Competition, Product and Process Innovation: an empirical analysis*

Carlos Daniel Santos†

University of Alicante, CEP and UECE

Abstract

Understanding how changes in competition affect innovation and productivity is an old and important economic question. The main contribution of this article is to separate the effects of competition on the two types of innovation: product and process. I introduce explicit assumptions about unobserved heterogeneity in a model of firm behavior that allow me to identify the causal effect. I use a sample of Spanish firms to estimate the model. The sample contains useful features to answer the question at hand, namely, detailed information about the types of innovation and, self reported measures of market structure.

Overall, I find that competition, measured by the number of competitors or market shares, has negative effects on product innovation and no effects on process innovation. I then explain why no effect on process innovation should be expected. By shifting demand, competition directly changes the optimality condition for product but not for process innovation. Thus, competition has no *direct* effects on process innovation or, as a consequence, productivity. Changes in competition operate only indirectly on process innovation, namely through firm size.

***Acknowledgments:** I would like to thank Pedro Albarran, Steve Bond, Bo Honoré, Eugenio Miravete and John Van Reenen and participants at the PEJ meeting (2010), ESWC (2010) and University of Alicante for helpful comments. Financial support from the Instituto Valenciano de Investigaciones Economicas, Ministerio de Ciencia e Innovación and FEDER funds under project SEJ-2007-62656 is gratefully acknowledged.

†University of Alicante, csantos@ua.es

1 Introduction

In this article I analyze the relation between competition and the two sides of innovation: product and process. Figure 1 presents the main evidence for a sample of Spanish firms - reported innovations decrease steadily with the increase in the number of competitors. However, as later reported, the effect disappears for process innovation once we condition on size. The significant size effect is consistent with larger firms doing more R&D and innovation and could be due to the existence of large fixed and sunk costs in the R&D process and unrelated with any competition effects. The results thus suggest that the effect of competition on process innovation operate through firm size (similarly to the conclusions in Gilbert (2006)).

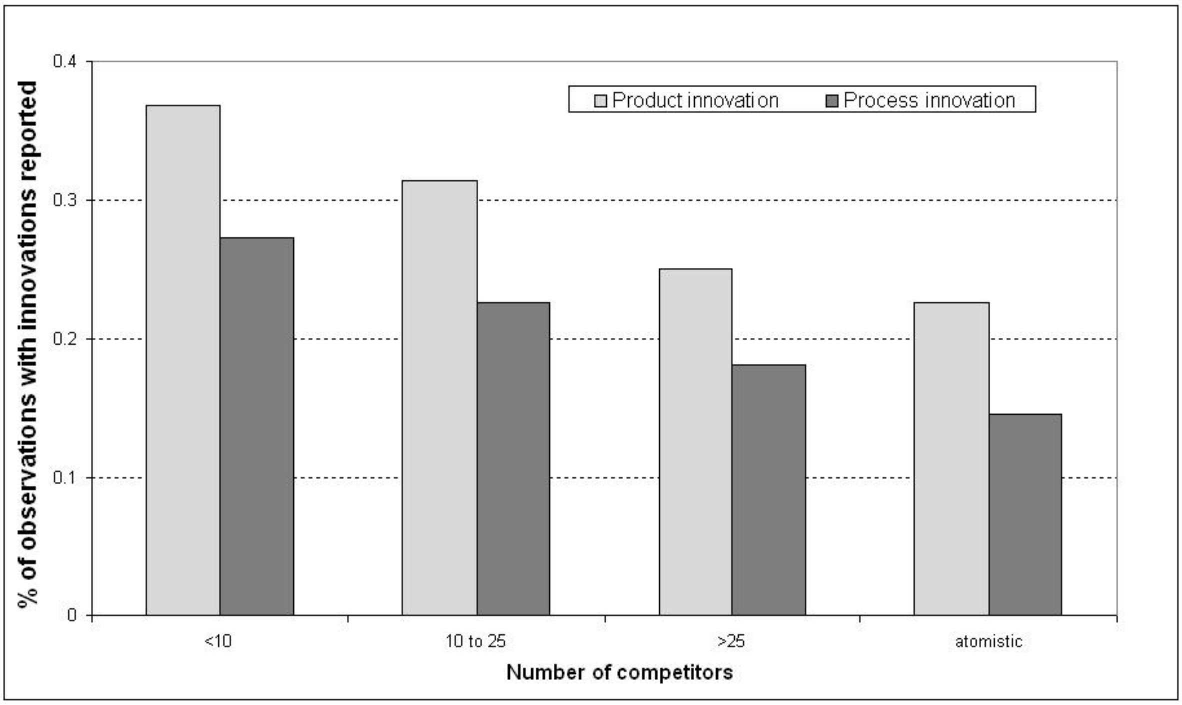

Figure 1: Percentage of firms with innovation by number of competitors

There are traditionally two main views regarding the effects of competition on innovation: for the Schumpeterian view monopolies favor innovation, whereas for the alternative view competition favors innovation (Arrow, 1962). In practice, both approaches have valid arguments (see the discussion in Schmutzler (2009)). For example, Vives (2008) characterizes the effects for a range of competition measures and market structures. Overall, he shows that even in a very simple setting the results change significantly depending on the precise measures of competition and market structure used. Thus, empirical results are important to understand and separate the mechanisms by which competition affects innovation (see the extensive surveys of Kamien and Schwartz (1975), Cohen and Levin (1989) and Gilbert (2006)). To address the problem I use a sample of Spanish firms (*Encuesta sobre Estrategias Empresariales*) containing very detailed information on innovation and market structure together with the usual accounting variables (e.g. sales, capital stock, investment, employment, profits). The measures of competition are directly reported by the firms and not imputed at the predetermined industry code, making them a very close measure of the competition perceived by individual firms. This has the advantage of being closer to the information contained in the firm's information set. Moreover, the large period covered in this panel (1990-2006) allows me to address some of the problems with non-linear panel data models. In particular, I estimate fixed effects and dynamic models where consistency results are valid for large time dimension.

The evidence reported here is consistent with the findings of previous studies that use the same dataset. For example, González et al (2005) find a positive effect of market share and concentration on R&D decision. Huergo and Jaumandreu (2004) analyze how product and process innovation change with firm size and age to find that both types of innovation increase with firm size. Cassiman and Martinez-Ros (2007) study the effect of product and process innovation on exports. Interestingly, they find that firm productivity is explained by product innovation but not by process innovation.[1] Finally, using British data Aghion and Grifith (2008) and Aghion et al (2005) estimate an inverted-U effect of competition on innovation (measured by patents). To address the endogeneity problems they use policy changes as instruments. However, their methodology does not explain how the policy changes affect innovation. The policy changes do not satisfy any exclusion restriction and affect the environment in several ways besides market structure.[2] The failure of the policy changes to meet an exclusion restriction is a well know problem in the program evaluation literature and one of the reasons why a structural approach is often required (Heckman, 2008).

The problem that remains in the empirical literature is how to address the equilibrium nature of both market structure and innovation. In particular, competition variables will depend on industry characteristics and so will innovation. Econometric models suffer from

[1] This result is at odds with what we would expect but can be rationalized if productivity is mismeasured. In the absence of firm level price data, revenues are used instead of physical quantities to estimate total factor productivity. The resulting productivity estimates include efficiency together with product quality and market power.

[2] One example is the use of privatization episodes that lead to a direct change in competition but also firm size. Changes in firm size, and not competition *per se*, can be the responsible for the increase (decrease) in innovations .

reverse causality and omitted variables that can lead to serious endogeneity problems. In this article I address these questions by writing down an explicit model of firm behavior and stating the assumptions about the information structure, endogenous and exogenous variables. In particular, I make explicit assumptions that allow me to identify the causal effect. I assume a timing between states and decisions to avoid the reverse causality and focus on the omitted variables problem. I then present conditions on the omitted variables such that the effects can be consistently estimated.

