universiteit hasselt
KNOWLEDGE IN ACTION

2014 | Faculty of Business Economics

DOCTORAL DISSERTATION

Innovation, competition and firm performance: An econometric analysis on the Dutch pharmaceutical sector

Doctoral dissertation submitted to obtain the degree of
Doctor of Applied Economic Science, to be defended by

Shreosi Sanyal

Promoter: Prof. Dr Mark Vancauteren | UHasselt
Co-promoter: Prof. Dr Ludo Peeters | UHasselt

universiteit hasselt
KNOWLEDGE IN ACTION

2014 | Faculty of Business Economics

DOCTORAL DISSERTATION

Innovation, competition and firm performance: An econometric analysis on the Dutch pharmaceutical sector

Doctoral dissertation submitted to obtain the degree of
Doctor of Applied Economic Science, to be defended by

Shreosi Sanyal

Promoter: Prof. Dr Mark Vancauteren | UHasselt
Co-promoter: Prof. Dr Ludo Peeters | UHasselt

D/2014/2451/65
ISBN number: 9789089130327

Doctoral committee:

Prof. Dr. Piet Pauwels (Hasselt University, Dean)

Prof. Dr. Mark Vancauteren (Hasselt University, Promoter)

Prof. Dr. Ludo Peeters (Hasselt University, Co-promoter)

Prof. Dr. Wim Vanhaverbeke (Hasselt University)

Prof. Dr. Roel Braekers (Hasselt University)

Prof. Dr. Glenn Rayp (Ghent University)

Prof. Dr. Boris Lokshin (Maastricht University)

Dedicated to my parents

Acknowledgements

"There is not a more pleasing exercise of the mind than gratitude. It is accompanied with such an inward satisfaction that the duty is sufficiently rewarded by the performance" ~ Joseph Addison

I would like to take this opportunity to express my heartfelt gratitude to the people who have played an integral role in my journey of realizing this PhD. First of all, I am genuinely grateful to my promoter Prof. Dr. Mark Vancauteren for providing me with the opportunity to pursue my doctoral research. Also his guidance throughout my research work has played an indispensable role in formulating my PhD. I would also like to express my sincere thanks and regards to my co-promoter Prof. Dr. Ludo Peeters for his constant co-operation, support and optimism at every stage of my PhD research.

I am indebted to Prof. Dr. Glenn Rayp, who has not only played a formidable role in accomplishing my PhD, but has also been instrumental in shaping my academic achievements in the course of time. I hold deep gratitude for his insightful suggestions and comments in my research studies during the entire phase of my PhD. I am also extremely thankful to Prof. Dr. Wim Vanhaverbeke for his support and encouragement. I truly appreciate his helpful guidance for my research and academic endeavors. Further, I owe my gratitude to the other members of the jury, which includes Prof. Dr. Boris Lokshin and Prof. Dr. Roel Braekers, for their invaluable inputs in my thesis manuscript.

I also extend my thanks to the members of the secretary office and the computer helpdesk, for their friendly co-operations throughout my PhD time. I admire their friendliness and good vibes. Also, I am thankful to have friendly colleagues from the KIZOK institute and the ECO group.

My stay in Hasselt would not have been a happy one without our international researchers' community, which was like a home away from home. Besides my Belgian friends, my earnest thanks goes out to all my friends from India, Italy, Iran, Columbia, Belarus, Greece, Germany, Czech-republic, Serbia, Portugal, Brazil, Philippines, China and the South African countries. Getting the opportunity to know so many different cultures was indeed a valuable experience. I learned that, although there exist so many different languages in the world, but the voice of love, friendship and brotherhood remains the same. A special mention is well-deserved for Filippo who has been genuinely supportive and encouraging during the hard times of my life. I am thankful to him from the bottom of my heart for being a true friend.

My deepest regard and heartfelt thanks to my parents and my brother for being with me all through this journey and experiencing my failure and success as much as I did, and even sometimes greater. Without their constant and daily moral support and encouragement, it would not have been possible for me to complete my PhD. I cannot thank them enough for loving me so dearly and being there for me every single moment. Also, a warm thank you to my sister-in-law for being a whole-hearted support and a genuine well-wisher. I also want my nephew (who is just a few months old) to know that his cherubic smile and innocence have sparkled happiness through me and have given me the strength to go through the difficult moments.

Most importantly, I would like to thank the almighty who has guided me through the correct path of life, and has given me the strength to overcome the obstacles that came my way.

This period of my PhD has not only been a huge learning experience in terms of my research, but has also been a lesson for the journey called life. Therefore, I would like to thank my life-my inner soul for traversing through all the hurdles with grit, courage and self-respect.

The journey in realizing my PhD has provided me with the broader horizon and outlook towards the world at large. It has instilled in me the quest to excel and the determination to face any challenge that crosses my way. And as I contemplate on the path I travelled so far, with all the happiness and heartaches, I owe to each and every single soul who have been in this journey at some point and have put their own portion of responsibility in shaping my today.

Shreosi Sanyal
Diepenbeek, December 2014

Abstract

This study mainly deals with the contemporaneous relation between innovation, competition and firm performance, conditioned by various firm-level determinants, for the Dutch pharmaceutical industry. Our analysis encompasses R&D investment and patents as the essential indicators of innovation; and investigates on their plausible linkage with competition, firm size, firm age, capital intensity and other variables, that eventually affect the productivity and growth in this knowledge-intensive industry. In addition to a comprehensive theoretical exploration, a number of empirical investigations are performed on a firm-level panel database for the Netherlands' pharmaceuticals.

