Soumission aux Journées de Recherche en Sciences Sociales – Nancy – Décembre 2015

Market size and innovation: An application to the French seed market for large crops[1*]

Sylvie Charlot,
INRA, UMR GAEL, Univ. Grenoble Alpes, France
sylvie.charlot@grenoble.inra.fr

Chokri Dridi
Institut du Développement et de la Prospective/IUT
Université de Valenciennes et du Hainaut-Cambresis
and
Regional Economics Applications Laboratory,
University of Illinois at Urbana-Champaign
chokri.dridi@univ-valenciennes.fr

Stéphane Lemarié
INRA, UMR GAEL, Univ. Grenoble Alpes, France
stephane.lemarie@grenoble.inra.fr

September 2015

Abstract

The aim of this paper is to analyse the impact of market size on innovation in the seed industry. The analysis is based on a panel dataset that covers 19 large crops in France during the period 1989-2012. Our econometric analysis is based on a negative binomial specification and we conduct both cross section as well as panel data analysis. We show that the French crop area always has a positive and significant effect on the number of innovations introduced each year. Market size of foreign countries may be either positive or negative revealing synergy or substitution effects. When hybrid crops are considered, the innovation is mainly determined by a positive and very significant fixed effect, crop area having no more influence. This last result can be interpreted as market size being mainly dependent on crop area for non-hybrid crop and dependent of price mark-up for hybrid crops.

Keywords Seed Innovation, Market Size, Crops, Count Data, France.

JEL code C01, L66, O31, Q16.

---

[*] The authors would like to thank Aline Fugeray-Scarbel and Cédric Lanu for their research assistance. This work, part of the BreedWheat project (ANR-10-BTBR-03), has received funding from the French Government managed by the Research National Agency (ANR) in the framework of the Investments for the Future. The authors are solely responsible for any errors contained in the article.

Introduction

Seed innovation is a key driver that explains productivity gains in agriculture and is expected to meet the needs of future durable agriculture. Investment in plant breeding research enables improvements in yield potential, as it has been observed since WWII, as well as pest resistance (decreasing in pesticide use) or quality traits. These investments are made, to a large extent, by private companies and represent more than 10% of their sales. In the seed sector, like many other research-intensive sectors, the incentive to invest in research is driven by the anticipated windfall from the sales of new innovative products (here new improved varieties). There is a consensus in the theoretical literature that market size has a positive impact on research investment (Vives, 2008), however empirical analyses related to this issue is scarce.

The aim of this paper is to analyse more precisely to what extent market size determines innovation in the seed sector. Research investment in this sector is rather specific to the species. If we consider the large crops, it is expected that research investment on corn genetic improvement, for example, to be related to the size of the corn seed market but rather independent of the size of other markets (wheat, barley, rapeseed, etc.). We hence carried out an econometric analysis to explain the number of new products introduced each year with respect to several indicators of market size. This analysis takes inspiration from various recent articles that address such issue in the pharmaceutical market (Acemoglu and Linn ,2004; Dubois et al., 2011). As far as we know, such analysis has never been carried out in agriculture and the seed sector more particularly.

Our analysis is based on a panel dataset that covers 19 large crops in France during the period 1989-2012. Innovations in this sector are mainly embedded in new products, and because of the regulatory framework, each innovation constitutes a new product. As a consequence, the count of new varieties introduced each year can be used to characterise the magnitude of the innovation. The proxy used to characterise the market size are the acreage of each crop (in France and other regions), adjusted by the share of farmer saved seed. We also control for the hybrid versus non-hybrid crops. Our econometric analysis is based on a negative binomial specification and we conduct both cross section as well as panel data analysis. The econometric analysis reveals positive and significant impact of the French market size on the innovation. Interesting results appears when hybrid and non-hybrid crops are distinguished, showing that the market works differently in these two cases.

The paper is organized as follow. We first present the general economic background on the relationship between innovation and market size and the related literature. We then present the context, the data, the model that is estimated and the results.

1. The related empirical literature

Literature addressing the determinants of innovation mainly distinguishes between four factors: technological opportunities, the industrial structure, the potential market size and finally the regulatory setting including intellectual property rights. The different features of the innovation and market structure (including supply, demand, and the regulatory settings) on which the new product or process is possibly sold affects the excepted benefits of firms that invest in R&D.

Market innovation may indeed lead to a (temporary) monopoly situation inducing a monopoly rent for the innovator. The importance of this monopoly power depends on whether the innovation is incremental or radical. In the latter case, innovation is supposed to be very profitable for the firm. However, two other conditions are also necessary to make an innovation profitable. First, the monopoly position induced by the new product or new process must be protected during a sufficiently large period, especially by (intellectual) property rights or by secret. Second, this innovation must correspond to a potential demand or has to create it. The existence of this (potential) anticipated demand and its size are therefore crucial in the decision to invest in the innovation process. Moreover, demand must exist but in many cases it has to be large enough, as a large market allows taking advantage of increasing returns to scale. Since innovation activities are often characterized by high fixed costs, inducing large returns to scale, investors and firms are possibly very sensitive to the size of the potential market when they decide to invest in R&D and to get involved in an innovation process.