In the final section I explain how the results fit with the theory. Changes in competition shift product demand. Thus, competition directly changes the optimality conditions for product innovation (revenue side) although it has no *direct* effect on process innovation (cost side). Measures of competition then satisfy an exclusion restriction in the process innovation equation, arising naturally from the additivity of revenues and costs. Technically, competition is required to shift revenues and process innovation to shift costs. Because the profit function is additive in revenues and costs, when the derivatives are taken with respect to process innovation, the competition term is dropped from the first order condition. The result implies that competition has no *direct* effects on firm level productivity, if the source of productivity improvement is process innovation. The distinction between *direct* (conditional) and *indirect* (unconditional) effects is important and has not been emphasized in previous work. Competition does have effects on process innovation that operate through an indirect channel, namely through firm size. Characterizing the unconditional (indirect) effects has been the main focus of the literature. Perhaps unsurprisingly, both the empirical and the theoretical literature have found no robust unconditional effect of competition on innovation. Although the unconditional effects are certainly relevant from a policy perspective, understanding how the mechanism operates is important to guide the choice amongst alternative policies. Understanding the mechanism is one of the several reasons why a structural approach is useful in interpreting the empirical results.

The structure of the article is as follows. Section two details the empirical strategy and section three discusses the data. Section four contains the results, section five provides a theoretical explanation and finally section six offers some concluding remarks.

2 The Model

Because exogenous variations in competition are rare events, I adopt a different approach to estimate the effect from competition on innovation. I start by formulating the firms' information set and decisions and analyze which restrictions will allow me to identify the competition effect. In particular, I vary the assumptions making them less restrictive leading to more complex estimation methods. As the assumptions cannot be tested I take the evidence from varying the estimation methods as supportive (or not) of the results from the simple models.[3]

There are two sets of variables - outcome variables chosen by the firm, and state variables observed by the firm. The outcome variables of interest are product and process innovation (PdI and PcI), R&D decision and intensity (rd and RD/Y). Let the vector of outcome variables be denoted by $Y = \{PdI, PcI, rd, RD/Y\}$. The state variables can be divided into firm level variables (X) and market structure variables (μ). We can include in the set of state variables firm size, productivity, market share or the number of competitors. There are also state variables which are observed by the firm but unobserved by the econometrician (ξ). The three sets of variables constitute the firm's information set in period t,

$$\Omega_{it} = \{X_{it}, \mu_{it}, \xi_{it}\}$$

The first assumption is about the timing of decisions. Firms observe the current state in period t and decisions take one period to materialize. Thus, we can abstract from any problems related with reverse causality.[4] The optimal decision of firm i in period t is

$$Y^*_{i,t+1} = f^*(\Omega_{it}) \tag{1}$$

For now I will abstract from how market structure is determined in equilibrium. This abstraction is irrelevant for estimation purposes if the agents are playing a dynamic Markov game with a stationary equilibrium (and assuming no structural changes in the sample). Optimal decisions in period t are the solution to the dynamic game taking as given the (Markov) equilibrium beliefs. Thus, equation (1) can be the policy function solving the dynamic problem faced by the individual firm

$$\begin{array}{c} \max_{\{Y^*_{is}\}_{s=t+1}^{\infty}} E\left[\sum_{s=t}^{\infty} \beta^{s-t} \pi(\Omega_{is}) | \Omega_{it}\right] \\ s.t.\ (\Omega_{i,s+1}) = h(Y_{i,s+1}, \Omega_{is}, \nu_{i,s+1}) \end{array}$$

where β is the discount factor, $\nu_{i,s+1}$ is a vector of stochastic variables, $\pi(\Omega_{is})$ are the

[3] These assumptions involve the relation between the observed and unobserved variables. The fact that the variables are unobserved makes the assumptions untestable.

[4] In some cases, omitted variables are confused with reverse causality. For example, in a dynamic setting, expectations about future market structure can affect current innovation decisions. Provided the information set is correctly specified, there is no reverse causality problem because the firm uses current information to infer about the future. However, the econometrician not observing the full information set generates an omitted variables problem. In these circumstances future market expectations are not correctly conditioned.

profits obtained in state (Ω_{is}) and a controlled first order Markov transition is imposed on the state variables. Notice that some of the choice variables can become a state variable in the next period (e.g. $PdI \in X$). I distinguish between firms decision (Y^*) and realized outcomes ($Y_{i,t+1} = f(\Omega_{it}, \varepsilon_{i,t+1})$) as some of the outcomes are inherently stochastic. For example, firms decide on innovation trials but innovation outcomes are stochastic and subject to random shocks $\varepsilon_{i,t+1}$. Such shocks are not part of the firm's information set.

In this setting, omitted variables are the main cause for econometric problems. In particular, any unobserved market level time varying variable is most likely correlated with the market structure, making it impossible to estimate the effect of μ on Y. For example, let ξ represent industry level technological opportunities. When the opportunities are better, innovation increases. However, technological opportunities also generate entry into the market causing market structure to change. Overall, the underlying technological opportunity can shift both the market structure and innovation and generate spurious correlation between the two variables. Thus, we cannot estimate the equation (1) because a part of the state space (ξ) is unobserved. I will focus on the endogeneity of market structure, i.e. when μ_{it} is not independent from ξ_{it}.

2.1 Omitted variables

Now I discuss conditions under which we can address the omitted variables problem. This discussion will guide me in the empirical section for the choice of estimation methods. Our interest is to estimate the policy function

$$Y_{i,t+1} = f(X_{it}, \mu_{it}, \xi_{it}, \varepsilon_{i,t+1}) \tag{2}$$

where by assumption $\varepsilon_{i,t+1}$ is unobserved and independent from any information at period t.

Model A) One trivial solution to the endogeneity problem is when either ξ_{it} does not enter equation (2) or ξ_{it} is independent from μ_{it}. In both of these cases equation (2) can be directly estimated.

Model B.i) With no loss of generality the unobserved component can be factored in the following way $\xi_{it} = \xi_i + \xi_t + \xi'_{it}$. When ξ'_{it} is independent over time and independent from μ_{it}, we have the traditional "fixed effects". This adds the restriction that technological opportunities cannot vary over time. Thus, we can simply add firm and time specific dummies to estimate equation (2) (see Arellano and Honoré (2001)).

Model B.ii) Alternatively, we can allow dependence between ξ'_{it} and μ_{it} but maintain independence over time. In this case previous lags, $\mu_{i,t-1}$ are valid instruments for μ_{it}.

Up to now we have only exploited statistical relations. For example, it is very hard to motivate economically why ξ_{it} is independent over time whereas lagged values of μ_{it} are still powerful instruments. I now exploit the economic structure of the problem.