In this PhD dissertation, chapter 1 deals with the theoretical overview and motivation of this study, with special emphasis on the review of the Dutch pharmaceutical industry. Chapter 2 involves a comparative theoretical study with the pharmaceutical industries of other countries of Europe and U.S. Additionally, the divergence in policy issues, regulations and financing with the developing countries, especially of India is analyzed succinctly. Chapter 3 comprises of a description on the data sources that we have used to construct our panel dataset, along with the corresponding variable constructions.

From chapter 4 to 7, the main topics of the thesis are discussed. Chapter 4 empirically investigates on how various firm characteristics affect a firm's decision to invest in R&D, using a generalized sample selection Tobit II estimation technique. Subsequently, chapter 5 analyses the effect of R&D intensity (fitted values obtained from Tobit II estimation) and other determinants on the innovation output in this particular sector in the Netherlands. The innovation output of the firms is indicated by the patents and citation-weighted patents, which are used alternatively as the dependent variables in a Zero-inflated negative binomial estimation framework. Chapter 6 emerges as an extension to Chapter 5, with a

deeper focus on the competitive framework in this industry. It provides with an in-depth analytical approach on various competition indicators, which includes the Herfindahl-Hirschman concentration index, the Lerner index, the mark-up with adjustment for factor elasticities and the profit elasticity as proposed by Boone (2000). Plus, the effect of various competition indicators on the innovation output of the pharmaceutical firms is analyzed using a non-linear specification. We have considered three empirical models, namely the Zero-inflated negative binomial, Hurdle negative binomial logit and the Poisson-pseudo maximum likelihood estimators, with quadratic specifications for the competition measures. In our estimation strategies applied for all the three mentioned chapters, we have incorporated the maximum likelihood approach following Wooldridge (2005), in order to handle the individual effects. Chapter 7 finally trajects the effective channels through which innovation, captured through R&D measure and citation-weighted patent counts, affect the overall productivity in the Dutch pharmaceutical industry, using the growth accounting approach. This analysis is supplemented with a production function approach, in which R&D investment enters the production function as a factor input. Random effect GLS and system GMM are the two estimation techniques used to serve this purpose. However, the latter approach also includes the semi-parametric Levinsohn-Petrin technique as an additional econometric tool.

As the last chapter of the thesis, chapter 8 puts forth the fundamental conclusions as inferred from the previous chapters, with relevant policy implications. From the results obtained in chapter 4, it was empirically established that young and small entrepreneurs, with adequate capital reserve and enjoying higher degree of monopoly are likely to invest in R&D. This entails for the encouragement of venture capital markets and size dependent R&D taxation or regulation. However chapter 5 suggests that bigger sized firms have a greater propensity to persistently patent their innovations, under prominent barriers to entry. Although R&D intensity is found to act as a major determinant for generating new patents, our

findings suggest that, not all innovating firms are deemed to patent their innovations. Persistence in patenting is a pivotal strategic measure that big firms undertake to form barriers in the market and evade competition. This can result in prominent market inefficiencies that lead us to question whether patenting is really desirable in the social context. Further, Chapter 6 establishes the sensitivity of the different competition measures in order to analyze its effect on innovation output. From the in-depth empirical analysis undertaken in this chapter, it is asserted that competition exhibits either a negative or a U-shaped relation in the Dutch pharmaceutical industry. Finally, chapter 7 provides evidence of both R&D investment and patenting performance having a positive and significant effect on productivity, while the former has a greater influence than the latter. This reaffirms the fact that not all innovations undergo the patenting process. Therefore, encouraging higher R&D investment is likely to propagate productivity and growth in this sector.

Samenvatting

Dit doctoraat bestudeert de simultane relatie tussen innovatie, competitie en bedrijfsprestaties voor de Nederlandse farmaceutische sector, waarbij gecontroleerd wordt voor determinanten op bedrijfsniveau. Meer bepaald worden R&D investeringen en patenten als belangrijkste indicatoren voor innovatie beschouwd, en wordt hun mogelijke verband met competitie, bedrijfsgrootte, bedrijfsleeftijd, kapitaal intensiteit en andere variabelen onderzocht. Uiteindelijk bestuderen we hoe dit verband productiviteit en groei in deze kennisintensieve sector beïnvloedt. Hiertoe wordt een uitgebreide theoretische studie uitgevoerd, alsook meerdere empirische analyses op een panel data set van Nederlandse farmaceutische bedrijven.

In het eerste hoofdstuk van dit doctoraat ligt de focus op het theoretisch overzicht en de onderzoeksmotivatie, met aandacht voor de context binnen de Nederlandse farmaceutische sector. Hoofdstuk twee bevat een vergelijkende theoretische studie met de farmaceutische sectoren uit andere Europese landen en de Verenigde Staten. Bijkomend wordt een beschrijving gegeven van het verschil met ontwikkelingslanden, zoals India, wat betreft beleid, regelgeving en financiering. De databronnen en de constructie van de panel data set worden besproken in hoofstuk drie.

De volgende hoofdstukken omvatten de kern van dit onderzoek. In hoofdstuk vier wordt onderzocht hoe bedrijfskenmerken van de farmaceutische sector een invloed hebben op de R&D-investeringsbeslissing aan de hand van de *generalized sample selection Tobit II* schattingstechniek. Vervolgens analyseert hoofdstuk 5 het effect van R&D intensiteit – aldus geschat via de *Tobit II* schattingstechniek – en andere determinanten op innovatie output via een *zero-inflated negative binomial estimation framework.* Patenten en citatie-gewogen patenten worden hierbij beschouwd als indicatoren voor innovatie output. Hoofdstuk zes vormt een