A large part of the empirical literature related to the determinants of innovation investments focuses on the role of firms' characteristics (size, sector), the industry's structure (number of firms), demand and regulatory settings such as intellectual property rights (Cohen 2010).

The size of potential demand and therefore of market size is part of the definition of market power, especially when innovation contributes to differentiate goods, services or production process. It indeed may be the case that an innovation creates a new demand that did not exist before and therefore increases the market size. Therefore, the innovation itself affects the market size. There is a reverse causality as market size also depends on innovation. As mentioned previously, empirical literature focuses on characteristics of the supply part of market power, mainly firms' characteristics, and on regulatory settings. Empirical papers trying to assess the effect of market size on the incentives to innovate are not numerous and, to our knowledge, they are all applied to the pharmaceutical sector.

Despite previous research focusing on the relationship between contemporaneous R&D spending and prices, Acemoglu and Linn (2004) were the first to investigate the impact of anticipated market size on firms' R&D spending. They first develop a theoretical model that they test on the drugs market in the US on the long run. In their model there is only one firm with the best technology for each drug leading to a monopoly. The quality of a new drug (innovation) depends on investments in R&D. When the latter are subject to increasing returns to scale, the decision to invest in R&D depends on the potential anticipated market size. Acemoglu and Linn test their model on the US drugs market for the period 1980-2000. In Acemoglu and Linn (2004), the potential market is assessed by using different datasets; a large survey on a panel of about 30 000 individuals as well as a survey conducted by doctors on drug use. They use demographic characteristics (age of population, income categories and mortality) to adjust their market size assessments. When they use lag variable and instrumental variables to control for the potential reverse causality described above, Acemoglu and Lin find an effect of future potential market only and no longer of current potential market.

More recently, Dubois et al. (2012), first notice that in previous models drugs within a given therapeutic group are only vertically differentiated meaning the best drug is assumed to capture the entire market. This is not observed in numerous therapeutic submarkets. They therefore develop a theoretical model with horizontal differentiation *à la* Salop and show that in this context the R&D firms' decisions, incumbent and possible entrants, depend on the value of the fixed costs necessary to develop a new innovation. The industry structure, especially its concentration, therefore depends on the fixed costs; if it is too large there will not be any new entrant. Dubois et al. (2011) then test this model and its predictions by using data containing detailed drug sales in 14 countries, between 1997 and 2007 as well as detailed therapeutic classification of drugs and new patents. This dataset is mixed with demographic data (population and mortality). In order to assess the future market size, Dubois et al. estimate the lifecycle of drugs and use it as a weight of the estimated market size by demographic data. They also use the population in different age categories by country and disease prevalence as instrumental variable, to control for the fact that a successful innovation generates large sales due to its quality and its novelty. Finally, deploying a count data econometric model, they find an estimated positive elasticity of innovation to expected market size of 25%.

2. The seed market context

The seed industry is a research-intensive sector where companies generally invest more than 10% of their sales in R&D (Fernandez-Cornejo, 2004). A large part of this research investment is dedicated to breeding programs that are crop-specific. As a consequence, we expect market size for each crop to influence the size of the corresponding breeding program and the number of innovations for that crop. Note also that seed companies often operate in different countries and define their breeding program at the international level. Hence the number of innovation in one country is related not only to the market size in that country, but possibly also to the market size in other countries. This mechanism might be moderated depending on the adaptation to the local agricultural and climatic conditions. In other words, a positive influence of the market size of a country is expected only when the seed developed for a foreign country is adapted to the domestic (here French) market.

Innovation in the seed sector is mostly embedded in products and there is almost no process innovation. The seed market is regulated, and this regulation defines a standard of what should be defined as a new seed variety (or product). More precisely, a product should be *Distinct, Uniform, and Stable* (DUS). *Distinct* means that a new product should be different from those that are already on the market. *Uniform* means that the seed producer should be able to produce a large quantity of seeds that are genetically identical. Lastly, *Stable* means that the units of seed product sold at different years should be also genetically identical. As a consequence, an innovation (i.e. genetic improvement) leads necessarily to the introduction of a new product.

A seed variety that meets this DUS standard can be protected through *Plant Breeders Rights* (PBR[2]) a *sui generis* intellectual property right. PBR provides a monopoly power for 20 years to its owner with two important exemptions: (i) farmers may use their saved seed (on-farm seed) to sow again the variety the year after, and (ii) other breeders may use this variety as a source of genetic material in their research program.

---

[2] PBR are also denominated as Plant Variety Protection (PVP) in the literature.

Note that the first exemption is not effective when the performance decreases significantly from one generation to the next one, as it is the case in particular with hybrid seeds (eg. corn, canola, sunflower, sugar beet). Finally, we expect that PBR is a source of incentive to invest in research in the seed sector, especially for the crop with hybrid seeds.

In some countries, like those in the European Union, new seed varieties have to be registered on an official catalogue. Catalogues are defined at the national level and there is a European catalogue that is the aggregation of the national catalogues. In the case of large crops, registration requires the new varieties to meet the DUS standard as well as a minimum level of performance with respect to different traits. The official catalogue is an essential source of information for our analysis because it provides a complete inventory of the innovations.

