R&D AND ITS DETERMINANTS: A STUDY OF THE PHARMACEUTICAL FIRMS IN THE NETHERLANDS

SHREOSI SANYAL[1]* and MARK VANCAUTEREN[2]

[1]Hasselt University
[2]Hasselt University, Statistics Netherlands
*Corresponding author: shreosi.sanyal@uhasselt.be. Tel. +32 11 268648

Abstract

R&D investment is the key driving force for long-term economic growth and is of prime importance to the knowledge intensive pharmaceutical industry. Using a structural panel data framework of the Dutch pharmaceutical sector, this paper empirically investigates on how a firm's competitiveness, size, capital intensity and age affects its decision to invest in R&D. Our findings are based on the firm level data of the entire Dutch pharmaceutical industry for the period 1996-2006, which is obtained from the Central Bureau of Statistics database of Statistics Netherlands and the REACH database. For the empirical estimations, we adopt a generalized sample selection Tobit II model, as information on the R&D expenditure of all firms is not available. The maximum likelihood (ML) approach following Wooldridge (2005) is applied for handling the individual effects. In order to identify whether persistence in innovation exists, past information on patents are used in our analysis. We also pay special attention to the robustness of our Tobit II estimations by using Tobit I as an alternative technique. Our empirical findings suggest that young and small entrepreneurs with adequate capital reserve, and enjoying a higher degree of monopoly, are more likely to invest in R&D.

Keywords: R&D intensity; Firm determinants; Lagged patents; Tobit II; Maximum likelihood

JEL codes: L65, O31, C33

1. Introduction

In order to develop new ideas and technologies, firms undertake R&D investments which are inherently dependent on the degree of competition, determinants of capital market, the level of financial development and other controlling factors. This paper undertakes a firm-level study of the Dutch pharmaceutical industry in the light of R&D activities, and relates this variable to various firm characteristics such as, firm competitiveness, size, capital intensity, age and other determinants that reflect the technological capabilities of the firms and the extent to which they affect a firm's decision on how much to invest in R&D.

Our analysis is confined to the pharmaceutical industry alone, as it is known to be one of the most knowledge intensive sectors. Innovation in the pharmaceuticals is a complex process of building a nexus between knowledge applications and tackling the creativity of scientific clinical concepts through management expertise. The continual increase in the complexity and scale of pharmaceutical R&D expenditures are characterized by the emergence of complicated diseases that calls for novel treatment targets and enormous research activities in the pharmaceutical periphery. Therefore, rigorous R&D activity is imperative for this sector.

Regarding the R&D expenditure of the firms, it is a well-known fact that many firms do not report their R&D expenditures. For instance studies that employ R&D data provided by the European Community Innovation Surveys (CIS hereafter) take R&D availability as a starting point. As a result, the paucity of data on R&D investment will bias results based only on firms which report their R&D. We attempt to correct for this biasness using a sample selection model with censoring (following Heckman, 1979). Our model consists of two equations, where the first equation quantifies whether a firm reports its R&D expenditure or not and the second equation explains the amount of R&D invested, given that R&D expenditure is reported. This second equation is a regression equation with censoring at zero, implying that positive R&D is reported for the concerned firm on a given year.

Besides analyzing several market characteristics that affect R&D intensity, we have also specifically focused on the link between competition and innovation investment in the Netherlands' pharmaceutical firms. The question of whether competition hinders or bolsters expenditures in research activities has always remained under scrutiny, which dates back to the early Schumpeterian view that stiff competition may offer little room for innovation activities. This investigation gained greater momentum in recent times owing to better data availability and analytical precisions (as in Cohen, 2010 or Gilbert, 2006). Hence, this paper

provides evidence of the extent to which competitiveness affects a firm's decision on how much to invest in R&D expenditures, after controlling for traditional factors like size, capital intensity and age of firms.

In addition, our panel data allows us to analyze the dynamics of the innovation process. In order to comprehend if there exists a persistence of innovation at the firm level, we have incorporated a lagged patent dummy for finding the possible relation between past patenting activities to current R&D activities. This idea stems from the fact that, once a patent is granted, the firms may need to invest in R&D such that the patent can be transformed into a more commercial innovation for obtaining benefits. Hence, we intend to take a closer look on whether past patenting boosts further R&D expenditures for the persistence of innovation process.

The following section of this paper provides a quick review of the literature dealing with R&D and its determinants. Section 3 describes the data used for the analysis, and section 4 offers a brief overview of the methodological underpinnings of the empirical model. The empirical findings of different versions of the model explaining innovation activities are then discussed and contrasted in section 5. Finally, section 6 concludes.

2. Literature review

Investment in R&D is of prime importance to the volatile pharmaceutical sector, in order to innovate and develop new products, fostering long-term economic growth. The pharmaceutical industry is one of the most R&D intensive sectors (OECD, 2003), and the costs on R&D expenditure have surged over the last decade (DiMasi *et al*., 2003). There is a significant amount of work done on the valuation and drivers of R&D expenditure. Here we review some of the determinants of R&D.

According to the Schumpeterian tradition, the size of firms acts as an important explanatory variable where the general consensus of bigger sized firms inducing higher innovation prevails. This Schumpeterian finding was backed by several subsequent studies which include Bound *et al*. (1984) and Mairesse and Mohnen (2002) among others, which explains the idea that bigger firms possess greater monetary stock to invest in R&D. According to Cohen and Klepper (1996a), large firms are more incited to engage in innovative activities as they can amortize these costs by selling more units of output. Moreover, larger firms are expected to have a greater stock of knowledge base, and hence, expected to be more innovative than

smaller firms. Statistical evidence of this relationship is also provided by Nilsen and Schiantarelli (2003), which includes much greater incidences of zero investments in small firms as compared to large firms. More recent works of Hennessy and Whited (2007) argue that large firms face lesser cost of external finance as compared to small firms, considering a fixed cost of investment. Hence large firms have greater propensity to invest in R&D. In the context of pharmaceutical research, studies distinguishing research budgets per programme held within the firm and firm size conclude that, there is a significant size advantage. For example, Henderson and Cockburn (1996) had found that, drug discovery programmes that are carried out in larger firms seem to significantly correlate with higher levels of innovation.

Despite having several studies that support the Schumpeterian hypothesis, there exists an array of research that portrays contrasting and divergent findings. In the early studies by Hamberg (1964) and Comanor (1967), a weakly decreasing relationship between R&D intensity and firm size was found. Likewise, in the context of individual pharmaceutical research programs, Jensen (1987) had posited that the R&D expenditures exhibit decreasing returns to scale. This may be due to the role played by complementary assets in innovation in the pharmaceutical sector. On the other hand, Scherer (1965) asserted that, R&D intensity increased with firm size up to a certain intermediate level, after which it decreased. Likewise, Bound *et al*. (1984) found significant non-linearity in the relation between R&D expenditure and firm size, wherein both very small and very large firms are more R&D intensive than average sized firms. According to Cohen *et al*. (1987), firm size has a very small, statistically insignificant effect on R&D intensity when either fixed industry effects or measured industry characteristics are taken into account. Numerous existing studies also infers that smaller sized firms are more prone to innovation. Acs and Audretsch (1991) finds that small firms contribute 2.4 times greater innovation per employee than large firms. According to Akcigit (2009), with the increase in size of the firms, a lesser percentage of their revenue is allocated in R&D activities. In view of the fact that, the resource base is the main driving force for the invention of new chemical compounds in a pharmaceutical firm, the resource base may be independent of the firm size. Moreover, the upsurge of venture capital markets instigates smaller firms to invest in R&D (Enzing and Kern, 2006). Besides, R&D activities can also take place in external research units outside the firm which therefore has little effect on the size of the firm (Symeonidis, 1996).

Furthermore, the variable age and the entry-exit dummies are likely to shed some light on the stability and potentiality of the industry. The technology and products of industries evolve in

accordance with the innovations that are introduced by the entrant, surviving and incumbent firms. It is generally assumed that as firm ages, they have a greater propensity to expand their capital investment and skilled labor force, thus increasing the R&D expenditure. This accrues from the idea that, as a firm remains in the industry for a longer time, it establishes a history of performance. Consequently outsiders become better informed about the firm's ability to succeed in R&D, such that the adverse effect of capital market imperfections gets abated over time. However, on the contrary, young firms may be more dynamic and exhibiting greater interest to engage in R&D activities in order to survive in the industry. Prusa and Schmitz (1994) studied the software industry and purported that younger firms are more innovative than their older counterparts. In addition, intangible resources in the form of research employees, can leave an established firm to implement their knowledge on start-ups, thereby launching spinoffs (Klepper and Thompson, 2010). As sunk cost deters new entrants in the market (Sutton, 1991), the entry of new firms in the pharmaceutical industry may be hindered due to the huge sunk cost in the pharmaceutical R&D and the probable failure of potential new drugs during clinical trials. Hence, based on this argument, the feasibility of younger pharmaceutical firms to indulge in the risk of R&D investment should be less. However, as discussed before, with the emergence of a well-developed private equity market in Netherlands, the availability and access to venture capital has increased recently (Enzing and Kern, 2006), which encourages not only the smaller sized firms but also the younger ones (usually the smaller sized firms are the budding ones in the market).

