Omega\left(I_t\right)\right) + c_2 \cdot x_{0t} \quad (7)$$

For convenient reason, we introduce the constraint that b is negative. Because we expect b to increase, the constraint is introduced by supposing that b is nil when the variety is older than a certain level (different level will be tested). With this constraints, exp(b) increases within the interval [0,1]. For that reason, exp(b) can be interpreted as a coefficient of diffusion, and it is generally accepted that it follows an S-shaped function (Griliches 1957; Karshenas et al. 1995). Some simple dynamic property of market share can be analysed by supposing that exp(b) follows a logistic function[9]:

$$\frac{\mathrm{d}\exp(b)}{\mathrm{d}t} = b \cdot \left(1-\exp(b)\right)\exp(b) \quad \Rightarrow \quad \mathrm{d}b = b \cdot \left(1-\exp(b)\right) \cdot \mathrm{d}t \quad (8)$$

By introducing that in equation (5), we obtain[10]:

$$\frac{\mathrm{d}\ln(s_{it})}{\mathrm{d}t} = \frac{\mathrm{d}b}{\mathrm{d}t} - \frac{\mathrm{d}a}{\mathrm{d}t} = \mathrm{d}b = b \cdot \left(1-\exp(b)\right) \cdot \mathrm{d}t - \frac{\mathrm{d}a}{\mathrm{d}t} \quad (9)$$

The following property can finally be observed :

$$\frac{\mathrm{d}s_{it}}{\mathrm{d}t} > 0 \quad \Leftrightarrow \quad b \cdot \left(1-\exp(b)\right) \cdot \mathrm{d}t > \frac{\mathrm{d}a}{\mathrm{d}t} \quad (10)$$

For a given increase of competition, only young products with low values of b have increasing market share. After a while, market shares will decrease because b is too high. Moreover, we can see that the faster the competition increases, the younger the product will turn to decreasing market share.

Substitution properties have also to be discussed. One of the main critics that have been addressed to the multinomial logit model lies on the property that when a new product is introduced, the market shares of other products decrease homothetically. Berry et *al.* (1995) illustrate this with the following example: supposing that a Lada has the same market share than a Mercedes, then the multinomial logit forecasts that the introduction of a BMW decreases the market share of both the Lada and the Mercedes by the same amount. However these critics has less effect when we apply the model inside a group of products where the substitution is more regular. We turn then to a nested logit (Trajtenberg 1989; Berry 1994). Our application corresponds to this last case because we have distinguished different earliness groups. Actually, the strong assumption of the model does not concern substitution within a group, but rather substitution between products of one group and external alternatives. The first external alternative is the product of the neighbouring group. We have seen that farmers on the frontier of two areas of adaptation may choose between products of different groups. This effect is generally taken into account in nested logit by compiling a matrix of substitution between groups. This has not been done yet, and we implicitly suppose that that substitution between groups is always nil. The other crops represent the second type of external alternative. Depending on the relative prices of

---

9 Logistic function is defined as follow: $\exp(u) = \dfrac{1}{1-\exp(a-b\cdot u)}$

10 x_{it} is suppose to be constant over time for simplicity.

products (corn, wheat, sugar beat, etc.), some markets could increase at the expend of others. This treatment as not been done yet also. Even though substitution hypothesis seems to be strong, it has to be noted that we try to explain market share instead of sales. In other words, total sales of one earliness group may increase or decrease, the model only suppose that market shares of all the products are still the same.

2.2. Results with the basic model

The data used for this study has been described before (see section 1). We do not present the result for the group 16, because the corresponding area covers the south of France and the north of Italy and/or Spain (the market shares in France is not representative of market share for all the area). The methodology for calculating estimation of x_{it} is described in details in the appendix A. Remember also that the regression has been made with the constraint that b is nil after a while. Different periods have been used for this limit. If we choose a long time (for example 10 years), estimation of b first increases, taking even positive values, and then decreases. The decrease is similar to the one that has observed in the descriptive model (see section 1), and represents increase of competition which is also capture by a_t in the basic model. To eliminate the redundancy, and to have tractable values of b we have taken the maximum length of time which still respects the hypothesis of increasing estimation of b.