The relationship between market size and innovation has not been addressed in the seed context. Frey (1996) presents data on research investment in plant breeding (measured by the number of researcher) for almost all the US crops. Simple comparison shows that research investments are more important for the major US crops like Corn, Soybean and Wheat, compared to minor crops such as Barley, Oat and Sorghum. However no statistical analysis has been made on these data to test for this relationship.

More attention has been paid to the impact of intellectual property rights on innovation in the seed industry. Eaton and Graff (2014) provide and extensive and recent review of this literature. Among the most important studies, Alston and Venner (2002) find no significant impact of PBR on wheat yield improvement in the US, Carew and Devadoss (2003) find weak positive effect on canola yield in Canada and Naseem et *al*. (2005) find positive effect on cotton yield in the US. Note that these studies estimate the impact of PBR on the evolution of yield over time, with control for different other factors. Each study focuses on one particular crop, considering possibly different states or regions. Because of this focus on one crop each time, it is not possible to assess the impact of the different market size corresponding to different crops, as we do here.

3. Data and descriptive statistics

This analysis is conduced on the basis of the dataset that covers 19 large crops in France during the period 1989-2012. The number of innovations is defined on the basis of the new varieties registered in the French official catalogue (source GEVES). The number of innovations is under-estimated because we do not take into account the seed varieties that are registered in other European countries[3].

Different proxies of market size are used. We first use the land areas for each crop. French areas are obtained from the French Ministry of Agriculture and foreign areas are obtained from FAO. In order to have a reasonable number of explanatory variables, we only retain large zones (group of countries) defined by the FAO. Also, because of the

---

[3] The European catalogue is the agregation of the national catalogues. Any seed variety registred in the European catalogue can be sold in any country of the EU. As a consequence, an innovation registered in spain can be sold in France. Note however that because of the adaptation to the climate the proportion of seed variety that are suitable for several counrties is rather small.

specificity of French agriculture, we only consider zones from Europe (4 groups[4]), America (3 groups) and Northern Africa (1 group).

As presented above, Plant Breeders Rights enable farmer to save their own seed but this practice is only implemented when the seed is not hybrid. As a consequence, for a given area of crop, market size is greater with non-hybrid seed. This characteristic is taken into account by introducing a dummy variable for the crops with hybrid seed: maize, sunflower, sugar beet and rapeseed. In addition, the French area were also corrected by removing the surface sown with on-farm seeds. This correction was not possible for the foreign countries because of the lack of data on the share of farmer saved seed for each country and each crop.

Table 1: Surface by crop and region (in thousand hectares)

Crop	***France***	***W. Europe***	***S. Europe***	***E. Europe***	***N. Europe***	***N. America***	***C. America***	***S. America***	***N. Africa***
Hybrid									
Maize	3196	644	3771	7924	2	30800	8981	18296	1206
Sunflower	806	69	1412	9435	0	1027	1	2729	166
Sugar Beet	426	722	468	2679	349	550	0	40	121
Rapeseed	1116	1232	78	1958	1014	5353	2	80	23
Non hybrid									
Barley	1168	29	114	19	35	4	37	127	423
Beans	49	2478	4183	19534	3796	5888	275	897	3653
Chicory	4	11	0	1	0	0	0	0	0
Flax	70	46	16	310	18	0	0	5	12
Lupins	6	13	14	95	4	0	0	29	13
Mustard Seed	2	7	0	209	1	231	0	0	0
Oats	132	329	809	6982	1040	2412	62	634	112
Peas	219	128	115	2045	169	1097	4	117	62
Potatoes	154	589	570	6736	441	650	85	908	304
Rice	23	0	384	266	0	1233	340	5508	598
Rye	38	807	242	7001	303	293	0	58	24
Sorghum	58	0	50	81	0	3155	1969	1567	5881
Soybeans	29	26	386	1148	0	28665	173	29232	18
Triticale	152	407	92	1314	139	17	0	31	8
Wheat	2903	3576	6610	39671	3799	33047	778	8476	6639

---

[4] The french area is removed from the surface of western area.