In regard to the entry-exit of firms, the works of Audretsch (1995) and, Huergo and Jaumandreu (2004) asserts that, entry of a firm is contemplated as the way in which firms explore the value of new ideas under uncertainty. As pointed out by Malerba and Orsenigo (1996), the turbulence in the market caused by the entry and exit of firms may result in lack of innovativeness. This, in a way, annotates that higher competition caused by greater market turbulence diminishes innovation.

Regarding the topical research on R&D-competition relationship, product market competition is assumed to play a significant role in determining the extent of R&D expenditure. In the early empirical literature, Schumpeter (1943) estimated linear cross sectional relationship and typically found a negative relation between competition and innovation, confirming the theoretical and anecdotal evidence of that era. In consonance with Schumpeter, Blundell *et al*. (1999) found a positive relation between market share and innovation. On the contrary, a prominent number of research works since then, spearheaded by Nickell (1996) and Blundell

et al. (1999), found a positive effect of competition on innovation with linear specification estimations. However, work by Scherer (1967) and subsequently by Aghion *et al.* (2005), allowed for additional non-linearities in a cross-sectional analysis and discovered a significant inverted U-shaped relation between them. A study by Danzon (1997) posited that the pharmaceutical industry is a monopolistically competitive market. Hence, the marginal cost is far lesser than the price in the short run, but becomes almost the same in the long-run, such that the economic profit is minimized. Glover (2002) explains that pharmaceutical companies resort to a high degree of intellectual property right protection for creating a secretive regime, as R&D investment in this sector involves huge cost, time and uncertainty.

Another concomitant determinant that affects the extent of R&D investment is the capital intensity that the firms possess. The capital market imperfection is a concern for both industry practitioners and policymakers, leading to financial constraint and consequently reducing investment in innovation below its desired level. Based on the study by Hottenrott and Peters (2011), higher capital intensity as reflected by a firm's overall collateral value, reduces the likelihood of a firm facing binding constraint, which results in more expenditure in R&D. This finding is congruent with the early empirical work by Bound *et al*. (1984) who postulated that, there is a highly significant complementarity between R&D intensity and capital intensity, which increases when the selectivity bias of R&D intensity was corrected. However, although the mainstream pharmaceutical industry is capital-intensive, it is constrained by high regulatory hurdles. This causes the profitability of any particular product a long term prospect.

This paper further sheds light on whether past patenting activities boost R&D investments such that the persistence in the innovation process in maintained. While Romer (1990) has assumed that innovation is persistent at the firm level to a very large extent, Aghion and Howitt (1992) has put forth the idea of the creative destruction process which leads to a perpetual renewal of innovation. According to Cefis (2003), the empirical knowledge about the dynamics in firms' innovation behavior is a tool to access various growth models. Economic theory provides different explanations of why innovation might demonstrate a true state dependence over time. In the early works of Mansfield (1968) and Phillips (1971), the "success breeds success" hypothesis has been emphasized. The second argument is based on the idea that knowledge accumulates over time (Nelson and Winter, 1982). According to evolutionary theory, technological capabilities are a decisive factor in explaining innovation. Firms' technological capabilities are in turn, determined by human capital. Since a firm's

absorptive capacity is a function of the level of knowledge, learning in one period will further permit a more efficient accumulation of external knowledge in the subsequent periods, thereby inducing state dependence in innovation behavior (Cohen and Levinthal, 1990). The third postulate is based on the fact that, if a firm decides to take up R&D, they have to incur start-up costs for the R&D department which is generally a sunk cost and irrecoverable (Sutton, 1991). Such sunk cost may prevent both entry and exit into R&D activities. Hence they prevent non-R&D performers from taking up such activities because the potential entrants have to consider the sunk cost in determining their prices. On the other hand, they represent a barrier to exit for established R&D performers because the R&D expenditure is not recovered when the firm stops R&D and the firm has to incur them again if it decides to re-enter.

In regard to the pharmaceutical industry, development of drugs is a risky affair as large number of drugs fail in the clinical trials resulting in huge sunk cost. Hence innovative pharmaceuticals are more susceptible to pursue their R&D efforts, also presumably causing a barrier to the incumbent firms. Besides, as pointed by Duflos (2006), the dynamics of innovation process is an innate feature of the pharmaceutical industry as patenting plays a preponderant role in drug innovation. The high importance of patenting in the pharmaceutical industry has also been asserted by Taylor and Silberson (1973), Levin *et al*. (1987) and Cohen *et al.* (1997). Hence, firms are prone to invest in R&D and subsequently patent their innovations, which in turn augment R&D investment in order to transform the patent into a more commercial innovation. This propensity is cyclical, and thereby plausible persistence in the innovation process exists. In the later works like Peters (2007, 2009), Raymond *et al.* (2009), Antonelli *et al.* (2010) also confirmed persistence in the innovation process.

3. Data

For this study, we employ a purpose built dataset based on a panel of firms located in the Netherlands, with annual data from 1996 through 2006. A number of data sources were employed to compile our panel dataset, which includes the Reach database of Bureau van Dijk and; Community Innovation Surveys (CIS), R&D surveys and the General Business Register of Statistics Netherlands (CBS).

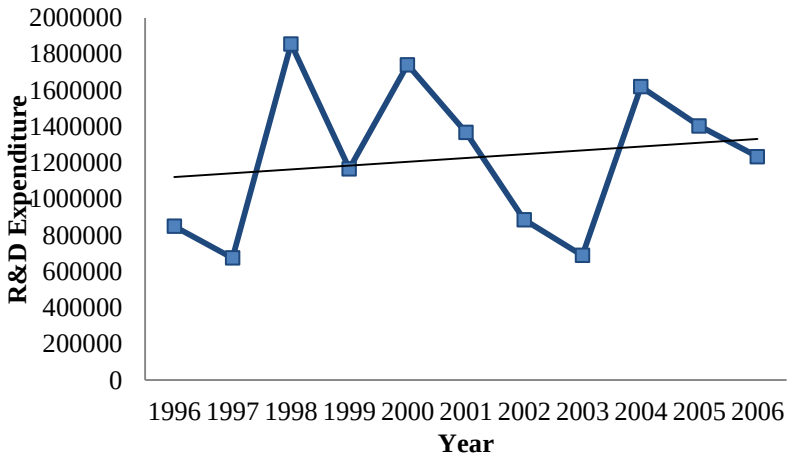
The Reach database provides detailed financial data including R&D expenditures and data on the ownership status of the firms in the Netherlands. We selected 673 firms that belong to NACE (Rev.2) sector code 21 (Manufacturing of Pharmaceutical products and

Pharmaceutical preparations). The ownership criteria to matching patents with firms are essential in the construction of the sample on patenting firms. Since firms register or report R&D expenditures under different names, the Algemeen Bedrijven Register (ABR) data, issued yearly by Statistics Netherlands, retrieves information on firms' ownership structure to find the names and the direct ownership (expressed in percentage) of all their subsidiaries, holding units, and their shareholders. In the sample of firms we define the possible (not necessarily ultimate) parent of the firm (enterprise) that is located in the Netherlands. By this selection, we obtained 520 firms.

However, we find that in the REACH database only a small proportion of firms publish their R&D expenditures. This relates to the fact that for accounting purposes, many firms combine their R&D expenditure with other related costs (i.e., general and administrative expenses) under the heading of intangible fixed assets or operational costs. We used two complementary R&D data sources. First, we used annual reports of online data of Dutch firms so to append any missing R&D data. Second, we extract R&D data from the Community Innovation Surveys (CIS) and R&D surveys that are collected by Statistics Netherlands. The R&D surveys report R&D expenditures in the odd years while each of the CIS waves measures R&D expenditures in the even years of our sample period.