uitbreiding op hoofdstuk vijf, waarbij meer nadruk wordt gelegd op het aspect competitie in deze sector. Er wordt een diepgaande analytische studie uitgevoerd van verschillende competitie indicatoren, zoals de *Herfindahl-Hirschman concentration index*, de *Lerner index*, de *mark-up* gecorrigeerd voor factor elasticiteiten en winst elasticiteit zoals voorgesteld door Boone (2000). Vervolgens wordt het effect van deze competitie indicatoren op innovatie output van de farmaceutische bedrijven geanalyseerd. De volgende empirische modellen, met kwadratische specificaties voor de competitie maatstaven, worden hierbij gehanteerd: *Zero-inflated negative binomial, Hurdle negative binomial logit* en *Poisson-pseudo maximum likelihood*. De schattingstechnieken in hoofdstukken vier, vijf en zes gaan allen uit van de *maximum likelihood approach* van Wooldridge (2005) om gepast rekening te houden met bedrijfs-individuele effecten. Hoofdstuk zeven vormt de laatste schakel in het onderzoek en geeft een beeld van hoe innovatie, gemeten door R&D en citatie-gewogen patenten, de productiviteit in de Nederlandse farmaceutische sector beïnvloedt aan de hand van een *growth accounting approach*. Deze analyse wordt aangevuld met een productiefunctie methode, met R&D investeringen als factor input van de productiefunctie. De schatting gebeurt hier op basis van *random effect GLS* en *system Generalized Method of Moments*. Deze laatstgenoemde techniek bevat ook de semi-parametrische *Levinsohn-Petrin* techniek als een bijkomende econometrische methode.

Het laatste hoofdstuk bespreekt tenslotte de fundamentele conclusies samen met relevante beleidsimplicaties. De empirische resultaten uit hoofdstuk vier tonen aan dat jonge en kleine ondernemingen meer geneigd zijn om te investeren in R&D als ze beschikken over voldoende kapitaalreserves en een hogere mate van monopolie hebben. De overheid kan hierop inspelen door de markt voor *venture capital* aan te moedigen en grootte-afhankelijke belastingen en regulatie voor onderzoek en ontwikkeling door te voeren. Hoofdstuk vijf concludeert daarnaast dat grotere bedrijven de neiging hebben om voortdurend hun innovaties te patenteren wanneer

belangrijke toetredingsdrempels bestaan. Hoewel gevonden wordt dat R&D intensiteit een belangrijke determinant is voor het genereren van nieuwe patenten, tonen onze resultaten aan dat niet alle innoverende bedrijven hun innovaties patenteren. Het is vooral een strategische zet die grote bedrijven ondernemen om toetredingsdrempels te creëren en competitie te omzeilen. Dit kan leiden tot belangrijke marktinefficiënties, waarbij de vraag gesteld moet worden of patenteren wel wenselijk is in een sociale context. De conclusies inzake competitie worden verder uitgediept in hoofdstuk zes, waar blijkt dat competitie een negatieve dan wel U-vormige relatie vertoont in de Nederlandse farmaceutische sector. Tot slot levert hoofdstuk zeven het bewijs dat zowel R&D investeringen als patent prestaties een significant positief effect hebben op productiviteit, waarbij het eerstgenoemde een grotere invloed heeft dan het laatstgenoemde. Dit bevestigt nogmaals dat niet alle innovaties het patenteringsproces ondergaan. Het is voornamelijk het stimuleren van hogere R&D investeringen dat dus gunstige gevolgen heeft voor de productiviteit en groei in deze sector.

Table of Contents

List of Tables:

List of Figures

Chapter 1

INTRODUCTION

1.1 Motivation and overview

The pharmaceutical industry is characterized by a complex creative process with a strong research base as well as an intricate system of managerial expertise. This sector is of utmost importance and crucial not only from the perspective of economic growth, but also for playing the quintessential role of extending human lives and providing better health conditions. Scherer (2007) delineated the history of drug discovery, which commenced with the early discovery of Quinine to combat malaria, or Edward Jenner's invention of small-pox vaccine. Several natural ingredients were tried and tested from ancient times, for eradicating human diseases and increasing the longevity of life. The trajectory of drug discovery has witnessed massive successes over time, with newer drugs and treatments for the improvement of life. In the new genera of medicinal expertise, a broad spectrum of medications are invented based on genetic recombinant methods, modern molecule screening and DNA-sequence identification. The modern scientific age witnesses a substantial flurry of new medicinal treatments for alleviating various diseases and health conditions.

However, the pace of invention of new drugs seems to have stagnated relative to the colossal improvement in scientific equipment and computer-aided designing of new molecular entities. This stems from the fact that, the discovery and perpetual development of new pharmaceutical entities is undeniably a complex and risky affair, which involves high levels of uncertainty and huge monetary investments. The introduction of a novel drug eventuates through a tedious and lengthy process,

ranging from the invention of a new therapeutic product to its certification of efficaciousness, and finally its commercialization and introduction in the market. Thus, the rubric of effective drug design and discovery is also implemented and accomplished in academic institutions and public research enterprises, in collaboration with the pharmaceutical companies. Reichert and Milne (2002) emphasizes that the drug discoveries within academic communities have a leverage on commercial enterprises, as they possess the scientific expertise to manipulate molecular entities for effective utilisation in therapeutic processes. Along with the linkage within the academic and industrial enterprises, numerous small biotech firms are found to be emerging in the vicinity of the academic research centres. The innovative pharmaceutical companies are increasingly channelizing their innovation prerogatives to these research units and licensing their discovered molecular entities for later stage commercial purposes. Additionally, the pharmaceutical firms often possess the full ownership right of the inventions of these academic units and research centres, or employ their own researcher in their invention process. This enables them to retain an adequate scientific inventory of new molecular entities at their disposal for their new drug development pipeline.