Model C.i) In some cases market structure is correlated with the unobserved market characteristics and the assumption of independence over time for ξ'_{it} is not plausible. However it is plausible to assume that conditional on lagged (or current) characteristics, market structure is independent from future values. For example, $\mu_{it}|\xi_{i,t-1}$ is independent from $\xi_{it}, \xi_{i,t+1}, ...$(or $\mu_{it}|\xi_{it}$ is independent from $\xi_{i,t+1}, \xi_{i,t+2}, ...$). The first case is more plausible for some market variables like the number of firms, because entry and exit take one period to materialize. Finally, assume that unobserved characteristics follow an exogenous first order Markov process

$$\xi_{i,t+1} = h^{\xi}(\xi_{it}, \nu^{\xi}_{i,t+1}) \tag{3}$$

In such case lagged values of the dependent variable can "control" for the unobserved component. We can write

$$Y_{it} = f(X_{i,t-1}, \mu_{i,t-1}, \xi_{i,t-1}, \varepsilon_{it})$$

and if the function f is monotonic in $\xi_{i,t-1}$ we can invert it[5]

$$\xi_{i,t-1} = f^{-1}(Y_{it}, X_{i,t-1}, \mu_{i,t-1}, \varepsilon_{it}) \tag{4}$$

Using equations (3) and (4), and replacing in equation (2)

$$Y_{i,t+1} = f(X_{it}, \mu_{it}, h^{\xi}(f^{-1}(Y_{it}, X_{i,t-1}, \mu_{i,t-1}, \varepsilon_{it}), \nu^{\xi}_{it}), \varepsilon_{i,t+1}) \tag{5}$$

As $\mu_{it}|\xi_{i,t-1}$ is independent from ξ_{it}, μ_{it} is also independent from ν^{ξ}_{it} (and from any previous lags). There is still a problem of dependence between Y_{it} and ε_{it}. The solution is simple because lagged values $Y_{i,t-1}$ are valid instruments for Y_{it}.

In the second case $\mu_{it}|\xi_{it}$ is independent from $\xi_{i,t+1}, \xi_{i,t+2}, ...$ and μ_{it} will be correlated with ν^{ξ}_{it}. Thus, equation (5) can no longer be estimated because of this correlation. However,

[5] Invertibility is violated when the outcome variables are binary. In this case we can use other outcome variables that also depend on technological opportunites, for example, investment. See case c.ii).

lagged values $(\mu_{i,t-1}, \mu_{i,t-2}, ...)$ are valid instruments for μ_{it}.

Model C.ii) The final case occurs when there are more than two policy variables. This is useful when the first policy is not invertible. Let Y^2 be a second policy function that depends on the same set of state variables (one example is capital investment)

$$Y_{it}^2 = f_2(X_{i,t-1}, \mu_{i,t-1}, \xi_{i,t-1}, \varepsilon_{it}^2)$$

If function f_2 is invertible we can write

$$\xi_{i,t-1} = f_2^{-1}(Y_{it}^2, X_{i,t-1}, \mu_{i,t-1}, \varepsilon_{it}^2)$$

and replace this, together with equation (3), in equation (2)

$$Y_{i,t+1} = f(X_{it}, \mu_{it}, g^{\xi}(f_2^{-1}(Y_{it}^2, X_{i,t-1}, \mu_{i,t-1}, \varepsilon_{it}^2), \nu_{it}^{\xi}), \varepsilon_{i,t+1})$$

As in the previous case Y_{it}^2 is correlated with ε_{it}^2 but lagged values are valid instruments. Also, when μ_{it} is not independent from ν_{it}^{ξ}, lagged values are also valid instruments.

Moving from Model A) to C) relaxes the restrictions on the unobserved component. As such, estimating the equation of interest with the different models will allow me to evaluate how important are the different assumptions. In particular, the assumptions are not valid if the estimated parameters are not robust. On the other hand if the estimated parameters are robust to the different assumptions this is evidence supporting their validity. In the empirical section below I present results using the different models in A) to C).

2.2 Econometric specification

I study the effects of size, market share and the number of competitors on each firms' individual R&D decisions and innovation outcomes using a linear index approximation to equation (2). Firms' decisions are R&D (binary and intensity) and innovation. R&D expenditures are sometimes a poor measure of innovative effort. Furthermore, we cannot observe the share of R&D expenditures dedicated to product and process innovation. For such reasons, observing both types of innovation will be extremely useful in the empirical section.

A natural choice for the information set (state variables) is to include the capital stock, productivity and labor. These variables are quite persistent and subject to adjustment costs so that they naturally become a state of the dynamic problem. Equilibrium variables (market structure) are also states of the dynamic equilibrium. In this case I use both the number of competitors and market shares as the observed measures of market structure. Finally I add individual fixed effects as further states to capture unobserved firm level differences. There is also marginal costs and product quality as (unobserved) state variables. Costs and quality will be captured by measured productivity and sales. I further add these to the vector of states.

The parameters of interest are the coefficients on market share, number of competitors and size. I allow for a quadratic term on market share to account for possible nonlinearities as found in the previous literature (e.g. Aghion et al, 2005).

Innovation outcomes The discrete decision to innovate is modeled as a standard logit model:

Product innovation

$$PdI_{it+1} = \begin{cases} 1 \text{ if } PdI^*_{it+1} > 0 \\ 0 \text{ otherwise} \end{cases}$$

$$\begin{aligned} PdI^*_{it+1} &= \alpha_0^{pdi} PdI_{it} + \alpha_1^{pdi} PcI_{it} + \alpha_2^{pdi} \mu_{it} + \alpha_3^{pdi} \mu_{it}^2 + \alpha_4^{pdi} N_{it} \\ &\quad + \alpha_5^{pdi} y_{it} + \alpha_6^{pdi} rd_{it} + \alpha_7^{pdi} k_{it} + \alpha_8^{pdi} l_{it} + \xi_{it+1}^{pdi} \end{aligned} \quad (6)$$

where $\xi_{it+1}^{pdi} = \xi_i^{pdi} + \xi_{t+1}^{pdi} + \xi_{it+1}^{\prime pdi}$ and $\xi_{it+1}^{\prime pdi}$ is an i.i.d. logistic error term. PdI and PcI are dummy variables for product and process innovations, μ is market share, N total number of competitors, y is the log of sales, k the log of capital and l the log of employment. rd is a dummy variable taking the value one if R&D is reported and zero otherwise.

Process innovation

$$PcI_{it+1} = \begin{cases} 1 \text{ if } PcI^*_{it+1} > 0 \\ 0 \text{ otherwise} \end{cases}$$

$$\begin{aligned} PcI^*_{it+1} &= \alpha_0^{pci} PdI_{it} + \alpha_1^{pci} PcI_{it} + \alpha_2^{pci} \mu_{it} + \alpha_3^{pci} \mu_{it}^2 + \alpha_4^{pci} N_{it} \\ &\quad + \alpha_5^{pci} y_{it} + \alpha_6^{pci} rd_{it} + \alpha_7^{pci} k_{it} + \alpha_8^{pci} l_{it} + \xi_{it+1}^{pci} \end{aligned} \quad (7)$$

where $\xi_{it+1}^{pci} = \xi_i^{pci} + \xi_t^{pci} + \ \xi_{it}^{\prime pci}$ and $\xi_{it}^{\prime pci}$ is an i.i.d. logistic error term.

R&D discrete decision The discrete decision to do R&D is also modeled as a standard logit model:

$$rd_{it+1} = \begin{cases} 1 \text{ if } rd^*_{it+1} > 0 \\ 0 \text{ otherwise} \end{cases}$$

$$\begin{aligned} rd^*_{it+1} &= \alpha_0^{rd} PdI_{it} + \alpha_1^{rd} PcI_{it} + \alpha_2^{rd} \mu_{it} + \alpha_3^{rd} \mu_{it}^2 + \alpha_4^{rd} N_{it} \\ &\quad + \alpha_5^{rd} y_{it} + \alpha_6^{rd} rd_{it} + \alpha_7^{rd} k_{it} + \alpha_8^{rd} l_{it} + \xi_{it+1}^{rd} \end{aligned} \tag{8}$$

where $\xi_{it+1}^{rd} = \xi_i^{rd} + \xi_{t+1}^{rd} + \xi_{it+1}^{\prime rd}$ and $\xi_{it+1}^{\prime}{}^{rd}$ is an i.i.d. logistic error term.