From the compiled data, the trend in R&D expenditure over the period 1996-2006 is represented in fig. 1. The highest level of R&D expenditure is observed for the periods 1998, 2000 and 2004 with the amount of R&D expenditure as 1855888, 1740747 and 1619932 in thousands of euro respectively. The linear trend line shows a slight increase in the R&D expenditure over time. However the diagram illustrates the R&D expenditures of those Dutch pharmaceutical firms that has reported their R&D investment. Hence a complete picture of all R&D performing firms cannot be demonstrated, owing to the limitations in the data.

Fig. 1: Trend in R&D expenditure in the Netherlands Pharmaceuticals



Information on output, value-added, net tangible fixed capital assets, sales, depreciation and wages, all expressed in thousands of euro, was extracted from Statistics Netherlands. From this raw data, we calculate our independent variables that explain the probability of reporting R&D. The various determinants that influences the level of R&D investment and has been used in our analysis includes, the number of employees (e_{it}), capital intensity (k_{it}), the age of the firms (a_{it}), market share using sales (ms_{it}), market share using employees (me_{it}), firm-specific Lerner index (l_{it}), time dummies (α_k), entry dummies (β_k) and exit dummies (γ_k). In addition, the dynamics of the innovation process is accounted for by the use of a lagged patent dummy variable (p_k), which equals 1 if a firm possesses patents in any of the past years in the concerned time frame, and 0 otherwise.

Table 1 captures the statistical summary of the variables used in this study. It may be noted that, we have incorporated the R&D intensity of the firms instead of their total R&D expenditure in the following table, since the former is used for our empirical analysis, in order to account for the relative importance of R&D investment based on the manpower of individual firms. Similar is the economic rationale for using capital intensity. In addition, we have deflated the nominal values of these two variables using gross fixed capital formation price index from the EU Klems database (where 1995 is considered as the base year).

Table 1: Summary statistics

				<----------Quantiles---------->				
Variable	**Obs**	**Mean**	**Std.Dev.**	**Min**	**0.25**	**Mdn**	**0.75**	**Max**
R&D intensity	792	219.42	2149.45	0	0.07	2.43	9.7	36910
Number of employees	3880	814.09	3308.31	1	1	7	66	26575
Age	5676	9.33	11.9	0	0	3	16	39
Capital intensity	1704	1077.94	31818.83	0	4	18	55.94	1300000
Lerner index	1978	0.15	0.2	0	0.03	0.09	0.17	1
Market share using sales	1978	0.54	1.37	0	0	0.01	0.13	9.39
Market share using employees	3880	1.65	6.76	0	0	0.02	0.14	42.99
Lagged patent dummy	5720	0.05	0.22	0	0	0	0	1
Entry	5676	0.45	0.50	0	0	0	1	1
Exit	5676	0.46	0.50	0	0	0	1	1

It is observed that lagged patent dummy has no missing values, due to the fact that our data has complete information on the patenting activities of each pharmaceutical firm. Also, most of the values for entry and exit year of the individual firms could be extracted, resulting in very less missing values for age and entry-exit dummies. However, the R&D intensity of firms has only 792 observations, with the values spreading from 0 to 36910 units (expressed in thousands of euro). Therefore, in our panel data, the missing values for R&D intensity amounts to 4928, which is 86% of the total data. The size of firms, as depicted by the number of employees, has values ranging from 1 to 26575, indicating that our data encompasses pharmaceutical firms of all sizes. In other words, our dataset involves the Dutch pharmaceuticals in its entirety. The maximum age of firms till the latest year 2006 is observed to be 39 years. The capital intensity of firms has the maximum value of 130,0000 thousand euro. Lerner index, market share and the entry-exit dummies are the different proxies for measuring competition, as used in our model. The Lerner index, as a measure of firm's market power, ranges from 0 to 1, with 0 denoting 'no competition' and 1 as 'perfect competition'. The market share is a concentration measure of competition, which is calculated using sales as well as number of employees as alternative techniques. The former has values ranging from 0 to 9.39 units, whereas the latter has values from 0 to 42.99 units.

An in-depth survey of our R&D data reveals that 191 firms report their R&D expenditure for at least one year from 1996-2006, of which 133 firms have non-zero R&D for minimum one year of the sample period. The non-R&D reporting firms are 329, that never reported their R&D expenditure throughout the sample period. Considering also the firms that report zero R&D expenditure for the concerned 11 years' time span, the total number of firms with zero and missing R&D adds to 387. Table 2 represents the number of R&D performing and non-R&D performing (or non-reporting R&D firms) for each firm size categorization. In this

context, the R&D performing firms are those that report a positive R&D investment at least once in the 11 years sample period. On the other hand, the R&D non-performing or non-reporting firms resembles those pharmaceuticals that has missing or (/and) zero R&D data for the entire period under consideration.

Table 2: R&D investment based on different firm size segregation

Number of Employees	R&D performing firms	R&D non-performing/ non-reporting firms
≤ 20	9	199
> 20 and ≤ 50	8	52
> 50 and ≤ 100	11	34
>100	105	102

However the evidence provided in the table cannot lead us to draw a concrete inference, since some non-R&D reporting firms can be R&D-performing firms. Therefore table 2 is merely a representation of the data at hand and is not conclusive.

On competition and innovation

Competition in the pharmaceutical industry prevails not only on the introduction of newly invented drugs into the market, but also on the imitative drug therapies. Within the periphery of a given therapeutic class, the family of medicines goes through a well-delineated life cycle. Amongst the inventive pharmaceuticals, dynamic competition can be witnessed in the breakthrough of new molecular entities, as well as in its incremental advances towards consumable drugs. With the expiration of patents from the invented products, imitative competition from the generic firms permeates. Consumers benefit from such competition due to the significant lowering of prices. Altogether, the existence of competition begets prominent social returns and consumer surplus.

Concerning the competition measures that are used as explanatory variables (me_{it}, ms_{it} and l_{it}), the market share indices reflect the concentration in the market whereas the LI indicates the market power of the firms. Within the periphery of the production line of individual firms, market share acts as a formative indicator of the trend in its growth in the long-run. However the LI encompasses a wider perspective, where the geographic and product boundaries of a market for the operating firm do not bear any significance.

The market share of each firm using sales is derived as the ratio of the sales of a firm over the total sales of the whole industry for time *t*. Following Nickell (1996), we correct for the

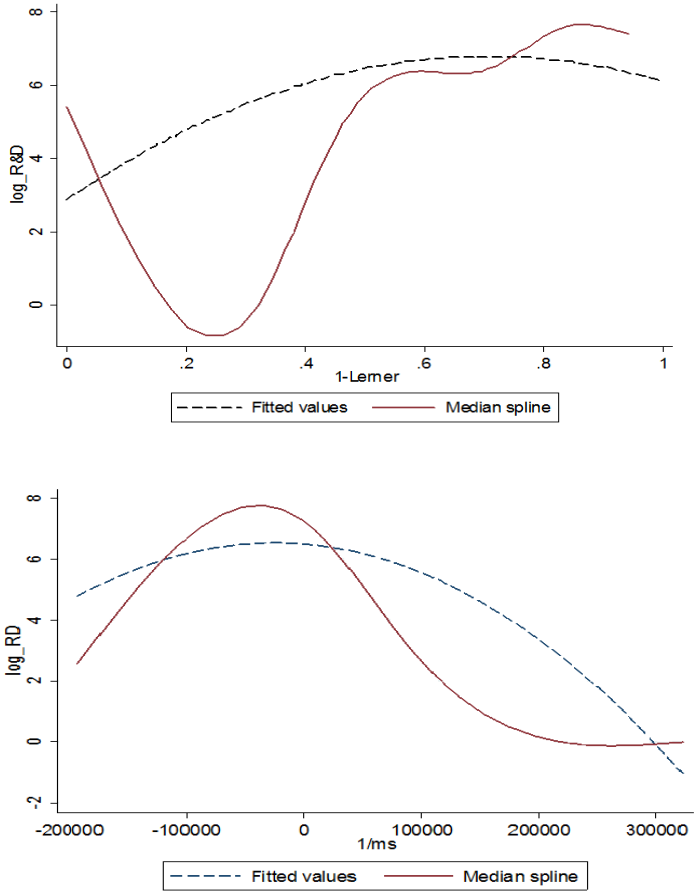
changing number of firms in each year. This is done by considering the total sales of firms as the average sales of a firm at time t, multiplied by the total number of firms chosen in the base year 1996. An alternative metric for market share measure is supplemented, where we used number of employees instead of sales. This measure of employment concentration to gauge competition and agglomeration is found in the study by Martin *et al.* (2010).