Table 2: Surface by crop in % total surface by region

Crop	France	W. Europe	S. Europe	E. Europe	N. Europe	N. America	C. America	S. America	N. Africa
Hybrid									
Maize	30.3	5.8	19.5	7.4	0.0	26.9	70.7	26.6	6.3
Sunflower	7.6	0.6	7.3	8.8	0.0	0.9	0.0	4.0	0.9
Sugar Beet	4.0	6.5	2.4	2.5	3.1	0.5	0.0	0.1	0.6
Rapeseed	10.6	11.1	0.4	1.8	9.1	4.7	0.0	0.1	0.1
Non hybrid									
Barley	11.1	0.3	0.6	0.0	0.3	0.0	0.3	0.2	2.2
Beans	0.5	22.3	21.7	18.2	34.2	5.1	2.2	1.3	19.0
Chicory	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Flax	0.7	0.4	0.1	0.3	0.2	0.0	0.0	0.0	0.1
Lupins	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.1
Mustard Seed	0.0	0.1	0.0	0.2	0.0	0.2	0.0	0.0	0.0
Oats	1.3	3.0	4.2	6.5	9.4	2.1	0.5	0.9	0.6
Peas	2.1	1.2	0.6	1.9	1.5	1.0	0.0	0.2	0.3
Potatoes	1.5	5.3	3.0	6.3	4.0	0.6	0.7	1.3	1.6
Rice	0.2	0.0	2.0	0.2	0.0	1.1	2.7	8.0	3.1
Rye	0.4	7.3	1.3	6.5	2.7	0.3	0.0	0.1	0.1
Sorghum	0.5	0.0	0.3	0.1	0.0	2.8	15.5	2.3	30.5
Soybeans	0.3	0.2	2.0	1.1	0.0	25.1	1.4	42.5	0.1
Triticale	1.4	3.7	0.5	1.2	1.3	0.0	0.0	0.0	0.0
Wheat	27.5	32.2	34.2	36.9	34.2	28.9	6.1	12.3	34.5

The average surfaces for each crop and zone are given in Table 1 in thousand hectares. Table 2 provides them in percentage of the total surface we observe in France and other regions. As can be seen from these figures, the distribution of crops among the different zones is quite uneven. If we focus on the main crops, it can be seen that Wheat, Beans and Sunflower are specific to Europe and Northern Africa while Maize and Soybean are specific to America. Within Europe, France is specific by having rather high share of Maize (similar to northern or southern America), Rapeseed and Barley and very low rate of Bean. Hence the geographic proximity with the rest of Europe does not necessarily lead to similar crop production.

Figure 1 shows the number of innovation as a function of the French area for all years, by crop. The surfaces and the number of innovation are very heterogeneous both between crops and between years for each crop. The number of new innovations introduced each year is much higher for maize, compared to other crops. When we only distinguish hybrid from non-hybrid crops (Figure 2), for a given level of surface, the hybrid crops correspond to a larger number of new seeds on average.

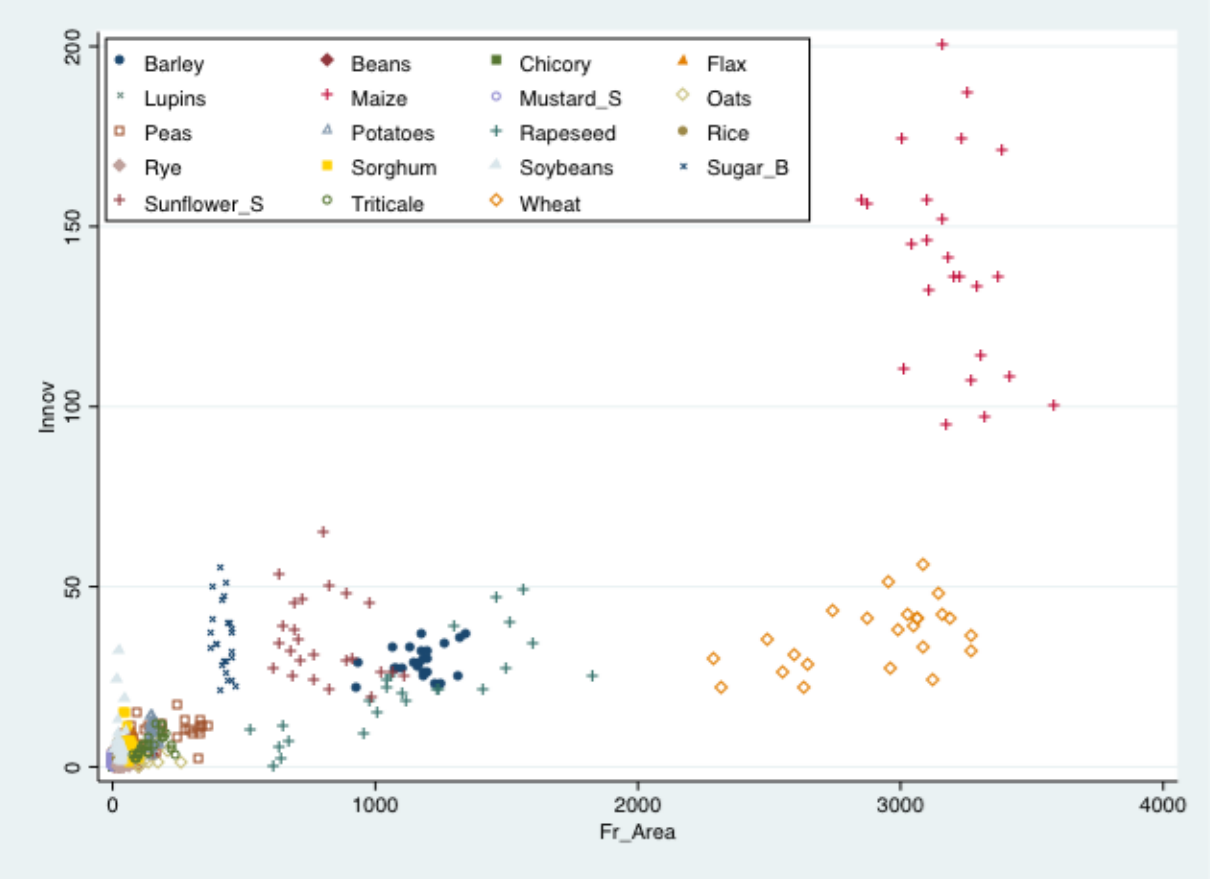

Figure 1: Surface in France by crop (thousand hectares)