However, despite the strategic dissemination of inventive procedures to other establishments, the final cost involved is excessively high. This entails the innovative firms to resort to intellectual property rights, in order to safeguard their innovation and combat generic intervention. However the generic firms create a niche for themselves once the patent expires. Albeit the complicated nature of drug development and the strategic mechanisms operating in this industry, the pharmaceutical sector will always have a profound importance for mankind, as it continues with the daunting task of abating the scourge of various diseases. Therefore, it was rightly posited by Levy and Wickelgren (2001), “It is hard to think of many industries that have contributed as much to human welfare as the pharmaceutical industry. The importance of the industry make the job of competition authorities that much more difficult and important”.

In the light of the above discussion, this PhD thesis can be briefly outlined in the following manner. Chapter 2 involves a comparative theoretical underpinning with the pharmaceutical industries of other countries of Europe, US and Japan. In addition, the divergence in policy and innovation criterion with the developing countries, especially of India has been theoretically analyzed in this chapter. Chapter 3 details on the descriptive statistics of the data sources that have been used for the construction of our panel dataset. Chapter 4 till 7 encompasses the main topics of our thesis, which essentially deals with the empirical investigations, based on a structured analytical framework. Chapter 4 deals with an econometric application on how various firm characteristics affect its decision to invest in R&D, using a generalized sample selection Tobit II estimation technique. Subsequently, chapter 5 analyses the effect of R&D intensity (fitted values obtained from Tobit II estimation) and other determinants on the innovation output in this particular sector in Netherlands, with the application of a zero-inflated negative binomial estimation strategy. Chapter 6 is an extension to Chapter 5, with a deeper focus on the competitive framework in this industry. Furthermore, the effect of various competition indicators on the innovation output of the pharmaceutical firms is analyzed using three empirical models, which includes, the Zero-inflated negative binomial, Hurdle negative binomial logit and the Poisson-pseudo maximum likelihood estimators, with quadratic specifications for the competition measures. Chapter 7 finally tracks the effective pathway through which R&D investment and patenting performance affect the overall productivity in the Dutch pharmaceutical industry, using the growth accounting approach as well as the production function approach, in a random effect GLS and system GMM estimation framework. As the final chapter of the thesis, chapter 8 provides a pedagogic conclusion for each chapters, with the critical ramification of the inferences obtained. It also succinctly elaborates on the relevant policy implications, as well as on the efficacy and the lack thereof on public policy in regulation and finance.

Taking into account the modalities and attributes of the corresponding chapters of this dissertation, it is seen that our research work is predominantly based on a micro-level investigation of the Dutch pharmaceutical industry. Therefore, in line with our primary research exploration, this chapter further elucidates on the innovation, competition and productivity criteria, and their subsequent policy issues that exemplifies and introduces this particular sector in the Netherlands.

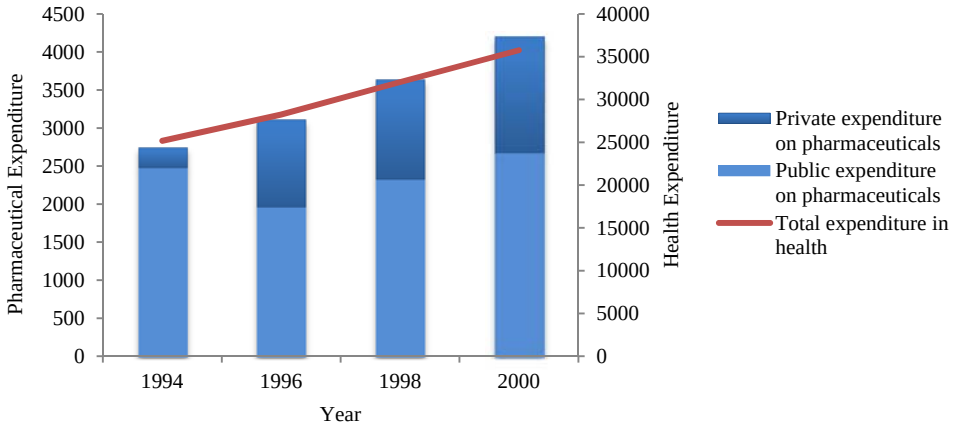
1.2 The pharmaceutical industry in the Netherlands

In spite of having a small geographic size with only over 16.5 million inhabitants, the Netherlands holds a strategic location in Europe, that allows it to expand its locus of trade interaction and possible collaborative ventures with the neighboring countries, as well as countries outside Europe. Pharmaceutical statistics by Nefarma (2011) reveals that the Dutch pharmaceuticals generated $4.88 billion in total revenues in 2009, employing 50,000 people of which 27% were in the R&D, and accounting for 9% of private Dutch R&D investment. The ingenious and collaborative characteristics of Dutch pharmaceutical firms and the health sector at large, combine the country's strong research potentialities with commercial opportunities through public-private partnerships and numerous other investment opportunities. The majority of the pharmaceutical firms in the Netherlands are subsidiaries of major foreign pharmaceutical companies that have their production, logistics and/or research facilities in the Netherlands.

By possessing a high quality infrastructure and adequate capital reserves, the Dutch pharmaceutical industry has the potential to transpire into the key industry for providing in-house employment base, upgrading therapeutic research and bolstering growth and development in the economy. According to Enzing *et al.* (2004), the expenditures on pharmaceutical products in the Netherlands have increased steadily over the last decade; with the rate of growth of the Dutch pharmaceutical spending transcending the rate of growth of total Dutch expenditures in health. The figure below confirms that the rate of growth of total

pharmaceutical expenditure have escalated during 1994-2000, being stronger than the expenditure in total health. Also, the diagram provides compelling evidence of a noticeable increase in private expenditure from 1996 onwards, with an estimated hike of $879 million dollars from 1994 to 1996. Enzing *et al.* (2004) opined that the changes in the pharmaceutical pricing policies and the introduction of expensive medicines, along with the increase in demand for pharmaceutical products have played a key role for its steady growth.