The PCM measures the profitability of firms, that is, the firm's ability to set its price above marginal costs. Hence the logic behind this measure is, if there are many competitors in a market with a low level of demand, then competition forces the firms to reduce prices until marginal costs. At the other extreme, a monopolist experiences no competition and hence can set a high price to maximize his profit. Thus, PCM falls in the range of no competition to perfect competition. We follow Aghion *et al.* (2005) for measuring PCM, which is termed as the Lerner Index. It is obtained as the difference between value added and total wages divided by total sales.

The Lerner index is considered to provide a more accurate inference than the concentration indices in the context of the output reallocation effect. Rojas (2011) asserted that the Lerner index is an appealing measure of competition as it specifies the positioning of market power of a firm within perfect competition and monopoly. Moreover, it testifies the role of demand elasticity in determining a firm's mark-up.

To shed some light on the relationship between competition and innovation, an inspection of the data suggests that there exists an inverted U-shaped relation between competition and the logarithmic value of R&D intensity, thus supporting the findings of Levin *et al.* (1985) and Aghion *et al.* (2005). This is depicted by the quadratic prediction plots in fig. 2. In the figure, competition (as 1-Lerner or 1/market share) is plotted in the x-axis against log of R&D in the y-axis. It is observed that, in both the cases, a non-linear inverted U-shaped curve is achieved.

Fig. 2: Quadratic Prediction plot (with median spline) of log(R&D) vs. Competition



Along with the market share and Lerner index, entry and exit dummies also essentially calibrate the level of competition in a market, through the turbulence caused by the inflow and outflow of firms. In this study, the entry-exit dummies are incorporated for analyzing how survival mechanisms influence heterogeneous mechanisms of R&D activities. The data on entry-exit is extracted from the ABR database. For the entry dummy, those firms which have entered the pharmaceutical market within the concerned period is assigned 1 and 0 otherwise. Likewise the exit dummy is calculated.

4. Empirical methodology

In the section on data analysis, it has been observed that 86% of data on R&D intensity is missing with only 133 out of 520 firms reporting R&D at least once within the period from 1996-2006. Since information on the R&D expenditure of most of the firms is not present in our dataset, we adopt a generalized sample selection Tobit II model applied to a panel data context, which is in line with the Heckman's two-step estimator model (Heckman, 1979). The insight of this approach is to solve the omitted variable problem, and consequently, to evade

the sample selection bias. Artés (2009) has applied the Heckman Tobit selection model, where the relation between R&D activity and market share is analyzed based on the short run and long run strategies of the Spanish firms, using a panel data.

The Tobit II model consists of a system of two equations. The first equation determines the probability of R&D so that possible selection bias can be corrected and the second equation determines the amount of R&D invested. The selection criterion for the panel data is such that we use data on firms that report R&D and compute the predicted R&D for those firms which do not report their R&D effort. In this framework we assume that the effect of no-R&D reporting firm is the same as R&D reporting firms. Since we distinguish between zero R&D and non-reporting R&D, we also assume that some non-innovating firms maybe R&D performers.

Let firm *i*'s R&D innovation effort at time *t* be written as:

$$R\&D_{it} = R\&D_{it}^{*} \; if \; R\&D_{it}^{*} = \alpha_{1i} + \beta_{1}^{'} X_{1it} + \varepsilon_{1it} \geq 0, \\ = 0 \; otherwise \qquad (1)$$

where $R\&D_{it}^{*}$ is a latent variable representing the firm's effort in R&D, α_{1i} is the firm specific unobserved heterogeneity, X_{1it} is a vector of independent variables and ε_{it} is a random error. Thus, we observe $R\&D_{it} = R\&D_{it}^{*}$ when $R\&D_{it}^{*} \geq 0$, i.e., when firm engages in R&D expenditure in year *t* and $R\&D_{it} = 0$ when $R\&D_{it}^{*} < 0$.

The second equation includes a binary variable z_{it} that is equal to one when R&D is reported for firm *i* in year *t* and zero otherwise. This can be written as follows,

$$z_{it} = 1 \; if \; z_{it}^{*} = a_{2i} + \beta_{2}^{'} X_{2it} + \varepsilon_{2it} \geq 0, \\ = 0 \; otherwise \qquad (2)$$

where z_{it}^{*} is the corresponding latent variable, a_{2i} is the firm specific heterogeneity, X_{2it} is the second vector of independent variables and ε_{2it} is a corresponding error term. Since the vector X_{1it} is not equal to X_{2it} we allow for an exclusion restriction which is typical for a sample selection model (see for example, Vella, 1998). Thus, sample selection arises because the observation on R&D is conditional on being observed, that is, conditional on $z_{it} = 1$. The

sample selection assumes that all sampled firms are probable R&D innovators, but only the firms where $R\&D_{it} \geq 0$ report this effort.

The model is completed by the assumption that the unobserved errors ε_{1it} and ε_{2it}, conditional upon X_{1it} and X_{2it}, follow a bivariate normal distribution having zero mean, variances $\sigma_1^2 (=1)$ and σ_2^2, and covariance $\sigma_{12} = \rho\sigma_2$, where $\rho = \text{cov}[\varepsilon_{1it}, \varepsilon_{2it}]$. The vector of independent variables X_{it} includes e_{it}, me_{it}, ms_{it}, l_{it}, k_{it}, a_{it}, α_k, β_k, γ_k and p_k, where the last four regressors represent categorical variables. It can be noted here, that the non-categorical variables are used in their logarithmic form in our regression estimations. These explanatory variables are incorporated in the main equation (as X_{1it}) or (/and) selection equation (as X_{2it}).

Since our data sample is a panel data, we apply the maximum likelihood approach following Wooldridge (2005), which enables us to exploit the unobserved heterogeneity dimension at the individual firm level. This approach assumes that the unobserved effects $(a_{1i} \,\&\, a_{2i})$ are distributed as follows for handling the individual effects,

$$a_{1i} = \alpha_{10} + \delta_{10} R\&D_{i0}^{*} + \delta_1' \overline{X}_{i1} + \xi_{1i}, \tag{3A}$$

$$a_{2i} = \alpha_{20} + \delta_{20} z_{i0} + \delta_2' \overline{X}_{i2} + \xi_{2i} \tag{3B}$$

where α_{10} and α_{20} are constants, $\overline{X}_{i1}, \overline{X}_{i2}$ are the vectors which includes the time averages of the variable $(e_{it} \text{ and } a_{it})'$, $R\&D_{i0}^{*}$ and z_{i0} are the initial values of log of R&D intensity and R&D probability respectively, δ_{10}, δ_{20}, δ_1' and δ_2' are the corresponding coefficients (vectors) to be estimated, and ξ_{1i} and ξ_{2i} are assumed to be independent, following normal distributions $\xi_{1i} \mid X_{i1} \sim N(0, \sigma_{\xi 1}^2)$ and $\xi_{2i} \mid X_{i2} \sim N(0, \sigma_{\xi 2}^2)$.

Although the constrained version of Wooldridge (2005) used the time averages $\overline{X_i}$, which allows for a reduction of explanatory variables, the within-means of the independent variables in this approach can be highly biased. This is due to the fact that, it includes the explanatory variables of all concerned time periods, including the initial period (as in Hesketh and Skrondal, 2013; Akay, 2012; and Conti and Pudney, 2011). As stated by Hesketh and

Skrondal (2013), the shortcoming to this constrained model is the direct dependence of the conditional distribution of the unobserved effects on the initial period explanatory variables rather than the explanatory variables of the other periods. In some cases, it depends solely on the initial period explanatory variables and the initial dependent variables causing a serious problem of biased results. The two probable solutions put forth in this study is either including the initial period of explanatory variables as regressors along with their within means of all periods; or excluding the initial period explanatory variables from the within-means. In our empirical analysis, we opt for the latter solution where we omit the initial period explanatory variables from the within means. Hence, equations (3A) and (3B) can be re-written as,

$$a_{1i} = \alpha_{10} + \delta_{10} R \,\&\, D^{*}_{i0} + \delta_1' \overline{X}^{'}_{i1} + \xi_{1i}, \qquad \text{(3A)'}$$

$$a_{2i} = \alpha_{20} + \delta_{20} z_{i0} + \delta_2' \overline{X}^{'}_{i2} + \xi_{2i} \qquad \text{(3B)'}$$

where, $\overline{X}^{'}_{i} = \dfrac{1}{T-1} \sum_{t=2}^{T} X_{it}$

As a robustness check, we have also applied the standard Tobit-I model. The Tobit-I is a special case of the Tobit-II model in which $R \,\&\, D^{*}_{it} = z^{*}_{it}$. This censored normal regression model performs the censoring from below at zero, and no transformation of the dependent variable occurs in this case.