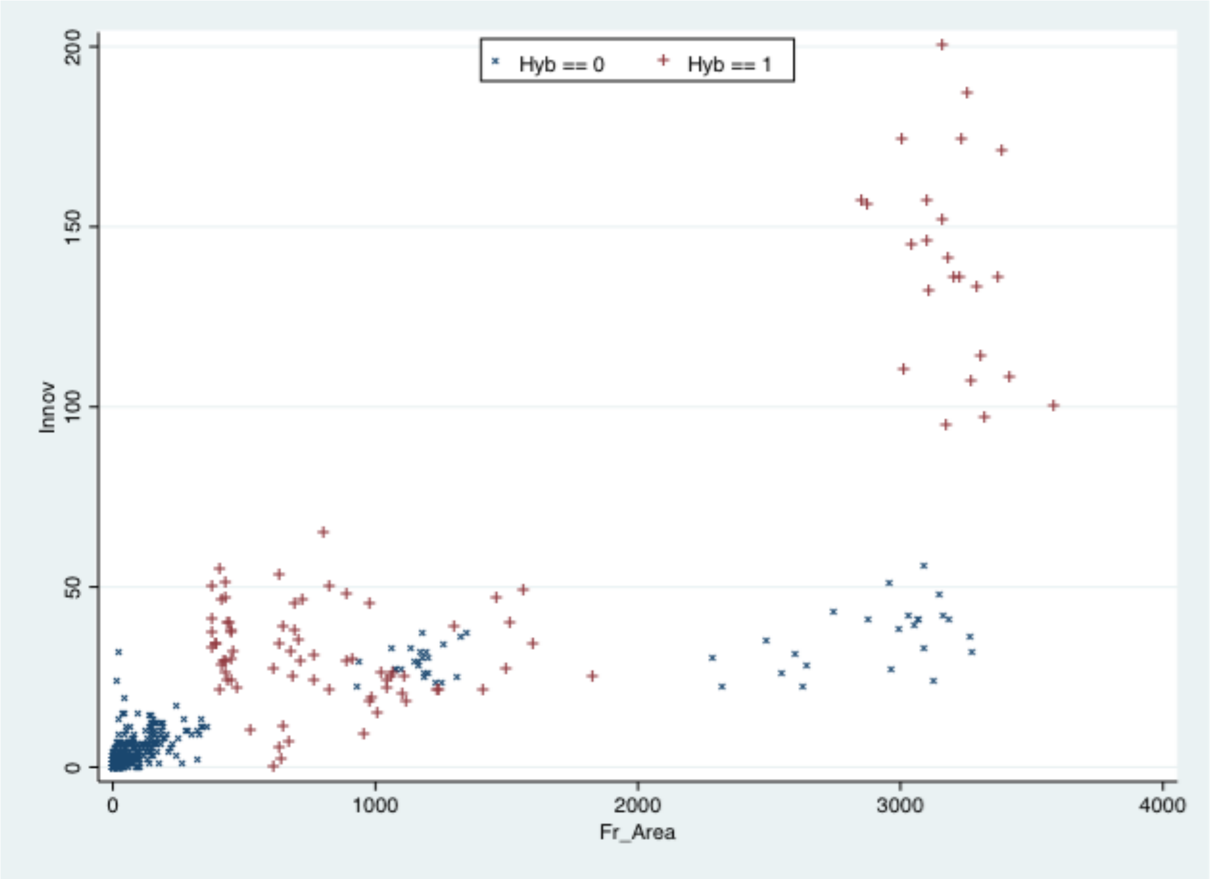

Figure 2: Surface in France by hybrid / non hybrid crop

4. Model

In the context of this paper the outcome of interest (i.e. the dependent variable) is the (simple) count data of yearly registrations of innovations for each of the 19 crops we retain for our analysis. We need therefore to estimate what is referred to as count-data models. However, the very real possibility that in some years no innovation is registered for some varieties and that data can be dispersed, imply therefore that the use of the *Poisson regression* can be problematic. Indeed, some of the problems encountered during the estimation of such models are: truncation in the count data, the presence of many (or excess of) zeros manifested by a low value of the mean, and the overdispersion of the data manifested by a variance-to-mean ratio greater than 1.

Our data is not censored and the zeros, while appearing not excessive, represent 13.38% of the 456 observations we have in the dataset, with the mean of the count data being 18.45 innovations. Since count-data models are usually estimated using the Poisson model, the number of occurrences of event y is therefore given by

$$Pr(Y = y) = \frac{e^{-\mu}\mu^{y}}{y!}, \qquad y = 0,1,2,\cdots$$

such that $E(Y) = Var(Y) = \mu$. From the above probability distribution the expected frequency of zeros when $\mu = 18.45$ is given by $Pr(Y = 0|\mu = 18.45) = e^{-18.45} \approx 0$ implying that there are relatively too many zeros in our dataset.