Table 3. Results with the basic model

	10	11	12	13	14	15
Nb obs	986	1194	579	378	397	480
Nb par.	22	22	23	21	21	22
a	0.075	0.083	0.107	0.125	0.091	0.131
b_0	-2.463	-2.418	-3.001	-2.499	-2.231	-2.693
b_1	-1.286	-1.586	-2.191	-1.269	-0.677	-1.443
b_2	-0.427	-0.700	-1.403	-0.724	ns -0.219	-0.552
b_3	ns -0.024	-0.275	-1.044	ns -0.162	0.000	* -0.427
b_4	0.000	0.000	-0.765	0.000		0.000
b_5			0.000			
R^2	0.322	0.314	0.463	0.263	0.251	0.348
Auto.	0.641	0.717	0.733	0.661	0.772	0.687

*Estimated parameters are always significantly different from 0 at the level of 5%, except for * where the level is 10%, and `ns' where it is non-significant at the level of 10%. `Auto' means autocorrelations, and is measured by the R^2 of the model explaining the residual at t as a function of the residual at t-1.*

The results are given in table 3 and graph 2. The estimation of the parameters have expected sign: a is significantly positive, b is negative and significant for the first years, a_t increases over time. Despite these expected signs, the model lets important residuals: R^2 varies from 0.25 to 0.50. This drawback has to be qualified because the number of parameters is no more than 7% of the total number of observations. Such a weighting is made with a test of Fisher, and then the model cannot ever be rejected. Residual are auto-correlated: for a given variety, the market share estimated by the model is either under-estimated or over-estimated, but rarely both. Because of the large number of degrees of freedom which are still available, there are probably some ways to improve the explicative power of the model.

From one group to the others, the model reflects correctly the variation of sensitivity to differences in performance. The estimation of *a* are lower for the north of France where the main use is silage, and sensitivity to price is greater.

The evolution of competition has been explained by the number of products and the average performances (see equation (7) for the model and table 4 for results). The model has a good explicative power (R^2 always superior to 0.75), and the effects are positives (when they are significant) as expected.
For the earliest groups (10 and 11), increase of yield does not explain evolution of competition, while its effect is significantly positive for later groups (13 to 15). For these last case, it is interesting to note that estimation of c_2 are closed to estimation of *a* in table 3: a unitary progress on the yield (x_{it}) has the same effect on sales than a drop of one unit of the average yield (x_{0t}).

For the analysis of the effect of the number of products, we have to consider separately the case of groups 10 to 12 where this number is tenfolded, and the groups 13 to 15 where it double. Comparison of group 11 and 12 shows clearly the opposition of using: the increase of competition is explained only by the number of products for the group 11, and only by the genetic progress for the group 12. The group 10 is closer to the group 11, even if the effects are not significant. For the groups 13 to 16, the effect of the number of products is generally significant but, compared to the effect of the average yield, it contributes to a less extend to an increase of competition.

Table 4: Analysis of evolution of competition (a_t)

	10	11	12	13	14	15
c_0	ns -3.743	ns -11.703	-19.492	-17.335	-19.180	-19.364
c_1	ns 0.036	0.068	ns 0.026	ns 0.032	0.022	0.109
c_2	ns -0.004	ns -0.172	0.131	0.140	0.136	0.122
R^2	0.703	0.753	0.883	0.860	0.858	0.973

3. Extension of the basic model

3.1. Hypothesis H1: the accident effect

- Presentation

One of the original facts in our model lies on the idea that performance is an estimation that can change from year to year depending on the new revealed information. Suppose that the yield in t is much greater than the estimation at the end of the year t-1 (x_{it-1}), then we will have $x_{it}<x_{it-1}$, and the model will predict a decrease of market share (*ceteris paribus*).
We want to test here a possible improvement the predictive power of the model in case of very bad revealed information for young varieties. This phenomena has been suggested by experts: "we know that if a product shows some problems during the first years, we will never be able to sell it after, even if we show that this bad event was quite unusual". This phenomenon can be interpreted by saying that the interaction between the climate and the variety is interpreted differently depending on the age of the variety. In more illustrative terms, if a young product behave badly it means that it is bad, but if the same bad information is revealed when the product is older, it means that the year was bad for this product.