R&D continuous decision Conditional on the decision, the R&D intensity (R&D to sales ratio) is

$$\begin{aligned} \left.\frac{RD_{it+1}}{Y_{it+1}}\right|_{rd_{it+1}=1} &= \alpha_0^{RDY} PdI_{it} + \alpha_1^{RDY} PcI_{it} + \alpha_2^{RDY} \mu_{it} + \alpha_3^{RDY} \mu_{it}^2 + \alpha_4^{RDY} N_{it} \\ &\quad + \alpha_5^{RDY} y_{it} + \alpha_6^{RDY} rd_{it} + \alpha_7^{RDY} k_{it} + \alpha_8^{RDY} l_{it} + \xi_{it+1}^{RDY} \end{aligned} \tag{9}$$

where $\xi_{it+1}^{RDY} = \xi_i^{RDY} + \xi_{t+1}^{RDY} + \xi_{it+1}^{\prime RDY}$.

Potential econometric problems There are several problems with panel data limited dependent variable models in short panels (see Arellano and Honoré (2001) for a review). In particular we have:[6]

1. *Strict exogeneity of the regressors*: This assumption is hardly ever met given the structure of the problem. Besides the system structure, which means all variables are endogenously determined, we also have the inclusion of lagged dependent variables (see below). However, the identification problem is less severe when T is large as in our case (Chamberlain, 1985; Honoré and Kyriazidou, 2000).

2. *Incidental parameters (fixed effects)*: This problem arises in short panels because the number of parameters to be estimated increases with the number of observations N. As for the previous case, the problem is less severe for large T.

3. *Lagged dependent variables*: The existence of lagged dependent variables directly violates the strict exogeneity assumption above. Although it is well know how to solve

[6] Notice that I am interested in the sign and not the magnitude of the coefficients. I am also not interested in distinguishing persistence due to unobserved heterogeneity from "true" state dependence. For this reason we can abstract from problems about identification of the marginal effects.

this problem in linear models, the issue is much more complicated in non-linear models where some estimators have been proposed (Honoré and Kyriazidou, 2000). However, with a large time-series, we can adopt an unconditional logit to mitigate the problem. For this reason again, large T helps to reduce potential bias in this case. Furthermore, results with and without the lagged dependent variables are presented to access their robustness.

The changes from random effects to the fixed effects specifications are, in most cases, not substantial. This could be because unobserved heterogeneity is well "controlled" by the set of state variables. For example, Blundell, Griffith and Van Reenen (1999) explore the existence of pre-sample data to "control" for the initial conditions.

3 Data

The data is part of the ESEE (*Encuesta sobre Estrategias Empresariales*) collected by the *Fundacion Empresa Publica*. It consists of an unbalanced sample of 30,466 observations for 4,094 firms over the period 1990-2006 for the whole manufacturing sector in Spain with an average 1,800 observations per year. The survey collects a variety of variables on R&D, innovation and market structure and has been used by other authors to study R&D and Innovation.[7] The dataset is particularly attractive for the empirical analysis of the relationship between market structure and innovation. For example, firms are directly asked about the number of competitors faced in the market. This is a direct measure of actual perceived competition and has advantages over imputing variables from a predetermined industry code. A description of the data and variable construction is contained in the Appendix and some descriptive statistics are presented in Tables 1 and 2. On average 35% of the firms report positive R&D expenditures, 30% introduce process innovations and 23% introduce product innovations. Thus, process innovations are much more frequent then product innovations. The fact that 34% of innovating firms report both types of innovation could be due to the complementarities as analyzed by Athey and Schmutzler (1995). There is also a decline in innovative output (both product and process) from 2000 until 2003 and a slight recovery in 2004. This is also noticeable in R&D expenditures which have declined from 1999 until 2002 and recovered afterwards. There is also a steady decline in reported market shares probably signaling the increased competition for the Spanish industry over this period.

[7] For example, Cassiman and Martinez-Ros (2007), Doraszelski and Jaumandreu (2007), Gonzalez, Jaumadreu and Pazó (2005) and Huergo and Jaumadreu (2004).

Variables	*Mean*	*Standard deviation*	*Min*	*25th percentile*	*Median*	*75th percentile*	*Max*
RD dummy	36%	48%	0%	0%	0%	100%	100%
Process Innov	32%	47%	0%	0%	0%	100%	100%
Product Innov	23%	42%	0%	0%	0%	0%	100%
Sales (EUR mio)	48.40	228.00	0.01	1.07	4.53	29.30	5,940.00
RD intensity	0.7%	2.8%	0.0%	0.0%	0.0%	0.4%	70.0%
Employment	263	816	0	18	47	257	25,363
Capital Stock (EUR mio)	21.60	92.80	0.00	0.30	1.67	13.20	3,240.00
Market share	12%	19%	0%	0%	0%	19%	100%
Age	25	22	0	9	19	33	270

Table 1: Descriptive statistics

Year	*Obs.*	*RD*	*Process Innov*	*Product Innov*	*Conditional RD intensity*	*Inv. rate*	*Sales growth*	*Conditional market share*	*VA growth*
1990	1,834	634	345	301	2.74%	.	.	34	.
1991	1,995	727	740	496	2.08%	15%	5%	30	5%
1992	1,954	678	671	487	2.03%	12%	3%	28	-2%
1993	1,831	626	637	440	2.24%	10%	-1%	28	3%
1994	1,837	645	654	478	2.01%	9%	15%	28	21%
1995	1,679	587	580	399	1.93%	11%	15%	29	14%
1996	1,706	590	577	402	1.92%	11%	5%	27	4%
1997	1,901	660	696	489	1.93%	11%	14%	28	10%
1998	1,766	668	676	461	1.89%	13%	11%	27	5%
1999	1,741	661	621	442	2.10%	14%	10%	25	2%
2000	1,849	693	720	508	2.07%	14%	10%	25	-1%
2001	1,709	620	570	363	2.01%	14%	6%	26	5%
2002	1,704	630	510	375	1.73%	11%	3%	26	-4%
2003	1,378	490	347	259	1.93%	10%	4%	24	4%
2004	1,373	507	380	282	1.81%	10%	7%	23	2%
2005	1,869	689	527	358	2.18%	13%	4%	23	6%
2006	1,988	689	545	370	2.04%	11%	8%	22	10%

Table 2: Aggregate descriptive statistics, averages per year

4 Results

In this section I present the estimates for equations (7) to (9). For the binary models I adopt a logit specification both for random and conditional logit fixed effects.[8] I have also directly estimated fixed effects (unconditional logit) with very similar results to the conditional fixed effects case. The sample for the fixed effects models is substantially smaller because several observations are lost (all zeros and all ones). For this reason whenever the results are similar, random effects are preferred.

Overall, the results show a robust negative relation between competition (as captured by the number of competitors) and product innovation and no relation between competition and process innovation or R&D decisions. Regarding market shares, I find evidence of an increasing and concave relation between market shares and product innovation but weakly significant for process innovation. I also find some evidence of a concave relation between market shares and the R&D discrete decision. Since the results condition for size (sales, capital stock and number of workers), market shares could be capturing either a form of market power or product quality. For this reason the number of competitors is a "cleaner" measure of competition.

A further relevant fact is that size (sales) is positive and significant for innovation (product and process) and R&D decisions but not for R&D intensity. This is consistent with the existence of large fixed and sunk costs for R&D (see for example Santos, 2010). Thus, larger firms are more likely to do R&D and innovate, but not necessarily dedicate a larger share to the R&D process. In this case the relation between size and innovation is unrelated from competition because size is not signalling any form of competition. In some cases more competition can even lead to larger firm size.