Furthermore, as in most applications using count-data models, there appears to exist an over dispersion in our data as the variance to mean ratio is greater than 1 (it is 57.23). Indeed a test of dispersion (Cameron & Trivedi, 2005; 670-671) on a basic version of the model using a robust Poisson Maximum Likelihood Estimation (MLE) shows the presence of significant overdispersion. The presence of significant overdispersion in our data implies that a robust negative binomial regression should be used.

The Negative Binomial (NB) regressions consist in estimating the relationship

$$\mu = exp(\boldsymbol{X}.\beta + \varepsilon) \Leftrightarrow Ln(\mu) = \boldsymbol{X}\beta + \varepsilon$$

The difference between the Poisson and the NB model lies in the probability distribution function.

As we observe 19 crops from 1989 to 2012, we can use this panel dimension to also introduce crop and time fixed effects. The crop fixed effects capture the impact of all unobserved and time invariant characteristics, which affect the innovation process for each crop. Time dummies control for annual unobserved variables that affect the innovation process of all crops in the same way.

Wooldridge (2005) show that the fixed effect approach is a very powerful mean to deal with endogeneity because it is still valid in case of correlation between unobservable variables and explanatory variables (see e.g. Heckman and Hotz, 1989). Wooldridge (2010) shows that this model is consistently estimated by using OLS after within transformation, i.e. de-meaned variables.

Below, we present results of a set of negative binomial models robust to heterocedasticity, introducing the time lag variables in order to account for delays in research and delays between innovation date and the date of inclusion in the seeds catalogue.

5. Results

Our results show that the quality of estimates improves when the time lag of surface is used as interest variable and when new control variables are introduced. From column 1 of Table 3 to column 3 of Table 4, the log-likelihood continuously decreases. Regardless of the specification, with fixed-effects or only hybrid control variables, replacing the surface by its lag of two periods strongly improves the quality of the model (column (1) to (2) in each table).

Introducing the time lag surface is a way to control for potential endogeneity linked to the reverse causality. It is possible that a new seed creates its own market if this innovation is very profitable for farmers. It is therefore also possible that some farmers switch from a crop to another once new varieties of seeds are sold on the latter market. This may cause a reverse causality and the surface of each crop becomes dependent on the innovations, thus making it endogenous.

Introducing the time lag of the surface -instead of the surface at the same time period as innovations- permits to tackle this potential endogeneity issue. Indeed, there is no reason for farmers to choose to change the crop they sow before very performing new seeds are sold. The potential reverse causality in this case is not linked to hypothetical expectancies but only to the opportunity a new seed offers once it is on the market. Using a time lag is therefore effective to deal with this endogeneity issue.

We tried different time lags (from 1 to 5 periods)[5]. Longer time lag leads to lower values of the log-likelihood but the estimates basically do not change. We therefore use the two years lag in order not to lessen the data set span too much and to maintain high degrees of freedom in estimated models.

For the models without fixed-effects, the average marginal effects of variables are computable. They are reported in Table 3bis only for our interest variable, i.e. the French area, following the estimates of the negative binomial model (Table3) allowing calculating them.

---

[5] Results are available upon request.

Table 3: Negative binomial model

	(1)	(2)	(3)	(4)
	Surface t	Surface t-2	Surface t-2 + Hybrid	Surface t-2 + Hybrid + Interact
Number of new seed varieties				
Fr_Area	0.00143***	0.00147***	0.00125***	0.00152***
	(14.97)	(15.01)	(15.77)	(18.96)
W_Europe_Fr	0.000566**	0.000601**	-0.000347*	-0.000451**
	(2.89)	(2.99)	(-1.98)	(-2.59)
S_Europe	0.000136	0.000182*	-0.0000519	-0.000153*
	(1.47)	(1.99)	(-0.68)	(-2.11)
E_Europe	-0.0000157	-0.0000236	0.00000238	0.00000221
	(-0.99)	(-1.46)	(0.21)	(0.21)
N_Europe	-0.000759***	-0.000773***	-0.0000859	0.000172
	(-5.24)	(-4.91)	(-0.60)	(1.15)
N_America	-0.0000369*	-0.0000457**	0.00000302	0.000000972
	(-2.01)	(-2.68)	(0.19)	(0.06)
C_America	-0.000162***	-0.000177***	-0.000187***	0.000216***
	(-4.21)	(-4.43)	(-5.78)	(4.32)
S_America	0.0000335*	0.0000412**	0.00000603	0.00000989
	(2.22)	(2.87)	(0.50)	(0.84)
N_Africa	-0.00000700	0.00000198	0.0000808**	-0.0000452
	(-0.27)	(0.07)	(3.03)	(-1.56)
Hyb			1.361***	2.407***
			(12.45)	(14.35)
HybxFr_Area				-0.00154***
				(-9.33)
Constant	1.708***	1.705***	1.456***	1.373***
	(21.66)	(21.55)	(21.98)	(20.66)
Pseudo LL	-1498.956	-1374.22	-1312.77	-1292.42
Wald test	1368.35	1247.83	1899.78	2773.73
Observations	456	418	418	418