Fig. 1.1: Expenditure in health and pharmaceuticals in the Netherlands, million US $ PPP

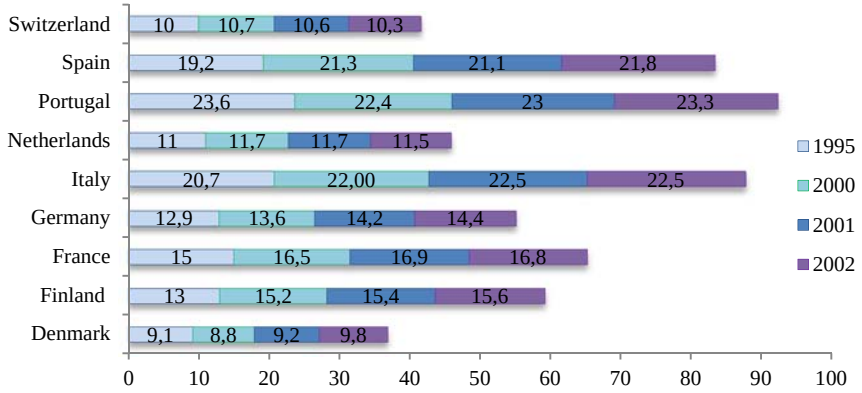


Source: OECD Health Data, 2002b

However, the total pharmaceutical expenditure in the Netherlands has been considerably lower relative to the other major EU countries that invest heavily on pharmaceutical and related products. The chart below shows that the percentage share of pharmaceutical spending on the total health expenditure for some EU countries for the periods, 1995, 2000, 2001 and 2002. It can be discerned from fig.1.2 that the Netherlands has a subservient position in terms of total expenditure, relative to most other prominent countries in this category that includes, Germany, France, Italy, Portugal and Spain, among others. The divergence in pharmaceutical expenditure amongst the different countries can emerge due to several factors, such as, the difference in demographic conditions, pricing policies, health insurance

facilities, influx of generics, the share of demand for prescription or over-the-counter drugs (Mossialos and Oliver, 2005).

Fig. 1.2: Expenditure on pharmaceuticals and other medical non-durables as a percentage of total expenditure in health



Source: OECD Health Data 2008, December 2008 update

Competitiveness of the EU market and Industry for Pharmaceuticals-Vol.1: Welfare implications of regulation

More recent data reveals that the Dutch pharmaceutical expenditure has contributed to just 9.6% of total healthcare expenditure in 2009, which clearly indicates that the proportion of pharmaceutical spending to total healthcare outlay is substantially less. However it accounts for a net amount of US$ 7,377 expenditure in the pharmaceutical sector in 2009, which is markedly higher than the previous years (Netherlands Pharmaceutical Profile Report, 2011). On reverting back to fig. 1.1, it is seen that the total pharmaceutical spending in 2000 amounted to US$ 4,205 only, suggesting a considerable augmentation in the Dutch pharmaceutical demand in the later years.

Based on the statistical data provided by Nefarma (2003), the Dutch pharmaceutical industry in 2002 comprised of 72.2% of branded pharmaceuticals, while only 18.5% engaged in generic production and 9.4% were obtained through parallel imports. However the generic firms are increasingly gaining prominence

over time, following the expiration of patents for numerous branded drugs and relatively lesser proportion of new breakthrough drugs in the market. Thus, the pharmaceutical sector witnesses several metamorphic transitions in terms of policy changes as well as competitive and innovative attributes, consequently affecting the overall firm performance. Therefore, the following subsections provide a concise introspection of the policy changes, and the concurrent characteristics of innovation and competition practices by the Dutch pharmaceutical industry.

1.2.1 Policy issues: Regulation and financing

The pharmaceutical industry in the Netherlands has undergone a panoply of policy changes over the years that transcends from a regulatory regime in the early 1990s to managed competition in 2006. A study by Boonen *et al.* (2010) puts forth the sequential policy transformations in this Dutch sector, where it is seen that the health insurance reform of 2006 has been an effective step to control the generic drugs prices, as compared to the former regulatory steps that were undertaken by the government. A synopsis of the most crucial policy changes for the pharmaceutical industry that was enforced by the Dutch government from 1991-2009 is provided below.

- 1991: Drug reimbursement system (GVS)
- 1996: Regulation of maximum price (WGP)
- 1998: Uniform retraction in the reduction of bonuses and discounts by the pharmacies
- 1999: GVS reimbursement limits recalculation
- 2004-2008: Nationwide voluntary covenants between the pharmaceutical industries, pharmacies, insurers and the government on generic drug price reduction.

Addition to the above governmental regulations, several reforms based on the market orientation also came into force, which is enumerated hereunder.

- 1992: Allowing the social health insurers the possibility of selective contracting in the retail market of pharmaceuticals.
- 1995-2003: The financial risk of the insurer's increased from 3% to more than 90% in case of outpatient medicinal expenditure.
- 1999: Pharmacies and insurance agencies were allowed to collaborate and integrate together.
- 2003: The insurance providers were permitted to select their drug coverage, where therapeutic equivalent drugs could be excluded.
- 2006: Introduction of the Health Insurance Act.
- 2008: Pharmacy fees deregulation in case of outpatient prescription drug delivery.

The Netherlands' healthcare system has aimed towards the amalgamation of the old sickness funds and private health insurances from as early as 1974. In the course of time, several endeavors for merging all healthcare schemes and other policy formulations were initiated, which acted as the building blocks for the final enactment of the Health Insurance Act of 2006. According to this reform policy, a single health insurance scheme came into effect which led to amelioration of the dual private and public health care system. Through this enforcement, the citizens of the country are obliged to purchase standard medical insurance schemes from private health insurance companies (Van de Ven and Shut, 2008). This led to the emergence of managed competition between the pharmaceutical providers and the insurers, whereby the patients have the liberty to choose from an array of healthcare provisions. This unique Dutch healthcare system has brought forth radical milestone for cost-containment and more efficient performance by the insurers and medical providers.