4.1 Innovation equations

The results for product innovation are presented in Table 3 and for process innovation in Table 4. Columns (i) to (iii) are valid for Model A) above. Column (iv) is the simple fixed effects valid in Model B.i). Model C.i) is not valid because the binary step function is not invertible and I use investment as the second policy of Model C.ii). The results with no instruments are reported in columns (v) and (vi). Finally the correct use of lagged values as instruments for the number of competitors and market shares are reported in column (vii), for the lagged dependent variable in column (viii), for lagged investment in column (ix), and for both the competition measures and investment in column (x).

[8] Adopting a probit specification does not change the results.

4.1.1 Product innovation

The number of competitors is negatively related with product innovation whereas market shares have a positive and concave relation.[9] There is also a strong size effect (sales) that together with the market share effect suggests advantages beyond the size effects.

The results for the number of competitors in column (iii) are robust to the introduction of fixed effects (conditional and unconditional) and the control function. As expected, R&D intensity is always positive and significant making the relation between competition and innovation conditional on R&D expenditures. The effect of competition on innovation operates on the split of the R&D expenditures between product and process innovation.

4.1.2 Process innovation

Figure 1 suggests that process innovation shares the same competitive effects as product innovation. As in the data only 26% of the firms do one single type of innovation (i.e. the remaining either do both or none) we would indeed expect very similar effects. From Table 4 we conclude that this is not the case. Indeed product innovation is a very strong explanatory variable probably due to potential complementarities in the R&D process (see the work of Athey and Schmutzler (1995)). However, the effect of the number of competitors on process innovation reported in columns (i) and (ii) completely disappears once we condition on size in column (iii). The market share effect is similar to the one for product innovation but it is not statistically significant. The results hold for the remaining specifications. We can conclude that there is no evidence of any causal effect from competition to process innovation.

The difference from the product innovation equation is quite striking. If we compare the results in columns (v) of Table 3 and 4 the coefficients are very similar on all but the market level variables suggesting that the forces of competition operate very differently on the two types of innovation.

The results seem to contradict the theoretical predictions. In Vives (2008) an increase in the number of firms leads to a decrease in cost reducing effort. However, the theoretical predictions are unconditional and operate through the size of the firm whereas the empirical results presented here are conditional on firm size. In fact, if we exclude firm size as in columns (i) and (ii) competition has a strong negative effect. Thus, the results are perfectly in line with the theoretical predictions. Once we condition on the correct set of variables the effect disappears and competition has no *direct* effects on process innovation.

[9] We cannot reject the hypothesis of concavity against the inverted-U relation proposed by Aghion et al. (2005). This is due to the few observations with very large market shares (i.e. above 50%).

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Dependent variable: PdI_{it+1}										
Model:		Logit			Conditional FE logit			IV logit		
PdI_{it}					0.93***			0.40***		
					(0.05)			(0.11)		
PcI_{it}		0.78***	0.69***	0.41***	0.16***	0.37***	0.41***	0.32***	0.31***	0.35***
		(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.07)	(0.12)	(0.12)
Mkt $\text{sh}_{it} * 100$	0.55***	3.17***	1.81***	1.00*	0.94*	0.93	-0.80	0.65	0.81	-1.35
	(0.17)	(0.41)	(0.46)	(0.53)	(0.55)	(0.58)	(0.76)	(0.57)	(0.60)	(0.84)
Mkt $\text{sh}_{it}^2 * 1000$		-0.39***	-0.26***	-0.14**	-0.14**	-0.13*	0.11	-0.10	-0.14*	0.15
		(0.05)	(0.06)	(0.07)	(0.07)	(0.08)	(0.11)	(0.08)	(0.08)	(0.12)
$N_{it} * 100$	-15.11***	-11.39***	-9.84***	-7.35**	-6.96**	-9.02**	-11.60*	-7.37**	-7.94**	-19.34***
	(2.80)	(2.81)	(3.03)	(3.43)	(3.56)	(3.81)	(5.92)	(3.64)	(4.00)	(6.87)
y_{it}			0.30***	0.19**	0.16*	0.20*	0.26**	0.17*	0.19	0.24*
			(0.06)	(0.09)	(0.09)	(0.11)	(0.10)	(0.10)	(0.12)	(0.13)
k_{it}			-0.03	-0.17**	-0.13*	-0.23**	-0.16**	-0.14*	-0.34*	-0.23
			(0.04)	(0.07)	(0.07)	(0.09)	(0.08)	(0.08)	(0.20)	(0.21)
l_{it}			0.15**	0.13	0.10	0.07	0.21*	0.19	0.06	0.13
			(0.07)	(0.11)	(0.11)	(0.12)	(0.12)	(0.12)	(0.13)	(0.14)
$R\&D/Y_{it}$			11.28***	4.99***	3.53***	4.59***	5.14***	4.69***	5.59***	5.55***
			(1.35)	(1.34)	(1.28)	(1.40)	(1.39)	(1.45)	(1.64)	(1.69)
Inv_{it}						0.04			0.28	0.17
						(0.03)			(0.24)	(0.25)
Nobs	22,408	22,405	19,770	10,054	9,833	8,523	8,971	9,260	8,127	7,263
Nfirm	3,418	3,418	3,149	1,074	1,058	973	979	1,005	934	863
Time dum.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ind. dum.	Yes	Yes	Yes	No	No	No	No	No	No	No

Notes: Columns (i) to (iii) report random effects logit estimates where the explained variable is R&D dummy. In columns (iv) to (vi) results are reported for a conditional fixed effects logit. Finally columns (vii) to (x) report IV estimates with market share and N_{it} instrumented with own lags in column (vii), PdI_{it} and Inv_{it} instrumented with own lag in columns (viii) and (ix) respectively and Market Share, N_{it} and Inv_{it} instrumented with own lags in column (x). *** significant at 1%, ** significant at 5%, * significant at 10%.
Results for an unconditional fixed effects (not shown) yeld identical results to the conditional specification.

Table 3: Product innovation estimates.

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Dependent variable: PcI_{it+1}										
Model:		Logit			Conditional FE logit			IV logit		
PcI_{it}					*0.86****			0.09		
					(0.04)			(0.12)		
PdI_{it}		*0.85****	*0.77****	*0.43****	*0.17****	0.41***	0.41***	0.39***	0.37***	0.36***
		(0.05)	*(0.05)*	*(0.06)*	*(0.06)*	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)
Market share$_{it} * 100$	*0.61****	*1.92****	*0.74***	*0.75*	*0.50*	0.76	-0.54	0.57	0.76	0.61
	(0.14)	*(0.34)*	*(0.37)*	*(0.46)*	*(0.47)*	(0.50)	(0.67)	(0.48)	(0.50)	(0.51)
Market share$^2_{it} * 100$		3.43	3.72	4.61	4.71	-0.12*	0.06	-0.08	-0.12*	-0.10
		(0.00)	*(0.01)*	*(0.01)*	*(0.01)*	(0.07)	(0.10)	(0.06)	(0.07)	(0.07)
$N_{it} * 100$	*-6.71****	*-4.65***	*0.16*	*1.59*	*1.85*	2.68	-2.33	1.87	2.68	0.22
	(2.24)	*(2.27)*	*(2.40)*	*(2.80)*	*(2.85)*	(3.11)	(4.75)	(2.91)	(3.11)	(3.23)
y_{it}			*0.16****	*0.15***	*0.13*	0.07	0.11	0.14*	0.08	0.08
			(0.05)	*(0.08)*	*(0.08)*	(0.09)	(0.08)	(0.08)	(0.10)	(0.11)
k_{it}			*0.17****	*0.01*	*-0.06*	-0.25***	0.02	0.00	-0.13	-0.13
			(0.03)	*(0.06)*	*(0.06)*	(0.07)	(0.06)	(0.06)	(0.15)	(0.17)
l_{it}			*0.00*	*0.09*	*0.05*	0.08	0.18*	0.11	0.14	0.20*
			(0.06)	*(0.09)*	*(0.09)*	(0.10)	(0.10)	(0.09)	(0.10)	(0.11)
RD/Y_{it}			*3.69****	*3.23***	*2.73***	3.19**	1.97*	2.69**	2.08	1.88
			(1.05)	*(1.29)*	*(1.23)*	(1.35)	(1.16)	(1.29)	(1.35)	(1.38)
						0.18***			0.16	0.21
						(0.02)			(0.18)	(0.19)
Obs	22,838	22,380	19,752	13,625	13,625	11,610	12,020	12,813	10,904	9,649
Firms	3,444	3,411	3,142	1,462	1,462	1,346	1,318	1,382	1,264	1,154
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	No	No	No	No	No	No	No