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3bis: Average marginal effects

	(1) Surface t	(2) Surface t-2	(3) Surface t-2 + Hybrid	(4) Surface t-2 + Hybrid + Interact
Number of new seed varieties				
Fr_Area	0.0014314^{***}	0.03155^{***}	0.02543^{***}	0.02972^{***}
	(14.97)	(9.60)	(12.16)	(15.50)
Hyb			28.810^{***}	79.542^{***}
			(9.09)	(6.15)
HybxFr_Area				-0.03004^{***}
				(-8.62)

The degree of freedom is not reduced in the crop fixed-effect model as the within transformation allows to estimate it without introducing explicitly these crop fixed-effects. This kind of models allows controlling for all time invariant characteristics of each crop; whether it is hybrid or not, for instance, but also all other features that can affect its innovation process.

To estimate the time fixed-effects, time dummy variables were explicitly introduced leading to a lower degree of freedom. These dummies control for all period exogenous chocks that affect all crops in the same way, as technological progress (genomics,…).

We now discuss the effects that are estimated, based on the Tables 3, 3bis and 4. For all the estimates and all the econometric models, French land areas (a proxy for market size) seem to have a positive and significant effect on the number of innovations. The market size seems to be strongly related to the frequency of innovations in the seed sector, in France during the period we observe. When delays, in research and delays between innovation date and the date of inclusion in the seeds catalogue, as well as the possible endogeneity are controlled for, on average, an increase of one hundred thousand hectares in the surface, sown with one crop, rises the number of new varieties of this crop of about 3 (Column 2 and 4 in Table 3bis).

Table 4: Negative binomial model with crop and time fixed effects

	(1)	(2)	(3)
	Surface t / Crop FE	Surface t-2 / Crop FE	Surface t-2 / Crop + time FE
Number of new seed varieties			
Fr_Area	0.000510^{***}	0.000487^{***}	0.000494^{***}
	(0.00011)	(0.00011)	(0.00013)
W_Europe_Fr	0.000561^{**}	0.000726^{***}	0.000523^{**}
	(0.00018)	(0.00019)	(0.00019)
S_Europe	-0.000140^{*}	-0.000141^{*}	-0.00000841
	(0.000065)	(0.000066)	(0.000077)
E_Europe	-0.00000356	-0.00000217	0.00000350
	(0.0000076)	(0.0000083)	(0.0000081)
N_Europe	-0.000176	-0.000394^{*}	-0.000373
	(0.00016)	(0.00017)	(0.00019)
N_America	0.0000294^{**}	0.0000196	0.0000118
	(0.000010)	(0.000011)	(0.000011)
C_America	0.00000947	0.0000604	0.0000850
	(0.000056)	(0.000055)	(0.000062)
S_America	-0.0000463^{***}	-0.0000483^{**}	-0.0000559^{***}
	(0.000012)	(0.000015)	(0.000015)
N_Africa	-0.0000251	-0.0000631	-0.0000251
	(0.000045)	(0.000051)	(0.000051)
Constant	1.992^{***}	2.166^{***}	2.668^{***}
	(0.19)	(0.19)	(0.24)
Crop fixed-effects	X	X	X
Time fixed-effects			X
LL	-1113.77	-1008.68	-984.99
Wald	90.17	74.11	139.02
Observations	456	418	418

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The effect of the foreign zones can be either positive or negative, the effect of the Western Europe surface being always significant. Positive estimates can be interpreted as a synergy between foreign and domestic research efforts. Negative estimates can be interpreted as a substitution effects: firms having limited resources to allocate for research, they tend to focus on the larger market. Note that these results concerning the effect of foreign market sizes should be taken with caution because the sign of these effects are not consistent depending on the model specification.

The positive and significant estimates of the French market size effect is about three times lower for the panel estimation, compared to the cross section estimates. Two interpretations of this result can be given. First, the crop and time fixed effects in the panel estimation enable capturing the effect of non-observable variables. This unobservable can be specific to each crop and stable over time (e.g. industry structure) or a time choc identical for all the crops (e.g. agricultural policy reforms). This first interpretation tends to say that, because of non-observable variables, the cross section tends to over estimate the impact of the French market size. The second interpretation is related to the difference of market size variation captured with these two types of estimation. More precisely, cross section estimation captured both inter-crops and intra-crop market size variation while the panel estimation captured only intra-crop variation, inter-crops variation being captured by the crop fixed effects. This second interpretation tends to say that the research investment of the industry is more sensitive to inter-crop market size difference rather than the intra-crop difference.