Papers like Audretsch (1995b) and Klepper (1996) provide theoretical insights into the nature of the dynamics of innovation process. As mentioned in Section 3, we have used a lagged patent dummy in both the Tobit I and Tobit II analyses, in order to investigate on how past patents affect future R&D process in the Netherlands pharmaceutical industry.

5. Empirical estimation

We discuss in this section the empirical results that relate R&D investment to its various determinants by using the Heckman's sample selection Tobit II estimation technique. We first look at the results of the static model, and in the next subsection we develop our model by incorporating dynamics. Finally we perform a Tobit I estimation as a control or robustness check.

5.1 Results of the static model

Table 3 represents our Tobit II estimations of the Heckman selection model, where four static models have been estimated. Each column in the table enumerates individual Tobit II estimations, comprising of the coefficients from the selection equation and the outcome equation. In the Heckman model, the outcome (or Tobit) equation is the equation of interest and hence, its results will be substantially explained. On the other hand, the selection (or Probit) equation serves for the purpose of only the selection process, in which the interpretation of its regression coefficients depends on the observed response variable (in our analysis, it is probability of R&D intensity) to take the value of either 0 or 1. In our model, we allow for the exclusion restriction, such that $X_{1it} \neq X_{2it}$. Based on the exclusion restriction, we have used the entry-exit dummies in the selection equation that is excluded from the outcome equation, which might probably reduce the problem of collinearity to a considerable extent. However, we have not incorporated in the outcome equation any extra explanatory variable that is not included in the selection equation for robustness reasons (similar to IV estimation method). In this regard we follow Wooldridge (2010), where it is mentioned that the independent variables in the outcome equation should be a strict subset of the variables included in the selection equation. Our estimations are done in STATA, which uses a maximum likelihood procedure for estimating both the selection and the outcome equations.

In regression model 1 (as represented by column 1), we assume no random effects. However in models 3, 4 and 5, we include the initial values as well as the averages of the explanatory variables in order to control for unobserved heterogeneity which accounts for full random effects. In the first three regression results, market share using sales is solely used as the competition measure. However in model 4 and 5, both market share and Lerner index has been used as competition measures. But in column 5, we have substituted market share using employees (instead of sales) as an alternative concentration measure of competition. In all the estimation models, entry and exit has been used as exogenous variables that are not included in the outcome equation. This exclusion restriction has been accounted for to assume more robust identification of our model. The rest of the explanatory variables along with time dummies have been used in both the selection and outcome equations in all the regression models in Table 3.

From the results obtained in Table 3, it is seen that the coefficient of firm size has a systematically negative and significant effect on R&D intensity. This suggests that smaller

pharmaceutical firms are more dynamic in fostering new innovation activity in order to promote growth. This finding is in stark contrast with the Schumpeterian hypothesis which suggests that larger firms undertake greater R&D activity. Although several studies approve the Schumpeterian hypothesis, there also exists a number of research works where no such relation has been found between them (as in, Klomp and Van Leeuwen, 2006 and Benavente, 2006). Although there exists an extensive empirical work on innovativeness and firm size, no conclusive inference have been obtained by far. Even in the works of Schumpeter, two contradictory views are found, where he asserts that 'entrepreneurs are most likely to innovate' and then contradicts with the statement 'large firms having some degree of monopoly are most prone to innovate (Afuah, 1998).

Nevertheless, several studies assert that firms invest smaller percentage of their revenue in R&D activities as they increase in size (like, Acs and Audretsch, 1988, 1991 and, Akcigit, 2009). Since smaller firms are funded by the venture capital markets, it may cause an important drive for R&D activities. Enzing and Kern (2006) has asserted that the pharmaceutical industry in Germany, France and Netherlands have a very well-developed private equity market compared to other EU countries, where the availability and access to venture capital market rose substantially since 1995. This encouraged smaller firms to take up their R&D activities more seriously. Also, the smaller pharmaceutical firms have a greater tendency to collaborate with research universities or other research units where the R&D activities of the pharmaceuticals are most often carried out. Further, as pointed out by Cohen *et al.* (1987), considering the business units are more relevant than the entire firm for calibrating firm size empirically, the size of a firm as a whole may not bear any significance.

However, the selection equation suggests that firm size has a positive and significant effect on the probability of R&D reported. Since the estimated dependent variable (probability of R&D reporting) in the selection equation takes either the value of 0 or 1, interpreting that the regression coefficient for the selection equation is complicated and does not yield any confirmatory inference. It may be noted that, in the findings of Crépon *et al.* (1998), the signs of the coefficients of number of employees for the selection equation and the outcome equation are positive and negative respectively in relation to R&D. Our results are similar to these findings, although the relation between R&D intensity and size was not significant in the outcome equation of Crépon *et al.* (1998) study.

Concerning the age of the Dutch pharmaceutical firms, a positive and significant effect on R&D intensity is observed when no random effect is assumed (in column 1). However, when we control for unobserved heterogeneity by assuming full random effects, age seems to have a negative relation with R&D intensity. This effect is found to be significant in the regression result of column 3. But it loses its significance when Lerner index is incorporated as an additional regressor in the last two regressions (column 4 and 5).

Hence, if we consider the effect of age on R&D after controlling for unobserved heterogeneity, it reveals than younger firms are more prone to R&D activities. Dujowich (2013) reveals that smaller and younger firms are more susceptible to R&D, which is in congruence with our findings. However the effect of age on R&D intensity does not provide a concrete result as the negative effect is insignificant in the last two regression results. Plus, it has a positive relation when full random effect is not assumed. For the probit equation, age exhibits a negative effect on the probability to report R&D, although the effect is not significant with very low elasticity.

We find that capital intensity has a positive and significant influence on R&D intensity, as exhibited in the outcome equations. This confirms a positive complementarity between capital intensity and R&D intensity of the Netherlands pharmaceuticals. This finding is in line with Bound *et al*. (1984), Hottenrott and Peters (2011) and many other literatures that unanimously asserted that, an increase in capital intensity encourages the adoption of new technologies, resulting in the increase in R&D expenditure.

In our analysis, market share has been used as the concentration measure to indicate the level of product market competition in all the regression models of Table 3. It is found to have a persistently positive and highly significant effect on R&D intensity and the probability to invest in R&D in the outcome and selection equation respectively. Market share using employees instead of sales has been used as an alternative measure of concentration in the last regression model. It is also found to be positive and significant in both the equations. Hence more concentration in the Netherlands pharmaceutical market causes more R&D investments. This confirms the Schumpeterian conjecture (Schumpeter, 1943) that, lower level of competition causes higher level of R&D intensity, and vice-versa.

In addition, the logarithmic value of the Lerner index is used as an ancillary to market share for measuring the level of competition in the last two regression analysis (column 4 and 5). The Lerner index exhibits a significantly positive effect on R&D intensity in the regression

result of column 4. However, it is no longer significant in the final regression analysis. The only difference between the last two regression analysis is the metric used for measuring market share, where the market share in the last equation is measured using the share of employees instead of sales. In case of the probit equations, it shows significant and positive effect in both the models.

An increase in Lerner index demonstrates higher market power, and thereby a lowering of competition. Hence, similar to the concentration measures used, the effect of the Lerner index on R&D explains the fact that lesser competition among the Dutch pharmaceuticals induces greater R&D intensity. A possible explanation for finding a negative relation between competition and R&D is that, firms generate higher innovation incentives due to larger monopoly profits which benefit the technology, depicting a negative relation between competition and innovation. However, it has not been examined in our empirical analysis if there exists a nonlinear relation between competition and R&D investment. Although the quadratic prediction plots in section 3 suggest an inverted U relationship between them, the finding was only suggestive, as no consideration was made for the non-reporting R&D firms.

Finally, the entry and exit dummies are included in the selection equation as additional regressors. A positive coefficient is obtained for entry dummies, reflecting a positive effect on the probability to report R&D. But the result is not deterministic as it is not significant in any of the models. Nevertheless, the exit dummy is found to be significantly negative on the probability of R&D reported when random effect is assumed. This might hint at a contradiction as more exit of firms may hint at a greater concentration which is incongruous to our earlier finding that more concentration causes greater R&D. However we cannot reach a clear consensus or a substantially justifiable inference with the effect of entry and exit dummies as they are only included in the selection equation and the effect is only on the probability of R&D reported. In practical viewpoint, exiting firms would have lesser likeliness to report on their R&D activities. Anyhow, since the entry dummy is nondeterministic and the exit dummy is negative, it is probable that sunk cost in pharmaceutical research investments may prevent both entry and exit of firms that are innovation intensive (Sutton (1991).