Notes: Columns (i) to (iii) report random effects logit estimates where the explained variable is R&D dummy. In columns (iv) to (vi) results are reported for a conditional fixed effects logit. Finally columns (vii) to (x) report IV estimates with market share and N_{it} instrumented with own lags in column (vii), PcI_{it} and Inv_{it} instrumented with own lag in columns (viii) and (ix) respectively and Market Share, N_{it} and Inv_{it} instrumented with own lags in column (x). *** significant at 1%, ** significant at 5%, * significant at 10%.
Results for an unconditional fixed effects (not shown) yeld identical results to the conditional specification.

Table 4: Process innovation estimates.

Because process innovation is normally regarded as productivity enhancing, this implies that competition has no *direct* effects on firm level productivity.

4.2 R&D equations

I separate the R&D decisions into the discrete decision and the choice of intensity. The reason for doing so is because these decisions seem to be determined by different factors. In particular, the presence of fixed or sunk costs would be important for the discreteness whereas particular R&D technology characteristics could be the drivers of the second decision. Modelling both decisions jointly with a tobit specification reveals that the results are mostly driven by the discrete decision.

4.2.1 Discrete R&D decision

The negative correlation between the number of competitors and R&D in column (i) of Table 5 disappears once size effects are accounted for in column (iii). Market shares have a positive and concave relation in the estimates of columns (iv) and (v). However, this is no longer significant after the introduction of fixed effects in columns (vi) to (viii). Instrumenting the competition measures and the control function in columns (ix) to (xii) reveals the same pattern - the binary R&D decision is mostly driven by internal factors (size, previous innovation).

4.2.2 R&D intensity

Results for R&D intensity are reported in Table 6 where the sample is restricted to observations with positive R&D. At first sight R&D intensity is decreasing with market share. However, the results loose significance once we allow for fixed effects. These results are not surprising given the lack of good empirical results in the literature for explaining R&D intensity. The reason for the poor fit is most likely because R&D intensity is determined by unobserved heterogeneity, thus, the fixed effects explain most of the variation.

5 Rationalizing the evidence

The previous results fit well in the theory. Changes in market structure have a direct effect on the demand for a firm's product but not on the cost structure (this excludes the possibility of spillovers as analyzed by Levin and Reiss (1988)). Naturally, product innovation shifts the demand curve whereas process innovation shifts the cost curve. Variations in demand then change the optimality condition for product innovation and leave the optimality condition

Dependent variable: rd_{it+1}

	(i)	*(ii)*	*(iii)*	*(iv)*	*(v)*	*(vi)*	*(vii)*	*(viii)*	*(ix)*	*(x)*	*(xi)*	*(xii)*
Model:			Logit				Conditional FE logit				IV logit	
Market share$_{it}$ $*$ 100	1.24***	4.44***	1.93***	1.40**	1.11**	0.75	1.02	0.60	-0.14	1.02	1.21*	-0.16
	(0.21)	(0.52)	(0.56)	(0.57)	(0.46)	(0.64)	(0.68)	(0.71)	(0.95)	(0.68)	0.73	1.10
Market share$^2_{it}$ $*$ 1000		-0.46***	-0.24***	-0.16**	-0.11*	-0.08	-0.11	0.00	0.10	-0.10	-0.16	0.11
		(0.06)	(0.07)	(0.07)	(0.06)	(0.08)	(0.09)	(0.10)	(0.10)	(0.10)	0.10	0.16
N_{it}	-0.1***	-0.08***	-0.05	-0.05	-0.03	0.01	0.03	0.03	0.06	0.05	0.03	0.07
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)	(0.07)	(0.04)	0.05	0.08
y_{it}			0.74***	0.70***	0.32***	0.56***	0.46***	0.57***	0.78***	0.57***	0.53***	0.66***
			(0.08)	(0.08)	(0.05)	(0.11)	(0.11)	(0.12)	(0.13)	(0.12)	0.15	0.16
k_{it}			0.27***	0.25***	0.09***	-0.02	-0.09	-0.09	-0.07	-0.06	-0.30	-0.37
			(0.05)	(0.05)	(0.03)	(0.08)	(0.09)	(0.11)	(0.10)	(0.09)	0.24	0.26
l_{it}			0.44***	0.38***	0.17***	0.10	0.06	0.03	-0.03	0.01	0.01	-0.06
			(0.09)	(0.09)	(0.06)	(0.12)	(0.13)	(0.14)	(0.14)	(0.13)	0.14	0.16
PdI_{it}				1.14***	0.69***	0.71***	0.39***	0.66***	0.79***	0.54***	0.70***	0.72***
				(0.08)	(0.07)	(0.08)	(0.09)	(0.09)	(0.09)	(0.09)	0.09	0.10
PcI_{it}				0.51***	0.22***	0.36***	0.13	0.37***	0.43***	0.27***	0.28*	0.29*
				(0.07)	(0.06)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	0.14	0.15
rd_{it}					3.42***		1.54***			0.87***		
					(0.07)		(0.06)			(0.11)		
Inv_{it}								0.09***			0.34	0.41
								(0.03)			0.29	0.30
Time dum.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ind. dum.	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No
Obs	22,661	22,661	20,122	19,753	19,673	6,884	6,805	5,900	6,056	6,364	5,506	4,869
Firms	3,428	3,428	3,173	3,143	3,137	740	731	682	664	680	632	581

Notes: Columns (i) to (v) report random effects logit estimates where the explained variable is R&D dummy. In columns (vi) to (viii) results are reported for a conditional fixed effects logit. Finally columns (ix) to (xii) report IV estimates with market share and N_{it} instrumented with their lags in column (ix), rd_{it} and Inv_{it} instrumented with their lag in columns (x) and (xi) respectively and Market Share, N_{it}, rd_{it} and Inv_{it} instrumented with their lags in column (xii). *** significant at 1%, ** significant at 5%, * significant at 10%. Results for an unconditional fixed effects (not shown) yeld identical results to the conditional specification.

Table 5: R&D dummy logit estimates.