The cross section estimate enables to isolate the effect of crops for which hybrid seed are commercialized (estimations 3 and 4 in Table 3 and Table 3bis). In Table 3, Column (3), a specific dummy variable is introduced for hybrid crop. This dummy is highly significant and the positive effect of the French area on the number of innovations is still significant. The positive effect of the hybrid dummy can be interpreted by the higher price markup made by companies for these types of crop. Indeed, farmers cannot save their own seed with hybrid crop. As a consequence, competition occurs only between seed suppliers, and product differentiation on this market enables companies to have higher price markup. Conversely, with non-hybrid crop, farmers can save their own seed, which introduces competition with the supply of the seed companies, leading to lower price markup. At last, note that the positive effect of the hybrid dummy cannot be interpreted by the lower quantity of seed sold because the areas sown with farmer saved seed has been removed to define the French crop area[6].

Two "hybrid effects" are introduced in Column 4 of Tables 3 and 3bis: a hybrid fixed effect and an interaction with the crop area in France. In such a framework, the fixed effect becomes then more important (compared to estimates in Column 3) and the interaction with the crop area becomes negative and almost equal to the (positive) effect of the area estimated for all the crops. This effect disappears for hybrid seeds (Column 4). However hybrid seeds have a strong positive impact on the number of innovations comparing to non-hybrid (79 on average, cf. Column 4 in Table 3bis), but this effect is not linked to the market size variations.

Hence, for hybrid seed, there is no impact of the surface on the number of crops. This result does not contradict the fact that the market size has a positive effect on the innovation. It simply reveals that this effect is observed through different channels, depending on the crop: for non-hybrid crops, this effect is based on the quantity sold (i.e. surface), while for hybrid crop, this effect is based on the important seed price mark up.

---

[6] For example the wheat production covers 6Mha in France and half of this surface is sown with farmer saved seed. We hence consider surfaces of about 3Mha in our estimations (see figure 1).

Conclusion

This paper is a first attempt to estimate the impact of the market size on the innovation process, in the seed market. The expected positive effect of the market size is confirmed on a dataset that cover 19 crops in France between 1989 and 2012. Descriptive statistic show that the number of new seed varieties introduced each year is much larger for crops with hybrid seed. This positive effect of hybrid seed is confirmed by the estimates. Interestingly, we show that, for these hybrid seeds, market size is no longer related to crop area but probably to price market up. The effect of the crop area in foreign country is taken into account in our analysis but the estimates are rather inconclusive for the time being; the sign of the estimated effects being sometime different, depending on the specification.

This analysis can be enriched in several ways to better understand the determinants of the innovation in the seed industry. One interesting way would be to conduct the analysis at the European level by using innovation count data from the European seed catalogue. The analysis can also be enriched by including other determinants, such as the seed industry structure and other factors that may influence the crop market size such as the agricultural policy or the commodity prices.

References

Acemoglu, D. and J. Lin (2004): "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry", *Quarterly Journal of Economics*, 1049-1090.

Alston, J. M., and R. J. Venner. (2002) “The effects of the US Plant Variety Protection Act on wheat genetic improvement”. *Research Policy*, 31(4), 527-42..

Cameron, A.C., P.K. Trivedi (2009): *Microeconometrics using Stata*, Stata Press

Carew, R., and S. Devadoss (2003) “Quantifying the Contribution of Plant Breeders’Rights and Transgenic Varieties to Canola Yields: Evidence from Manitoba”. *Canadian Journal of Agricultural Economics*, 51(3), 371-95

Cohen, W.M. (2010): "Fifty Years of Empirical Studies on Innovative Activity and Performance", *Handbooks in Economics*, Elsevier, Vol. 1, Ch. 4, 129-213.

Dubois, P., W. Man Wynne Lam, O. de Mouzon, F. Scott-Morton and P. Seabright (2012): "How Does Innovation in the Pharmaceutical Industry Respond to Policies that Influence Expected Market Size?", *SCIFI-GLOW: SCience, Innovation, FIrms and markets in a GLObalized World, CEPR Policy Paper*.

Eaton, D. and G. Graff. (2014) “The dynamic IP system in crop genetics and biotechnology”. In S. J. Smyth, P.W.B. Phillips, and D. Castle (eds) *Handbook on Agriculture, Biotechnology and Development*, 425-449. Edward Elgar Publishing.

Fernandez-Cornejo, F. (2004): *The Seed Industry in US agriculture: An exploration of data and information on crop seeds markets, regulation, industry structure, and research and development*. United States Department of Agriculture, Economic Research Service, Agriculture Information Bulletin No. 786, Washington, January 2004.

Frey, K. J. (1996) “National plant breeding study - I : Human and financial resources devoted to plant breeding research and development in the United States in 1994”. *Special Report 98. Iowa State University, Iowa Agriculture and Home Economics Experiment Station*.

Naseem, A. and J. F. Oehmke (2005) “Does Plant Variety Intellectual Property Protection Improve Farm Productivity? Evidence from Cotton Varieties”. *AgBioForum*, 8(2&3), 100-107.

Vives X. (2008): "Innovation and competitive pressure", *Journal of Industrial Economics*, 56(3), 419-469.

Wooldridge M. J. (2005) Fixed-Effects and Related Estimators for Correlated Random-Coefficient and Treatment-Effect Panel Data Models, *The Review of Economics and Statistics* 87(2), 385-390.

Wooldridge M. J. (2010) *Econometric analysis of cross-section and panel data*, Second Edition, The MIT Press Cambridge, Massachusetts London, England.