In addition to the estimation of the two equations, the Heckman model estimates ρ (actually the inverse hyperbolic tangent of ρ) which is the correlation of the residuals of the two equations, and σ (actually the log of σ) is the standard error of the residuals of the R&D

equation. The λ is the Inverse Mill's Ratio which is the product of ρ and σ. The Inverse Mill's Ratio is used by the Heckman's sample selection model to estimate the outcome equation. In the last three regression models we find ρ to be significant, which implies that we can reject the null hypothesis that ρ=0 and lies within the confidence interval. Additionally, both σ and λ are also found to be significant in the last three models. Hence the sample selection model performs well in the last three regressions and we consider them as the preferred models.

Table 3: Static Tobit II Estimations

Dependent variable	Probit (R&D =0/1)	Log of R&D per employee	Probit (R&D =0/1)	Log of R&D per employee	Probit (R&D =0/1)	Log of R&D per employee	Probit (R&D =0/1)	Log of R&D per employee
Log(Employees)	0.409***	-0.532***	0.436***	-0.595***	0.428***	-0.645***	0.406***	-0.606***
	[0.030]	[0.089]	[0.043]	[0.096]	[0.044]	[0.088]	[0.044]	[0.083]
Log(Age)	-0.012	0.140*	-0.057	-0.234**	-0.059	-0.146	-0.060	-0.152
	[0.043]	[0.073]	[0.044]	[0.105]	[0.044]	[0.105]	[0.042]	[0.104]
Log(Capital Intensity)	0.107***	0.192***	0.075***	0.176**	0.063**	0.137*	0.055**	0.178**
	[0.026]	[0.057]	[0.027]	[0.076]	[0.026]	[0.074]	[0.025]	[0.073]
Market Share Using Sales	0.202***	0.205***	0.093	0.537***	0.144*	0.564***		
	[0.057]	[0.057]	[0.063]	[0.071]	[0.081]	[0.069]		
Market Share Using Employees							0.057**	0.083***
							[0.025]	[0.010]
Log(Lerner Index)					0.241***	0.229**	0.220***	0.029
					[0.068]	[0.106]	[0.066]	[0.109]
Entry	0.201		0.255		0.153		0.075	
	[0.159]		[0.161]		[0.159]		[0.157]	
Exit	-0.103		-0.295**		-0.327***		-0.336***	
	[0.120]		[0.127]		[0.127]		[0.120]	
Intercept	-2.786***	3.768***	-2.722***	4.192***	-2.184***	4.846***	-2.083***	4.138***
	[0.201]	[0.848]	[0.239]	[0.696]	[0.276]	[0.608]	[0.269]	[0.558]
Time Dummies	YES		YES		YES		YES	
Initial[Log(R&D/Employee)]				0.428***		0.456***		0.594***
				[0.037]		[0.038]		[0.043]
Initial[R&DProbability]			0.797***		0.928***		0.929***	
			[0.119]		[0.125]		[0.123]	
Random Effects	NO		YES		YES		YES	
Log-likelihood	-1250.001		-1177.835		-1166.158		-1167.778	
ρ	-0.185		-0.596***		-0.681***		-0.768***	
	[0.203]		[0.135]		[0.101]		[0.088]	
σ	1.473***		1.440***		1.457***		1.512***	
	[0.054]		[0.070]		[0.068]		[0.075]	
λ (Inverse Mill's Ratio)	-0.272		-0.858***		-0.993***		-1.161***	
	[0.304]		[0.226]		[0.183]		[0.179]	
N Observation	1436		1436		1436		1436	
Censored	995		995		995		995	
Uncensored	441		441		441		441	
Estimation Method	**Heckman Tobit II Sample Selection**							

*** denotes 1% significance level, ** denotes 5% significance level and * denotes 10% significance level

5.2 Extension to dynamic model

The dynamics of a firm's innovation behavior is an important assumption for endogenous growth models (Aghion and Howitt, 1992). Past research studies that considered the patent-R&D relationship show evidence of persistence in innovation. Leeuwen (2002) and Raymond *et al.* (forthcoming) have investigated the dynamic relation between patent and R&D in reference to Netherlands. Both studies confirm a persistence in innovation over time. In the work of Raymond *et al*. (forthcoming), individual effects are accounted for and modeled in a dynamic Tobit II panel selection model. It was found that, with respect to the input innovation, past R&D/sales expenditures affect current R&D activities and this dynamic relationship also holds with respect to output innovation, i.e, share of innovative sales to total sales. This result is also confirmed by the study of Van Leeuwen (2002), where innovation input (R&D expenditures/ sales) is linked to innovation output (share of innovative sales/ total sales) and the innovation output to firm performance (revenue/employee). However, a major drawback in the latter study is that individual effects are not accounted for.

It is assumed that in the R&D equation, the past level of patents, represented by a patent dummy (as a measure of innovation output) affects current R&D expenditure (as a measure of innovation input). In the pharmaceutical industry, once a patent is granted, firms may need to invest in R&D so that they can transform the patent into a more commercial innovation in order to accrue benefits from it. This in turn helps the R&D sector to develop other novel inventions which can be patented, thereby maintaining a dynamics in the innovation process. Hence to account for the dynamics in the innovation process, we use lagged patent dummies.

The results, after the incorporation of dynamics using lagged patent dummy in our framework, have been summarized in Table 4. We have assumed full random effect (by including the initial values and averages), to control for unobserved heterogeneity in a dynamic setting of both the regressions in Table 4. In both the models, the lagged patent dummy is found to exhibit a positive effect on R&D intensity. But the effect is not significant in any of the two regression results. Similar results are obtained for the selection equation. Nonetheless, a positive coefficient between lagged patent dummy and R&D suggests, to some extent, a persistence in the innovation process, although the relation cannot be confirmed owing to the insignificant results.

Regarding the other explanatory variables, their effect remains the same on R&D intensity and they do not exhibit any prominent divergence from the results of the static model.

Summarizing the effect of these estimated coefficients in the outcome equation of the dynamic panel framework, the size of the firms and age has a negative effect on R&D intensity, where only the former variable is significant. The coefficients of capital intensity of the firms in the outcome equations are consistently positive and significant. The effect of market share and Lerner index is found to be significantly positive, again confirming a negative relation between competition and R&D investment. However, the Lerner index loses its significance in the second model of Table 4.

Table 4: Dynamic Tobit II Estimations

Dependent variable	Probit (R&D =0/1)	Log of R&D per employee	Probit (R&D =0/1)	Log of R&D per employee
Log(Employees)	0.422***	-0.654***	0.385***	-0.439***
	[0.045]	[0.088]	[0.050]	[0.102]
Log(Age)	-0.058	-0.144	-0.068	-0.154
	[0.044]	[0.105]	[0.045]	[0.109]
Log(Capital Intensity)	0.065**	0.138*	0.054**	0.235***
	[0.026]	[0.073]	[0.027]	[0.078]
Market Share Using Sales	0.145*	0.554***		
	[0.081]	[0.070]		
Market Share Using Employees			0.068**	0.047***
			[0.027]	[0.008]
Log(Lerner Index)	0.234***	0.205*	0.205***	0.076
	[0.070]	[0.110]	[0.069]	[0.113]
Entry	0.148		0.116	
	[0.158]		[0.164]	
Exit	-0.329***		-0.374***	
	[0.126]		[0.127]	
Patent dummy	0.077	0.144	0.067	0.230
	[0.181]	[0.190]	[0.187]	[0.192]
Intercept	-2.200***	4.848***	-2.065***	3.295***
	[0.277]	[0.600]	[0.279]	[0.758]
Time Dummies	YES		YES	
Initial[Log(R&D/Employee)]		0.455***		0.513***
		[0.038]		[0.043]
Initial[R&DProbability]	0.936***		0.965***	
	[0.126]		[0.128]	
Random Effects	YES		YES	
Log-likelihood	-1165.670		-1176.685	
ρ	-0.695***		-0.604***	
	[0.099]		[0.161]	
σ	1.463***		1.456***	
	[0.069]		[0.076]	
λ (Inverse Mill's Ratio)	-1.017***		-0.880***	
	[0.182]		[0.270]	
N Observation	1436		1436	
Censored	995		995	
Uncensored	441		441	
Estimation Method	**Heckman Tobit II Sample Selection**			

*** denotes 1% significance level, ** denotes 5% significance level and * denotes 10% significance level

Again keeping parity with the static model, the coefficients of the variables in the selection equation have the same signs and significance as the static model. The entry and exit

dummies are included in the selection equation as an exclusion restriction, where the former shows a positive and insignificant effect while the latter shows a negative and significant effect on the probability to do R&D. The value of ρ is found to be significant in both the cases which proves that the use of the sample selection model for this data is justified.