1.2.2 Innovation

Our research analysis is confined to the pharmaceutical sector alone, as it emphatically reflects on the essential concoction between innovation input and

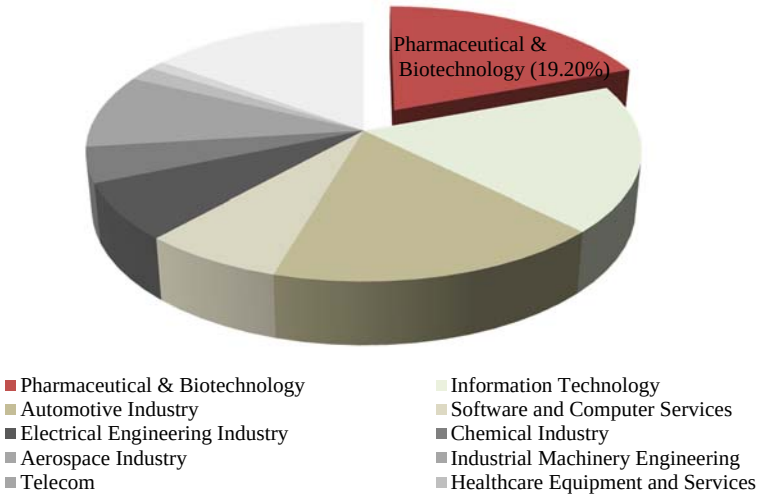
innovation output as an integral part of its performance. Despite being marked as one of the most knowledge intensive sectors, the innovation in the pharmaceutical industry is a complex process of building a nexus between knowledge application and tackling the creativity of scientific clinical concepts through management expertise. The continual increase in the complexity and scale of pharmaceutical innovation 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 innovation activity is imperative for this sector.

R&D expenditure as the innovation indicator

The latest upsurge of new research tools like pharmacogenomics and high throughput screening has led to promising advancements in clinical science research and technology. Statistical evidences suggest that the discovery and development of new drugs has led to an increase in life expectancy and has successfully improved the quality of life by eradicating the causal effects of various diseases. Studies pertaining to the cost-benefit analysis of pharmaceutical R&D (e.g. Lichtenberg, 2001) hints at significant positive externalities and social welfare from the innovation of new drugs. Cutler and MacClellan (2001) have concluded in their survey that, “in most of the cases we analyzed, technological innovations in medicines are on net positive. Technology often leads to more spending, but outcomes improve by even more”.

According to the European commission report by Guevara *et al.* (2008), the three main sectors that contribute for more than half of the total expenditure in R&D (54.5%) are information technology, pharmaceuticals and biotech companies, and the automotive. Out of these three sectors, the pharmaceutical and biotech industry account for $1/5^{th}$ of the total world-wide investment in R&D. The figure below illustrates the share of R&D investment in pharmaceutical & biotech sector, in comparison to the other sectors.

Fig. 1.3: Share of worldwide R&D investment in different sectors



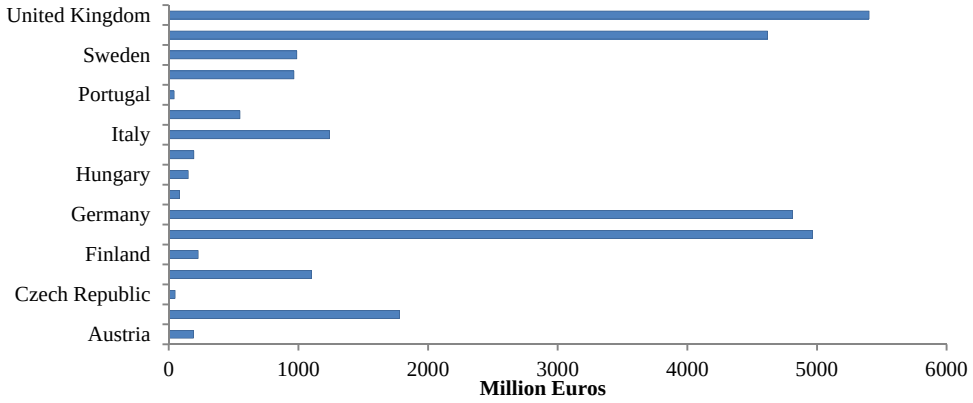
Source: European Commission (2008), EU Industrial R&D Investment Scoreboard

Regarding the Dutch pharmaceutical sector, its regulatory framework during the early 1990s was envisaged with several hurdles for innovation propensity, in comparison with the other prominent EU countries. Most of these disadvantages comprised of the prolonged time taken for application and decision making processes, the lack of transparency and predictability of the procedures undertaken, and the overlapping of tasks and evaluation frameworks by the official authorities. However, R&D investments in this Dutch sector gained momentum from the latter half of 1990s. Based on the study by Enzing *et al.* (2004), the gross domestic R&D expenditure in the Netherlands' pharmaceuticals increased prominently, with a 34% growth rate from 1994 till 2000. In spite of having an appreciable improvement in the R&D investment level, a perpetual slack was noticed over time, with only a 3% growth rate from 1999 to 2000. By and large, the main reason behind the spur in pharmaceutical R&D expenditure was the initiation of various programs from the late 90's onwards, which were oriented towards fundamental research and mostly funded by the Netherlands Organization for Scientific

Research (NWO). Additionally, the Netherlands Federation for Innovative Pharmaceutical Research (FIGON) and the Steering Group Orphan Drugs were the two parallel schemes that were started to stimulate research in the Dutch pharmaceutical sector.