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Dependent variable: RD/Y_{it+1}								
Model:	RE				FE		GMM	Sys. GMM
Market $\text{share}_{it} * 10^3$	-0.05**	-0.14***	-0.17***	-0.18***	-0.07	-0.09	-0.05	-0.25**
	0.02	0.05	0.05	0.06	0.06	0.06	0.05	0.11
Market $\text{share}^2_{it} * 10^5$		0.13**	0.24***	0.25***	0.08	0.12	0.11	0.39**
		0.07	0.07	0.07	0.08	0.08	0.08	0.18
$N_{it} * 10^3$	-0.16	-0.27	-0.57	-0.64	-0.53	-0.54	0.11	-0.45
	0.38	0.39	0.41	0.42	0.46	0.46	0.18	0.66
$y_{it} * 10^3$			-7.16***	-7.46***	-6.70***	-5.93***	-7.17**	-6.72**
			0.86	0.88	1.19	1.19	3.22	3.41
$k_{it} * 10^3$			0.23	0.35	1.05	0.39	-0.18	-2.11
			0.60	0.61	0.97	0.98	1.94	3.30
$l_{it} * 10^3$			4.99***	5.21***	3.61**	3.66**	1.50	12.13*
			1.03	1.05	1.44	1.44	2.46	6.46
$PdI_{it} * 10^3$				2.62***	1.93***	1.67**	1.39	3.58*
				0.69	0.73	0.73	0.85	1.99
$PcI_{it} * 10^3$				-0.86	-0.59	-0.53	-0.98	-1.29
				0.66	0.70	0.70	0.67	1.21
RD/Y_{it}						0.08***	0.08	0.17*
						0.01	0.10	0.09
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	No	No	No	No
R2	9%	10%	12%	12%	2%	17%		
Obs	8,144	8,144	7,152	6,915	6,915	6,866	6,866	6,866
Firms	1,620	1,620	1,490	1,470	1,470	1,464	1,464	1,464
AR1							-2.28	-2.16
p-val							0.02	0.03
AR2							1.15	1.16
p-val							0.25	0.25
Sargan							793	404
p-val							0.00	0.00
Sargan difference								389
p-val								0.06

Notes: Only observations with R&D intensity above zero and below 70% are used. Columns (i) to (iv) contain the results for the random effects logit estimates. Columns (v) and (vi) contain the results for a fixed effects model while in columns (vii) and (viii) a dynamic panel data model is estimated (GMM).
data model is estimated (GMM). *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 6: R&D intensity estimates.

for process innovation unaltered. Thus, changes in competition are expected to have direct effects on product innovation but no direct effects on process innovation. The effect is *conditional* on the relevant set of variables and due to the additive structure of the profit function in revenues and costs. As market structure changes, the whole system rearranges and unconditional effects on process innovation take place. Vives (2008) provides a detailed analysis of several models under different competitive regimes. Like most theoretical results in the literature he focuses on the unconditional effects (see also Boone (2000)). Although the relevant effects from an economic perspective are the unconditional, the conditional effects allow us to understand and isolate the channel by which competition affects innovation.

To illustrate this define a simple two period N -player game.[10] In the first period firms choose their marginal costs (ω) and the quality of their products (χ). In the second period firms set quantities (or prices) conditional on the strategies of the other players. Each firm i faces an inverse demand function $p(q_i, \chi_i, \xi_i^p, \chi_{-i}, q_{-i})$ and a cost function $c(\omega_i, q_i, \xi_i^c)$, where ξ_i^p and ξ_i^c is heterogeneity in the demand and cost functions. This heterogeneity is assumed to be common knowledge but we can easily allow it to be private information. In the case of private information we need to specify the distribution from which heterogeneity is drawn and allow players to form the corresponding beliefs. Market structure and competition are determined by $(\chi_i, q_i, \chi_{-i}, q_{-i})$. For example, market share is equal to $\frac{q_i}{\sum_{j=1}^{N} q_j}$ and the number of active players is equal to $\sum_{j=1}^{N} \mathbf{1}(q_j > 0)$. Let's assume the demand and cost functions satisfy the conditions for existence of a solution. Besides these regularity conditions allow the profit function to satisfy the following:

i) The profit function is additively separable in revenues and costs,

ii) Competition measures are excluded from the cost function,

iii) Process innovation is excluded from the revenue function.

Conditions i) to iii) are satisfied for the most common problems. For condition ii) to be satisfied we cannot have spillovers or network economies. The optimal quantities (or prices) will be a function of the underlying N dimensional vector of costs and qualities chosen in the first stage. In the second stage firm i chooses

$$\max_{q_i \in [0, q_{\max}]} \quad \widetilde{\pi}_i = [p(q_i, \chi_i, \xi_i^p, \chi_{-i}, q_{-i}) - c(\omega_i, q_i, \xi_i^c)]q_i$$

[10] The results can be extended to a more general infinite horizon game with the introduction of appropriate conditions.

And the solution is

$$q_i^* : p(q_i^*, \chi_i, \xi_i^p, \chi_{-i}, q_{-i}) - c(\omega_i, q_i, \xi_i^c) + (p_1^N(q_i, \chi_i, \xi_i^p, \chi_{-i}, q_{-i}) - c_2(\omega_i, q_i, \xi_i^c))q_i = 0 \text{ for all } i \tag{10}$$

The solution to the set of N equations is the N dimensional vector of optimal quantities as a function of the vector of qualities ($\widetilde{\chi} = (\chi_1, .., \chi_N)$), productivity ($\widetilde{\omega} = (\omega_1, .., \omega_N)$) and heterogeneity ($\widetilde{\xi} = (\xi_1^p, ..\xi_N^p, \xi_1^c, .., \xi_N^c)$)

$$q^*(\widetilde{\chi}, \widetilde{\omega}, \widetilde{\xi}) = (q_1^*(\widetilde{\chi}, \widetilde{\omega}, \widetilde{\xi}), ..., q_N^*(\widetilde{\chi}, \widetilde{\omega}, \widetilde{\xi}))$$

In the first period firms choose the quality and marginal costs. For this they have to pay an investment cost $g(\chi, \omega, \xi^g)$ which again can be heterogeneous across firms.

$$\max_{\omega_i, \chi_i} \pi_i = [p(q_i^*, \chi_i, \xi_i^p, \chi_{-i}, q_{-i}) - c(\omega_i, q_i^*, \xi_i^c)]q_i^* - g(\chi_i, \omega_i, \xi_i^g)$$

The "open-loop" strategies for each firm i satisfy the first order condition[11]

$$p_2(q_i^*, \chi_i, \xi_i^p, \chi_{-i}, q_{-i})q_i^* - g_1(\chi_i, \omega_i, \xi_i^g) \quad = \quad 0 \quad (\chi_i^*) \tag{11}$$

$$-c_1(\omega_i, q_i^*, \xi_i^c)q_i^* - g_2(\chi_i, \omega_i, \xi_i^g) \quad = \quad 0 \quad (\omega_i^*) \tag{12}$$

It is clear from equation 12 that conditional on (q_i, χ_i), competition measures do not affect the first order conditions. As such, when competition changes, for a given level of (q_i, χ_i), the optimal level or process innovation remains unaltered. In this case all the effects from changes in competition operate on process innovation through the choices of quantity and quality.

Summarizing, changes in competition are reflected in shifts of the demand curve. They have a direct effect on optimal product innovation whereas the effect on process innovation is indirect (in this simple case via optimal quantities or product innovation). If we condition for quantity and product innovation, competition has no effect on process innovation. Notice the restriction that competition must not enter the cost function, thereby directly excluding externalities or network economies from innovation.

[11] Restricting to "open-loop" solutions is sensible since it is unlikely that players observe their oponents' decisions of investment in process innovation.

6 Conclusion

I have presented evidence on the effect of competition on innovation by exploring variation of this effect across the two types of innovation. The results suggest no competitive effects on process innovation. When put into perspective the results are rationalizable. Competition has a direct effect on product market demand and subsequently a direct effect on product innovation. Because the optimality conditions for process innovation are not directly influenced by competition, there is no direct effect of competition on process innovation. The fact that competition effects on cost reducing efforts disappear once we condition on size, illustrates the need to be careful when analyzing the role of competition on productivity. Moreover, competition will have unconditional effects on process innovation and productivity, and selection effects on **average** industry productivity. I have just focused on the conditional (direct) effects. As shown by the myriad of theoretical and empirical results, the unconditional effects from competition on innovation are hard to characterize generically and have to be understood on a case by case basis. Characterizing the unconditional effects thus requires a full description of the equilibrium responses that can only be achieved by resorting to a structural model.