5.3 Robustness check and further issues

In order to check the robustness or sensitivity of our estimations, we perform the Tobit-I estimation, both in the static and dynamic framework, which is presented in Table 5. The first regression in Table 5 provides estimates of a static model, while the second regression introduces dynamics with lagged patent dummies. It is to be noted that for the market share measure we have used the measure using sales instead of employees in both the regression models, as no prominent changes were observed when market share using employees were used in our previous estimates. Hence we adhere to the market share using sales only, as our concentration measure.

Table 5: Tobit I Estimation

	Log of R&D per employee	Log of R&D per employee
Log(Employees)	-0.221**	-0.235***
	[0.090]	[0.091]
Log(Age)	-0.083	-0.075
	[0.150]	[0.150]
Log(Capital Intensity)	0.225**	0.227**
	[0.090]	[0.090]
Market Share Using Sales	0.260***	0.253***
	[0.065]	[0.065]
Log(Lerner Index)	0.055	0.045
	[0.099]	[0.100]
Entry	0.391	0.398
	[0.525]	[0.525]
Exit	-0.213	-0.167
	[0.250]	[0.255]
Patent Dummy		0.189
		[0.220]
Intercept	1.749***	1.721***
	[0.516]	[0.516]
Initial[Log(R&D/Employee)]	0.396***	0.395***
	[0.045]	[0.045]
Random Effects	YES	YES
Log-likelihood	-815.724	-815.356
N Observation	433	433

*** denotes 1% significance level, ** denotes 5% significance level and * denotes 10% significance level

From the estimated results, we find that, in both the static and dynamic estimations, firm size has a negative and highly significant relation with R&D intensity. The age of firms bears a

negative relation with R&D intensity, although the effect is insignificant. The capital intensity and market share of firms reveal a positive and significant relation, whereas the Lerner index has a positive but insignificant relation with R&D intensity. The entry-exit dummy, depicting the level of turbulence in the Dutch pharmaceutical market, provides insignificant estimates. However, the entry dummy and exit dummy has a positive coefficient and a negative coefficient respectively. Finally, the coefficient for patent dummy in the second equation is observed to be positive but insignificant. Although an innovation persistence is hinted at, it remains a suggestive inference owing to insignificance of the estimated result (similar to the Dynamic Tobit II estimations).

Altogether it can be concluded that the results of the Tobit I estimation in table 5 yields similar results as the Tobit II estimations in table 3 and 4, inspite of the difference in the estimation techniques (the former treats zero R&D expenditure as no R&D activity while the latter considers it as unobserved). Despite the similarity in the results of Tobit I and II, we prefer the Tobit II estimation technique as it represents a joint distribution for the censoring method and its possible outcome, and subsequently finds the implied distribution on the observed outcome (Cameron and Trivedi, 2009).

In addition, we have tested for the presence of endogeneity between product market competition and R&D intensity, using the structural model approach (as in Cameron and Trivedi, 2009). It is observed from the analysis that endogeneity is rejected at 5% level. Hence we can consider our estimates as robust.

6. Conclusion

The prime objective of this paper is to empirically investigate and elucidate how various economic determinants affect a firm's predisposition to engage in R&D investment, using a panel data of 520 Dutch pharmaceutical firms for the period 1996-2006. The structural model framework, estimated using a sample selection Heckman's Tobit II regression technique, disentangles the impact of competition measures, capital intensity, firm size and age of firms on the R&D investment.

By focusing on the substantive empirical results, smaller pharmaceutical firms are found to have greater inclination to engage in R&D. This negates the Schumpeterian hypothesis that, bigger firms are more conducive to R&D investment. However, several research have contradicted the Schumpeterian theory which includes Acs and Audretsch (1988, 1991),

Akcigit (2009) and Dujowich (2013). Owing to the huge cost of R&D investment in pharmaceuticals, collaborative strategic efforts between the private pharmaceuticals and public research units might be a strategic option to promote R&D investment in the Netherlands pharmaceuticals. This connotes to the idea that the research units of a firm is solely responsible for the R&D performance and not the entire firm size. Furthermore, the existence of smaller pharmaceutical spin-offs that have the expertise to undertake research initiatives can also engage in R&D activities. The presence of strong venture capital markets operating in the Dutch economy encourages these smaller firms to invest in pharmaceutical research. Additionally, the procurement of exclusive R&D tax incentives provided by the Dutch government, which includes special allowances and deductibles in pharmaceutical research, can play a major role in infusing R&D activities amongst the smaller sized pharmaceuticals.

Our empirical results also infer a negative relation between the age of firms and R&D intensity, which establishes that the young Dutch pharmaceuticals are more dynamic and have greater susceptibility to engage in R&D activities, in order to gain a foothold in the market. Hence the assertion by Dujowich (2013) that small and young firms attribute to more R&D activities is substantiated by our empirical findings.

On the other hand, the effect of capital intensity on R&D investment under ceteris paribus condition, is found to be consistently positive in our empirical results. This germane to the fact that pharmaceutical research incurs huge costs, for which it is quintessential for the firms to have adequate capital reserve for engaging in research activities. Drug discovery and development is a complex risky task involving clinical and preclinical trials, which involves huge technological and financial capacity.

Likewise, the concentration (as measured by the market share) and market power (as measured by the Lerner Index) of the firms has a systematically positive and significant relation with R&D intensity. Since greater concentration or market power implies lesser competition among the firms, our finding is found to be analogous to Schumpeterian viewpoint that lesser competition encourages more R&D activities. It is evident that the intellectual property right protection is fundamental in the pharmaceutical sector in their decision to invest in R&D. Intellectual property right protection causes monopoly power of the firms, thereby reducing the level of competition. In addition, competition among the drug manufacturers is also influenced by the insurance plans or brand valuation, and hence

institutional and regulatory frameworks largely determine the competition in this specific sector.

In the extension of our model to a dynamic framework, persistence in innovation is perceived, although the results are not deterministic. However the positive magnitude of the lagged patent dummy implies that past innovation output in the form of patenting can positively affect the present investment in R&D. But the insignificant result raises the question of the technological value of patents. Based on a perpetual inventory method, a 15% depreciation rate means that a patent value is close to zero after 20 years (Duflos, 2006). Hence the depreciation of patent value over the years can have little or no influence on the R&D investment.

Hence, we may infer from our main results that, smaller and younger Dutch pharmaceuticals engage in R&D investments in a non-competitive regime, which is presumably conditioned on their past innovation output. Young firms in a less competitive environment have the opportunity and time to build up the 'immunity' for themselves in the long-run, and are therefore encouraged to undertake plausible risks in R&D investments.

However, our results solely focus on the pharmaceutical industry within the geographical periphery of Netherlands. Therefore, the inferences drawn from this study may vary for different sectors or countries, due to the characteristic deviations in their respective technology and market conditions or governmental policies. Nevertheless, our research work is novel in the sense that, the role of various economic determinants to explain R&D in the Netherlands pharmaceutical industry is explored for the first time in a detailed analysis.

References

Akcigit, U. (2009), "Firm Size, Innovation Dynamics and Growth", Society for Economic Dynamics, Meeting Papers No. 1267.

Acs, Z.J. and D. Audretsch (1988), "Innovation in large and small firms", *American Economic Review*, **78**, pp. 678-690.

Acs, Z.J. and D. Audretsch (1991), "R&D, Firm Size, and Innovation Activity", in: Z.J. Acs and D. Audretsch eds., Innovation and Technological Change: An International Comparison, New York, NY: Harvester Wheatsheaf.

Afuah, A. (1998), "Innovation management: Strategies, Implementations and Profits", Oxford University Press: New York.

Aghion, P., and P. Howitt (1992), "A model of growth through creative destruction", *Econometrica*, **60(2)**, 323-351.

Aghion, P., N. Bloom, R. Blundell, R. Griffith and P. Howitt (2005), "Competition and Innovation: An inverted U relationship", *The Quarterly Journal of Economics*, **120(2)**, 701-728.

Akay, A. (2012), "Finite-sample comparison of alternative methods for estimating dynamic panel data models", *Journal of Applied Econometrics*, **27(7)**, 1189-1204.

Antonelli, C., F. Crespi and G. Scellato (2010), "Inside Innovation Persistence: New evidence from Italian Micro- Data", Laboratorio di economia dell' innovazione "Franco Momigliano", Universita di Torino, Italy, Working Paper No. 10/2010.

Artés, J. (2009), "Long-run versus short-run decisions, R&D and market structure in Spanish firms", *Research Policy*, **38**, 120-132.

Audretsch, D. (1995), "Innovation and Industry Evolution", MIT Press.