But latest statistical evidence from EFPIA report (2012) suggests that, although the pharmaceutical sector in Europe remains as one of the top performing high-technology sectors, the R&D expenditure in Netherlands pharmaceutical industry still lags far behind compared to the other important European countries. According to the 2009 EFPIA data, it can be observed from figure 1.4 that, countries like U.K, Switzerland, Germany, France, Italy and even Belgium is far ahead of Netherlands in terms of pharmaceutical R&D expenditure.

Fig 1.4: Pharmaceutical R&D expenditure in some EU countries



Source: EFPIA member associations official data

Note: Figures relate to R&D expenditure in the pharmaceutical sector for each country extracted from 2009 data

However various promising innovation projects in the Netherlands are ongoing, with potential synergies between intramural and extramural associations. Additionally, the penchant for therapeutic innovation is progressively harnessed upon with the help of academic and research institutes. Hence, a more robust

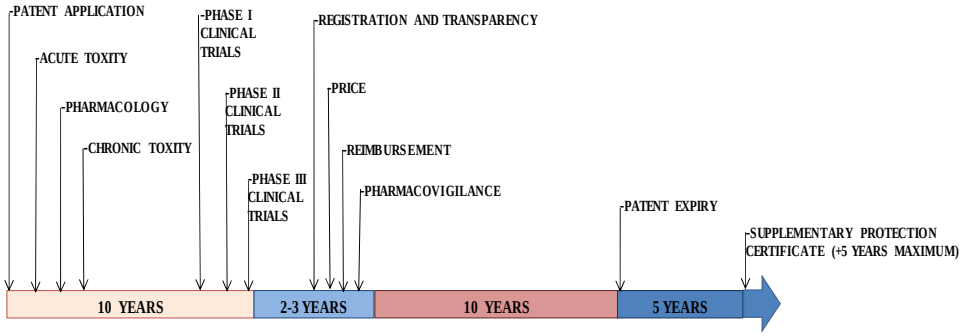
innovation portfolio for the Dutch pharmaceutical firms can be conjectured in the near future.

Nevertheless, the R&D investment in this sector is undoubtedly a complicated and long-standing process that calls for immense scientific and economic potentials by the pharmaceuticals. Optimization of this procedure involves a systematic development and discovery of new clinical compounds from a scientific concept to the final ratification by the regulatory authorities and the delivery of the new medicinal products for consumption. In addition, patenting of pharmaceutical innovation plays a pivotal role to accrue the costs of R&D and is of paramount interest for safeguarding their innovation from the generic drug producers.

Patents as the innovation indicator

Based on the preceding discussion, the following diagram represents the stylized phases encompassing the route from patent application to its expiry. As can be seen, the duration spanning the time of patent application to its final authentication is around 12-13 years, comprising of various stages of clinical trials, testing, pharmacovigilance and other administrative obligations. In the course of this intricate procedure, only one or two new chemical or biological entities, from a wide range of 10,000 synthesized substances, are successful in passing through all the development stages, based on their efficacy and safety. Furthermore, the cost incurred to develop a single therapeutic substance was estimated to be a lofty 1,172 euro in 2012 (Mestre-Ferrandiz *et al.*, 2012).

Fig. 1.5: Pathway from patent application to its expiration

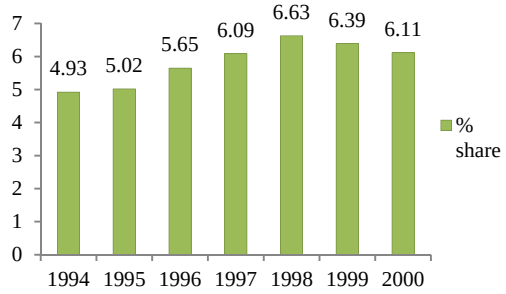


Source: EFPIA: Medicines for Mankind

Although the application for patent protection involves a lengthy period, it is an indispensable strategic step to recoup the enormous R&D investments undertaken. A number of studies have found that patents are more important to pharmaceutical firms in appropriating the benefits of innovation, compared to other high-tech industries. Both Levin *et al.* (1987) and Cohen *et al.* (2000) found that, innovation plays a crucial role in the pharmaceutical industry. In contradiction, other research-intensive industries, like the computer and semi-conductor industry, place more importance to lead-time and improving the production efficiencies through improved technological know-how. According to Arundel *et al.* (1998), about 80% of all pharmaceutical products and about 45% of all processes are patented.

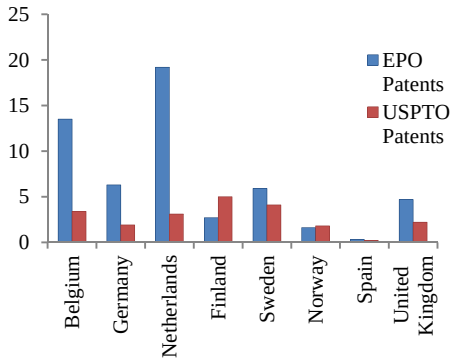
Concerning the Dutch patenting statistics, the OECD (2008) report suggests a strong patent pipeline in the Netherlands, bagging the top position in Europe in terms of total patent intensity during 1995-2005, with more than 260 family of patents per billion US$ of R&D investment. However, although it ranks as one of the topmost European countries in ICT patents, its pharmaceutical patent score is not as inspiring. Conforming to these Dutch patent statistics, fig.1.6 provides evidence to a considerably lower percentage share of pharmaceutical patents on the total number of patents.

Fig. 1.6: Percentage share of pharmaceutical patenting on total patenting in the Netherlands



Source: OECD Patent Database, Feb., 2003.