This hypothesis is tested with a dummy variable q_{it} which default value is 0, and take the value 1 if the estimation falls more than y percent, and the variety is less than ΔT years old. More formally:

$$if \begin{cases} x_{it} < (1-y)\cdot x_{it-1} \\ t < T_i + \Delta T \end{cases} \qquad \text{then} \quad q_{it'} = 1, \forall t' > t \tag{11}$$

The effect of this dummy is then added to the basic model (equation (5)), and the accident effect is defined by a parameter d to be estimated:

$$\ln(s_{it}) = a\cdot x_{it-1} + b_{it-T_i} - a_t + d\cdot q_{it} \tag{12}$$

Generally, fall of x_{it} does not exceed 10%, and y have to be less than 5% to have sufficient frequency of accident. Two other values of y are also considered (3% and 1%). The effect of ΔT can not be tested here because technical values are only available during the 4 first years, and x_{it} is stable thereafter. In other words, because of the data used here, accident can only be identified during the first three years, so it is not necessary to defined limit age.

- Results

Comparing to the basic model, R^2 increases a little bit (no more than 5%), The sign of d is always negative (as expected) but the significance varies, depending on the group and the value of y (table 5). Before analysing the values, we need to have some marks on the importance of this effect in terms of market share. We can translate the effect of d by an equivalent fall of yield of d/a (because the effect of the yield is multiplied by a). Note that such a simple translation can be made because the estimations of other parameters are similar in the two models (equation (5) and (13)). If we take *q*=-0.5 and *a*=0.1, the equivalent fall of yield is equal to 5 points, which is important compared to the maximum differences of yield between varieties at one time (between 5 and 10 points). The accident effect is such that it cancels all the competitive power brought by the increase of performance of one product (the experts' expression given at the beginning of the paragraph is confirmed).

Accidents appear to be more frequent in the earliest group. This can be explained by the more constraining climate in the North of France. As we can expect, the estimations of *d* is more important when y is higher. For groups 13 to 15, *d* is never significant with y equal to 5% and 3% because accidents are not frequent enough. *d* becomes significant with y equals to 1%, and in that case, the estimations are more important than for the groups 10 to 12. Finally information has to be more unfavourable in the North of France than in the South to have some effects on sales. This is probably because information are more frequently unfavourable in the North. However, when it happens, the effect is such that the variety is eliminated from the market.

Table 5. Test of the "accident effect" (H1)

		10	11	12	13	14	15
y=5%	Freq.	11%	6%	4%	3%	3%	5%
	d	-0.434	-0.674	-1.067	* -0.779	ns -0.171	* -0.517
	R^2	0.330	0.326	0.481	0.269	0.251	0.352
y=3%	Freq.	25%	18%	13%	9%	4%	10%
	d	-0.362	-0.574	-0.561	ns -0.075	ns -0.500	ns -0.033
	R^2	0.333	0.335	0.474	0.263	0.254	0.348
y=1%	Freq.	53%	50%	50%	45%	18%	31%
	d	-0.205	-0.238	-0.330	-0.535	-0.946	-0.513
	R^2	0.326	0.319	0.470	0.282	0.296	0.364

3.2. Hypothesis H2: the firm effect

- Presentation

In the basic model, and even with the H1 extension, two products with the same release date and the same revealed information have the same sales during all their life cycle. We consider here that, depending on the type of firm which sales it, the market share can be different. The hypothesis H2 can be justified by three arguments:

- A firm with large sales has generally a good reputation, because farmers know the products that it has proposed before[11]. Because of the parental link between products proposed by the same firm, the information on old good products influences favourably estimation of the value of new products of the same firm.