References

[1] Arrow, K. (1962): "Economic Welfare and the Allocation of Resources for Invention," NBER Chapters, in: *The Rate and Direction of Inventive Activity: Economic and Social Factors*, 609-626, NBER

[2] Aghion, P., Bloom, N., Blundell, R., Griffith, R., and P. Howitt (2005): "On Competition and Innovation: An Inverted U Relationship", *Quarterly Journal of Economics*, 120 (2), 701-728

[3] Aghion, P., and R. Griffith (2008): *Competition and Growth: reconciling theory and evidence*, Zeuthen Lectures, MIT Press

[4] Arellano, M., and B. Honoré (2001): "Panel Data Models: Some recent developments", in *Handbook of Econometrics*, vol. 5, ch. 53, ed. by J. Heckman and E. Leamer, Elsevier, 3229–3296.

[5] Athey, S., and A. Schmutzler (1995): "Product and Process Flexibility in an Innovative Environment", *Rand Journal of Economics*, 26 (4), 557-574

[6] Boone, J. (2000): “Competitive Pressure: The Effects on Investments in Product and Process Innovation”, *Rand Journal of Economics*, 31, 549–569.

[7] Blundell, R., Griffith, R. and J. Van Reenen (1999): "Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms, *Review of Economic Studies,* 66 (3), 529-554

[8] Cassiman, B., and E. Martinez-Ros (2007): "Product Innovation and Exports: Evidence from Spanish Manufacturing", manuscript

[9] Dasgupta, P., and J. Stiglitz (1980): "Industrial Structure and the Nature of Innovative Activity", *Economic Journal*, 90, 266-293

[10] Doraszelski, U., and J. Jaumandreu (2007): "R&D and productivity: Estimating production functions when productivity is endogenous", manuscript

[11] Cohen, W., and R. Levin (1989): "Empirical Studies of Innovation and Market Structure", *Handbook of Industrial Organisation*, edited by R. Schmalensee and R.D. Willig, 1059-1107

[12] Gilbert, R. (2006): “Looking for Mr. Schumpeter: Where Are We in the Competition-Innovation Debate?” In A. B. Jaffe, J. Lerner, and S. Stern (eds.): *Innovation Policy and the Economy*, vol. 6. MIT Press

[13] González, X., Jaumandreu, J. y Pazó, C. (2005): “Barriers to innovation and subsidy effectiveness”, *Rand Journal of Economics*, 36

[14] Griffith, R., Huergo, E., Mairesse, J., and B. Peters (2006): “Innovation and productivity across four European countries”, *Oxford Review of Economic Policy,* 22(4)

[15] Heckman, J. (2008): "Econometric Causality", *International Statistical Review,* 76(1), 1-27

[16] Honoré, B., and E. Kyriazidou (2000): "Panel data discrete choice models with lagged dependent variables", *Econometrica*, 68 (4), 839-874

[17] Huergo, E., and J. Jaumandreu (2004): “How does probability of innovation change with firm age?”, *Small Business Economics*, 22, 193-207

[18] Kamien, M., and N. Schwartz (1975): "Market Structure and Innovation: A Survey", *Journal of Economic Literature*, 13(1), 1-37

[19] Levin, R. C. and P. C. Reiss (1988): “Cost-Reducing and Demand-Creating R&D with Spillovers”, *Rand Journal of Economics*, 19 (4), 538–556

[20] Santos, C. (2010): "Sunk Costs of R&D, Trade and Productivity: The Moulds Industry Case", mimeo

[21] Schmutzler, A. (2009): "Is Competition Good for Innovation? A Simple Approach to an Unresolved Question", *Foundations and Trends in Microeconomics*, 5(6), 355-428

[22] Sutton, J. (1998): *Technology and Market Structure*, MIT Press

[23] Vives, X. (2008): "Innovation and Competitive Pressure", *The Journal of Industrial Economics*, 56(3), 419-469

A Data Appendix

A.1 Data and sample construction

Some notes on the ESEE: The dataset has been collected by the *Fundacion Empresa Publica* since 1990 for a panel of Spanish manufacturing firms. There has been an effort to avoid attrition in the dataset by bringing back to the sample firms that have dropped reporting for some reason not related with exit. The data is collected using direct interviews with a questionnaire. The sampling procedure includes all manufacturing firms with more then 200 employees. Firms with 10 to 200 employees are randomly sampled by industry and size strata, holding around a 4% of the population. Firms with less than 10 employees are excluded from the survey. The ESEE is representative of Spanish manufacturing firms classified by industrial sectors and size categories and includes exhaustive information at the firm level, especially regarding exporting and innovation activities.

Representativeness: The sample is representative of the whole industry and an effort is done to introduce new firms into the sample in order to maintain representativeness.

Variable construction:

Firm level deflators (index) for sales and inputs (materials and services) are constructed using reported variations in prices for sales and inputs. Missing values in the reported price variations are "filled" in two ways. If the missing value is only for one year, an average of the reported price changes in the years immediately before and after is used. If the missing value is for more then one year or a starting or end year industry level value added deflators collected from the OECD/STAN Database for Structural Analysis are used. The deflator indices are therefore time and firm specific.

Industry and aggregate deflators - Industry deflators (value added and gross fixed capital formation) were collected from OECD/STAN Database for Structural Analysis. Unit labor costs for the whole manufacturing sector were collected from the OECD.

Capital stock is constructed separately for land, buildings and other fixed assets using the perpetual inventory method

$$K_{it+1}^{j} = (1 - depreciation^{j}) * K_{it}^{j} + I_{it}^{j} \qquad j = land, build, other$$

$$K_{it+1} = K_{it+1}^{land} + K_{it+1}^{build} + K_{it+1}^{other}$$

The depreciation rate used is 2.5% for buildings, 15% for other fixed assets and 0% for land. Deflators for the capital stock were collected from the OECD/STAN Database for Structural Analysis.

Value added is equal to deflated sales subtracted from deflated materials and deflated external services expenditures (deflator construction explained above)

$$VA_{it}^{def} = Y_{it}^{def} - M_{it}^{def} - ESE_{it}^{def}$$

R&D dummy variable takes a value equal to one whenever positive R&D is reported and zero otherwise.

Other variables used are

Market share is constructed using two questions from the survey. Firms are first asked if they have a significant market share. If not, a zero is automatically attributed, otherwise firms are asked to report their market share. Due to this, zero reported market shares represent a significant proportion of the data (53%).

Number of competitors are directly reported by the firms. It is collected as an ordinal variable that takes four possible values. 1 - Less than 10 competitors; 2 - 10 to 25 competitors; 3 - More than 25 competitors; 4 - Atomistic market.

Cleaning Original data consists of an unbalanced sample of 31,470 observations for 4,357 firms over the period 1990-2006 for the manufacturing sector in Spain. The firms who only report in 1990 are dropped (248). After cleaning for missing values and firms with inconstant reporting (i.e. firms who leave and re-enter the sample) we are left with an unbalanced panel of 30,466 observations for 4,094 firms.

Variable	**Unit**
Family Ownership	Dummy
Foreign Ownership	Dummy
Production Methods used	Dummy
(CAD, Numerical Control, Robotics)	
Product Standardization	Dummy
Market Share	Percentage
Number of Competitors	Ordinal
[1 (<10); 2 (11-20); 3(>25); 4(Many)]	
Number of Products	Integer
Number of Markets	Integer
RD Expenditures	Euros
Product Innovation	Dummy
Process Innovation	Dummy
Employees	Integer
Long term debt (stock)	Euros
Cost of LT debt (stock)	Percentage
Long term debt raised (new debt)	Euros
Cost of LT debt raised (new debt)	Percentage
Equity	Euros
RD successfull financing	Dummy
Operational Profit	Euros
Patents (Spain)	Integer
Patents (External)	Integer
Capacity Utilization	Percentage

Table 7: List of variables