Benavente, J.M. (2006), "The Role of Research and Innovation in Promoting Productivity in Chile", *Economics of Innovation and New Technology*, **15(4/5)**, 301-315.

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

Bound, J., C. Cummins, Z. Griliches, H.B. Hall and A. Jaffe (1984), "Who does R&D and who Patents?", in Griliches, Z. (ed.), R&D, Patents, and Productivity, Cambridge, MA, 21-54.

Cameron, C.A. and P.K. Trivedi (2009), "Microeconometrics Using Stata", College Station: Stata Press.

Cefis, E. (2003), "Is there persistence in Innovation activities?", *International Journal of Industrial Organisation*, **21(4)**, 489-515.

Cohen, W.M. (2010), 'Fifty years of empirical studies of innovative activity and performance", in: Hall BW, Rosenberg N. (ed.), Handbook of the economics of innovation, Elsevier, Amsterdam, 129-213.

Cohen, W.M., R.R. Nelson and J. Walsh (1997), "Appropriability conditions and why firms patent and why they do not in the U.S. manufacturing sector", Carnegie Mellon University Working paper, Pittsburgh.

Cohen and Klepper (1996a), "Firm Size and the nature of Innovation within Industries: The Case of Process and Product R&D", *Review of Economics and Statistics*, **78(2)**, 232-243.

Cohen, W.M. and D.A. Levinthal (1990), "Absorptive Capacity: A New Perspective on learning and innovation", *Administrative Science Quarterly*, **32(1)**, 128-158.

Cohen, W.M., R.C. Levin and D.C. Mowery (1987), "Firm Size and R&D Intensity: A Re-Examination", NBER Working paper 2205.

Conti, G., S. Pudney (2011), "Survey design and the analysis of satisfaction", *Review of Economics and Statistics*, **71(5)**, 659-681.

Crépon, B., E. Duguet and J. Mairesse (1998), "Research, Innovation, and Productivity: an Econometric Analysis at Firm Level", *Economics of Innovation and New Technology*, **7**, 115-158.

Danzon, P.M. (1997), "Price Discrimination for Pharmaceuticals; Welfare Effects in the US and the EU", *International Journal of the Economics of Business*, **3(3)**, 301-321.

Duflos, G (2006), "Persistence of innovation, technological change and quality-adjusted patents in the US pharmaceutical industry", DRUID summer conference.

Dujowich, J.M.P. (2013), "Product Innovations by Young and Small Firms", *Small Business Administration*, Office of Advocacy, Fountainville, PA, 1-28.

DiMasi, J.A., R.W Hansen and H.G. Grabowski (2003), "The Price of Innovation: New Estimates of Drug Development Costs", *Journal of Health Economics*, **22(2)**, 151-85.

Enzing, C. and S. Kern (2006), "Structure, Dynamics and performance in National Biopharmaceutical Innovation Systems", in: Innovation in Pharmaceutical Biotechnology Comparing National Innovation Systems at the Sectoral Level, OECD Publication, 1-187.

Gilbert, R. (2006), "Looking for Mr. Schumpeter: Where are we in the competition-innovation debate? in: Jaffe AB, Lerner J., Stern S (eds), Innovation policy and the economy, **6**, 159-215.

Glover, G.J. (2002), "Competition in the Pharmaceutical Marketplace", PhRMA Statement, Washington DC.

Heckman, J.J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, **47(1)**, 53-161.

Henderson, R.M. and I. Cockburn (1996), "Scale, Scope and Spillovers: The Determinants of Research Productivity in Drug Discovery", *Rand Journal of Economics*, **27(1)**, 32-59.

Hennessy, C., and T.M. Whited (2007), "How Costly is External Financing? Evidence from a Structural Estimation", *Journal of Finance*, **62**, 1705-1745.

Hesketh, S.R., A. Skrondal (2013), "Avoiding biased versions of Wooldridge's simple solution to the initial condition problem", *Economic Letters*, **120**, 346-349.

Hottenrott, H., and B. Peters (2011), "Innovative Capability and Financing Constraints for Innovation: More Money, More Innovation?", Center for European Economic Research, Discussion Paper No. 09-081.

Huergo, E., and J. Jaumandreu (2004), "How does Probability of Innovation change with Firm Age?", *Small Business Economics*, **22**, 193-207.

Jensen, E.J. (1987), "Research Expenditures and the Discovery of new Drugs", *Journal of Industrial Economics*, **36(1)**, 83-95.

Klepper, S. and P. Thompson (2010), "Disagreements and Intra-industry Spinoffs", *International Journal of Industrial Organisation*, **28(5)**, 526-538.

Klomp, L. and G.V. Leeuwen (2006), "On the Contribution of Innovation to Multifactor Productivity Growth", *Economics of Innovation and New Technology*, **15(4)**, 367-390.

Leeuwen, V. G. (2002), "Linking Innovation to Productivity Growth using two waves of the Community Innovation Survey", OECD Science, Technology and Industry Working Papers, 2002/8, OECD Publishing.

Levin, R.C., Cohen W.M., and D.C. Mowery (1985), " R&D appropriability, opportunity and market structure: new evidence of some Schumpeterian hypothesis, *American Economic Review*, **75**, 20-24.

Levin, R.C., Klevorick, A.K., Nelson, R.R., and S.G. Winter (1987), "Appropriating the returns from industrial research and development", *Brookings Papers on Economic Activity*, **3**, 783-831.

Mairesse, J. and P. Mohnen (2002), "Accounting for Innovation and Measuring Innovativeness: An Illustrative Framework and an Application, *American Economic Review*, **92(2)**, 226-230.

Malerba, F and L. Orsenigo (1996), "Schumpeterian Patterns of Innovation Are Technology-Specific", *Research Policy*, **25**, 451-478.

Mansfield, E. (1968), "Industrial research and technological innovation: An Econometric Analysis", New York.

Martin, P., T. Mayer and F. Mayneris (2010), "Spatial concentration and plant-level productivity in France", *Journal of Urban Economics*, **69**, 182-195.

Nilsen, O.A., and F. Schiantarelli (2003), "Zeros and lumps in investment: Empirical evidence on irreversibilities and nonconvexities", *the Review of Economics and Statistics*, **85**, 1021-1037.

Nelson, R., and S. Winter (1982), "An Evolutionary Theory of Economic change", Cambridge.

Nickell, S. (1996), "Competition and Corporate Performance", *Journal of Political Economy*, **104(41)**, 724-746.

OECD (2003). Science, Technology and Industry Scoreboard 2001, OECD: Paris.

Peters, B. (2007), "Nothing's Gonna Stop us Now? AN Empirical Investigation on the Innovation Success Breeds Success' Hypothesis", mimeo.

Peters, B. (2009), "Persistence of Innovation: Stylized Facts and Panel data Evidence", *Journal of Technology Transfer*, **34**, 226-243.

Phillips, A.. (1971), "Technology and Market Structure: A Study of the aircraft industry", Lexington.

Prusa, T.J. and J.A. Schmitz (1994), "Can Companies maintain their Initial Innovative Thrust? A Study of the PC Software Industry", *Review of Economics and Statistics*, **76(3)**, 523-540.

Raymond, W., P. Mohnen, P. Mohnen, P. Franz and S. V. Sybrand (2009), "Innovative Sales, R&D and Total Innovation Expenditures:Panel Evidence on their Dynamics", Cirano Scientific Series.

Raymond, W., P. Mohnen, F. Palm and S. Schim der Loeff, "Persistence of Innovation in Dutch Manufacturing: Is it Spurious?", *Review of Economics and Statistics* (forthcoming).

Rojas, C. (2011), "Market power and the Lerner Index: A Classroom Experiment", *Journal of Industrial Organization Education*, **5(1)**, 1-19.

Romer, P. (1990), "Endogenous technological change", *Journal of Political Economy*, **98(5)**, 71-102.

Schumpeter, J. (1943), "Capitalism, Socialism and Democracy", London: Allen Unwin.

Scherer, F. M. (1967), "Market Structure and the employment of scientists and engineers", *American Economic Review*, **57(3)**, 524-531.

Scherer, F.M. (1965), "Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions", *American Economic Review*, **55**, 1097-1123.

Symeonidis, G. (1996), "Innovation, Firm Size and Market Structure", OECD Economics Department Working Papers, No. 161, OECD Publishing.

Wooldridge, J. (2010), "Econometric Analysis of Cross Sections and Panel Data", 2nd Edition, Cambridge: MIT Press.

Wooldridge, J. (2005), "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity", *Journal of Applied Econometrics*, **20**, 39-54.