[11] Note that the two hypothesis that are necessary for the emergence of reputation effect in game theory are present here: (i) uncertainty on gains, (ii) repeted games (Shapiro 1983).

- The firm with a large market share has generally more powerful and more adequate means to promote its products (experience of marketing division, advertising budget, etc.).
- Such a firm is more powerful in front of distributors, because its products represent a larger share of distributors activity.

The total market share of the firm on the market appears to be a good criterion. It can be compiled easily by making a sum of the sales of all the firm's products (remember that each product is most of the time produced by one firm). Because of the similarity between some groups, we have compiled the total sales of each firm for the North of France (group 10 to 13) and for the South of France (group 14 to 16). Firms have been classified in four categories which limits are 1%, 5% and 10% (see table 6). A given firm may change of category if its total market share increases or decreases[12]. The category 0 corresponds to the smallest firms, and the constraint of no firm effect has been introduced in that case. Three dummy variables have then been defined: γ_{ikt} ($k \in \{1,2,3\}$) which default value is 0, and takes the value 1 if the variety i is sold by a firm of category k at time t. The new model to test is then:

$$\ln\left(s_{it}\right) = a \cdot x_{it-1} + b_{it-T_i} - a_t + \sum_{k=1}^{3} e_k \cdot g_{ikt} \qquad (14)$$

Table 6. Definition and size of categories of firms

Category (k)	Definition (market share on the area)	Number of seed companies					
		Northern area (gp 10 to 13)			Southern area (gp 14 to 16)		
		Mini	Maxi	Moy	Mini	Maxi	Moy
0	< 1%	19	25	21.1	11	16	12.4
1	1% à 5%	3	11	7.7	1	8	6.0
2	5% à 10%	1	5	2.7	1	4	2.3
3	> 10%	1	4	2.5	1	4	2.4

- Results

Compared to the basic model, the introduction of the firm effect improves substantially the R^2 (table 7). The estimations of firm category effects are positive and significant (at least for the largest firms). The hypothesis H2 can not be rejected. The firm effect captures a part of the yield effect, because estimation of a are lower here than in the basic model. Nevertheless, estimations of the other parameters are similar.

To go farther in the analysis, it is useful to have a clear idea of this firm effect in terms of sales (as we have done before for hypothesis H1). For that, the firm effect is translated in an equivalent supplement of performance in the basic model, in order to predict the same sales. The calculus is more complicated here because the estimation of a has changed. An adjacent calculus enables us to estimate this equivalent supplement of yield as at least 5 points for the largest firms. In other

[12] Some adjustments have been made by hand in order to maintain an inertia and avoid, for example, a change from category 2 to 1 followed just after by a change from category 2 to 1.

words, the firm effect appears as important as the effect of the technical progress brought by a new variety.

As we have seen in the beginning of this paragraph three main arguments can explained the origin of the firm effect. Unfortunately, the data set used here does not permit to differentiate those results. Experts generally agree to say that promotion of products is generally more important in early groups (10 and 11) because the technical differentiation is harder (cf. the lower estimation of *a*). Conversely, reputation seems to have more effect in the case of late groups oriented toward grain production, because the most implanted firms have proposed better products for more than 10 years. The relationship with distributors is generally independent of the group, but depends more on the local context (distributors have a power if they are in monopoly faced to farmers).

Table 7. Test of the "firm effect" (H2)

	10	11	12	13	14	15
a	0.077	0.076	0.091	0.115	0.061	0.108
e_1 (1%-5%)	0.846	0.378	ns 0.036	ns 0.240	0.781	0.669
e_2 (5%-10%)	0.744	0.520	0.585	ns -0.010	1.235	0.864
e_3 (> 10%)	1.201	1.166	0.732	* 0.994	2.299	1.747
R^2	0.346	0.373	0.494	0.304	0.450	0.447

Conclusion

For the last 10-20 years, there was an increasing concern in the economic literature on the analysis of differentiated markets. This paper has tried to fill some inadequacy, and to test new propositions with an application to the case of the french market of hybrid corn. It can be summarized by three main ideas:

1) When a new product includes technical innovations, it experiences a diffusion period during the first years after its release. Learning of demand, which is here the main explanation of diffusion, lies on two phenomena: individual learning and mimetism. This was include in the basic model first by considering technical characteristics as estimations[13], and second by including age effect in a kind of diffusion coefficient. Extension of the model show also that if a very bad information is revealed on a young variety, then diffusion is stopped even if it can be shown thereafter that this accident was quite unusual.

2) Two identical products reach different market shares depending on the firm which sales it. This effect corresponds somehow to unobservable characteristics of products (for example, the trademark is one aspect of the style of the car in the works of Berry et *al.* (1995)). Nevertheless, the firm effect is different in the two kind of analysis. When it is considered as unobservable characteristics of products', the aim is to solve problems of estimation with endogeneity of prices

[13] Some theoretical works has also consider technical characteristics as imperfect estimations in models differentiation. For example, Bergemann and Välimäki (2001) consider competition between two products (one old known product and one unknown new product), and analyse pricing policy implications.

(non independent residual). Here, the firm effect (based on reputation, advertising or relationship with distributor) makes diffusion easier.
3) Innovation is a major aspect of the evolution of competition, measured here with a synthetic variable. In most early groups, differentiation is more difficult and competition increases mainly because of increasing number of products. In late groups, competition is more sensitive to technical progress because differentiation is easier.

The model used here is very simple, but it introduces several constraints. Substitutability constraints are managed by analysing independently earliness groups. But the analysis of substitution between groups has to be done, to consider that we are using a nested logit model. Because prices where unfortunately unavailable for all products, we have to suppose that they are linear function of yield (at the equilibrium). The explicative power of this model can be improve because we have observed autocorrelation of residual, and a large number of degree of freedom are still available. In the addition to the better management of constraint described just before, three extension of the model are conceivable.

Some other technical characteristics can influence the choice of farmer. Here, we ha ve supposed that yield is a synthetic variable which include these effects. But it is imperfectly true, and perhaps some improvements are possible if we consider also the effect of resistant to diseases, or resistance to lodging.
A significant improvement should be done also by taking the network of distribution into account. In France, distributors have probably an important effect, as well as the range of products proposed by each firm. This enhancement needs some preliminary theoretical exploration on the strategic equilibrium between seeds companies and distributors. This is a part of the agenda proposed by some author (Foncel et al. 1996). For the case of hybrid corn seeds, it would be very interesting to compare the case of France and United States whose networks of distribution are quite different.
Diffusion can also be treated in a very different way. Based on Arhur's models of increasing return to adoption (Arthur 1989), De Palma et *al*. (1998) has proposed another way to take diffusion into account in the theory of discrete choice. A new formulation of indirect utility is proposed: $u_{it} = x_{it} - 1 + r \cdot s_{it-1}$. Utility is both a function of technical characteristics and market share at the last period. The larger is sales, the larger is the amount of information available on the product, and this is a source of competitive advantage, as well as technical improvement. With such a model, we may found some lock in effect as in Arthur's model. One important exploration would be to test the permanence of this effect.

References

Anderson, S. P., A. De Palma et J.-F. Thisse (1992). Discrete choice theory of product differentiation. Cambridge, London, MIT Press.

Arthur, W. B. (1989). “Competing technologies, increasing returns, and lock-in by historical events.” Economic Journal **99**(934): 116-131.

Bergemann, D. et J. Välimäki (1996). Market diffusion with two-sided learning. New Haven, Yale University**:** 39 p.

Bergemann, D. et J. Välimäki (2001). Entry and vertical differentiation. Yale, Yale University, Cowles Foundation for Research in Economics**:** 41 p.

Berry, S. (1994). “Estimating discrete choice models of product differenciation.” Rand Journal of Economics **25**(2): 242-262.

Berry, S., J. Levinsohn et A. Pakes (1995). “Automobile prices in market equilibrium.” Econometrica **63**(4): 841-890.

Chamberlain, E. (1933). The theory of monopolistic competition. Cambridge, Harvard University Press.

de Palma, A., K. Kilani et J. Lesourne (1998). How network externalities affect product variety. Advances in self-organization and evolutionary economics. J. Lesournes and A. Orléan. London, Economica**:** 57-76.

Feder, G. et G. T. O'Mara (1982). “On information and innovation diffusion: a Bayesian approach.” American Journal of Agricultural Economics **64**(1): 145-147.

Foncel, J. et M. Ivaldi (1996). “Econométrie de la concurrence imparfaite sur les marchés à produits différenciés.” Revue Economique **47**(3): 477-486.

Griliches, Z. (1957). “Hybrid corn : an exploration in the economics of technical change.” Econometrica **25**(4): 501-522.

Griliches, Z. (1980). “Hybrid corn revisited : a reply.” Econometrica **48**(6): 1463-1465.

Hausman, J., G. Leonard et J. D. Zona (1994). “Competitive analysis with differenciated products.” Annales d'Economie et de Statistique **34**: 159-180.

Hotelling, H. (1929). “Stability in competition.” Economic Journal **39**: 41-57.

Jensen, R. (1982). “Adoption and diffusion of an innovation of uncertain profitability.” Journal of Economic Theory **27**(1): 182-193.

Karshenas, M. et P. Stoneman (1995). Technological diffusion. Handbook of the economics of innovation and technical change. P. Stoneman. Oxford, Cambridge, Blackwell Publisher**:** 265-297.

Lancaster, K. J. (1966). “A new approach to consumer theory.” Journal of Political Economy **74**: 132-157.

Shapiro, C. (1983). “Premiums for high quality products as rents to reputation.” Quarterly Journal of Economics **98**: 659-679.

Trajtenberg, M. (1989). “The welfare analysis of product innovation, with an application to computed tomography scanner.” Journal of Political Economy **97**(2): 444-479.

A. Calculus of technical performance x_{it}

Gross data are published by trial. One trial gives results on 10 to 20 varieties from the same group. One trial reports the average performance of repeated plot in a geographic area for which the products appear to be adapted. Different technical characteristics are reported, but we have only considered the yield.

x_{it} has been compiled by progressive adjustments, as for a bayesian process (see Feder and O'Mara (1982) for a use of this process in case of a new variety diffusion). Suppose that we are at period t, and that the values x_{it-1} are known for all the varieties i. Three steps are necessary to compile x_{it}.

- Revealed performance

Basically, the performance measured in a trial (phenotypic value) includes a year effect, which has to be eliminated in order to have coherent values from year to years. The revealed performance is defined relatively to some reference varieties, which are explicitly defined in each trial. Generally, these references correspond to the most used varieties, and they are between 5 and 10 years old. We defined y_{itn} as the yield measured for the product i at year t in trial n. J is the set of reference varieties for the same trial. The revealed performance is defined by z_{itn} and is compiled in the following way:

$$z_{itn} = \frac{1}{\Omega\left(J\right)} \sum_{j \in J} \frac{y_{itn}}{y_{jtn}} \cdot x_{jt-1} \tag{15}$$

- Mean revealed performance

$\overline{z}_{it}$ is the mean of z_{itn} for all the trials n where the variety has been tested.

- Adjustment of estimation

Estimation at the end of year t is defined as the average of all the mean revealed performances until t (including t). The first revealed performance is in T_i-1, just before the release. Formally, estimation is defined as follow:

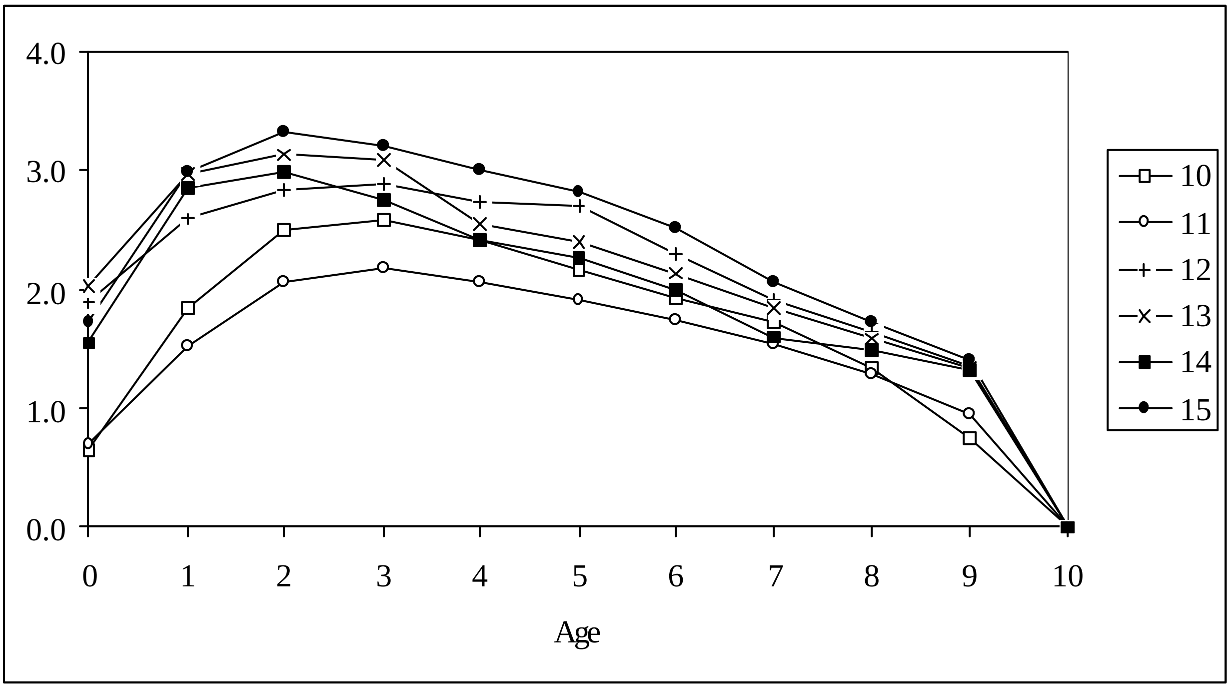
$$\begin{aligned} x_{it} &= \frac{1}{2+t-T_i} \cdot \sum_{u=T_i-1}^{t} \overline{z}_{iu} \\ &= \frac{1}{2+t-T_i} \cdot \left[\left(1+t-T_i\right) \cdot x_{it-1} + \overline{z}_{it}\right] \end{aligned} \tag{16}$$

This formula is applied while the variety is tested. When it is not tested at all, we have no value for $\overline{z}_{it}$, and the estimation are then stable (x_{it}=x_{it-1}).

All the published data of CTPS and AGPM from 1975 to 1994 have been used. For the first year, no adjustment of estimation can be made because x_{i1974} were unknown. For each group, we

suppose that x_{i1974}=100 for the most diffused variety. Starting from this value, estimations of all the following varieties released until the mid 90's are ranged from 100 to 150.
Sometimes we do not have any published information on the value of some varieties with positive sales. It happens mainly for before 1980, with varieties released at the beginning of the 60's. The varieties for which estimated performances can be compiled represent from 60% to 90% of total sales (for all France) before 1980, and more than 95% after 1980. Results has been compiled both by taking into account or not the years before 1980. No difference ha s been observed, so the first years were finally kept.

Graph 1. Estimation of the age effect (B_t)



Graphique 2. Evolution of the year effect (a_t) in the basic model