Fig. 1.7: Biotech patents in some EU countries



Source: European Commission Report, 2003

Note: Biotech patenting activity is corrected for the size of country.

As evident from fig. 1.7, the number of EPO and USPTO patents, corrected for the size of the individual European countries, places Netherlands in the foremost position in terms of EPO patent applications. Although the Dutch pharmaceutical and biopharmaceutical firms lag behind the overall patenting in the country, they perform relatively well in the category of pharmaceutical patent applications in Europe, exhibiting a steady increase over time (Enzing *et al.*, 2004). However, Nefarma (2011) report asserts that, one of the serious impediments in the Netherlands' pharmaceutical patenting activity is the excessive meticulous nature of the therapeutic innovation process as well as the administrative conduct, which causes the clinical research, registration and the final introduction in the market to exceed the statutory term. Consequently, investment in R&D for the Dutch pharmaceutical industry becomes less lucrative and therefore, companies tend to steer away from patenting.

Although the maximum life of a patent is 15 years after it is granted, the effective time of the patent to remain in the market is around 11 to 12 years before it expires. With the expiry of patent protection, the generic pharmaceuticals come to the fore

and actively reproduce cheaper imitations of the branded drugs. Moreover, even before the influx of the generics, the first in class branded products face intense competition from other alternative branded drugs next in line in the process chain. Therefore, it sometimes becomes tedious for the innovative pharmaceuticals to repatriate the R&D cost through patent protection. In case of Netherlands, the pharmaceutical companies encounter even more obstacles as their effective pharmaceutical patent life gets further curtailed due to the delay in the introduction of a therapeutic drug in the market.

However, albeit all the confronting hurdles, it is evident that the introduction of new pharmaceutical products incurs enormous benefits in extending human lives and eradicating or reducing many diseases (Lichtenberg, 2004 and Murphy and Topel, 2006). Hence, it is crucial to innovate new therapeutic products, where the intellectual property protection is fundamental to commensurate the underlying risks, costs and time involved in the process of drug discovery.

1.2.3 Competition

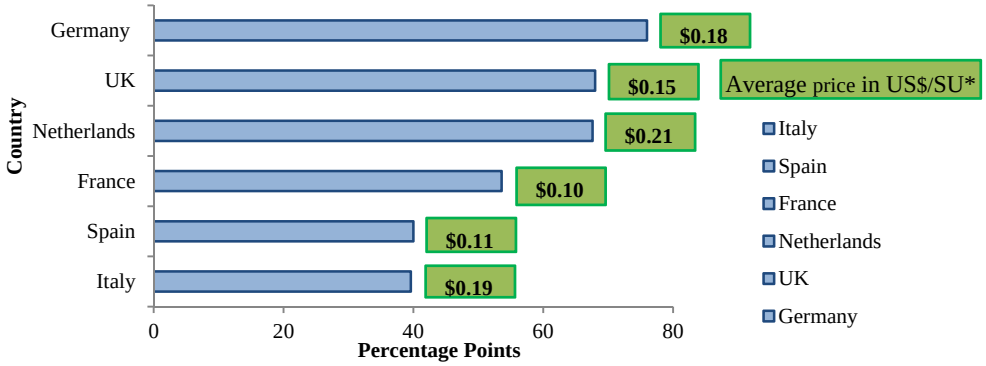
The Dutch economy is highly internationally oriented and an open economy. However the pharmaceutical sector is unique in its stature and cannot be typecast with other manufacturing industries. The multifarious effects due to the co-existence of generic and innovative drug companies, as well as the contending innovative firms and their strategic intellectual property protection, makes competition strategies of this industry even more complex and intriguing. Due to the vast heterogeneity in the pharmaceutical regulatory framework in different countries, their competitive strategies differ. Focusing on the Netherlands' pharmaceutical industry, it has undergone a complete overhaul by gradually moving from a supply-side regulation to managed competition.

Bevan and Van de Ven (2010) posited that, the first precondition to fulfill managed competition is price regulation. While most developed countries have moved from

tight regulatory measures for drug prices to price negotiations, the Netherlands has witnessed an exceptional situation. In 1990s, Netherlands have moved towards drug price regulation, thereby attempting to force the price of drugs down by reference pricing system. Taking into account the Medicine Price Act in 1996 that based drug prices on Belgium, France, UK and Germany, this system appears to be an effective method of cost containment in the context of Netherlands. Since an enormous cost containment pressure makes it difficult to bring products to market with a favorable return, price regulation seems to be a correct measure by the Dutch pharmaceuticals.

Despite pursuing drug price regulation, the Netherlands witnesses high drug costs as, cheaper alternative drugs are replaced by their more expensive alternatives (Mossialos and Oliver, 2005). But statistical evidence suggests that there exists a significant generic uptake in the Dutch pharmaceuticals in recent times, which is mainly due to the high margins enjoyed by the pharmacists when they dispense a generic drug. Based on the CBI market survey (2010), a stable and well-established generic pharmaceutical market is gaining a foothold in the Netherlands, which is steadily uprising. Hence, a growth in prescribing generic drugs and also the steady fall in the price of the older drugs have kept the pharmaceutical prices in the Netherlands under control over the past decade. It is evident from fig. 1.8 that Netherlands has a high generic penetration in the pharmaceutical market compared to the other five prominent European countries, ranking just after Germany in terms of generic penetration. Nevertheless, the information provided in the diagram also entails a much higher generic drug price in the Dutch market as compared to the other EU countries.

Fig. 1.8: Generic market volume penetration with generic price by country: Netherlands vs EU5



Note: *SU: market segmentation universe
Source: IMS Health, MIDAS, Market Segmentation MAT (September, 2010)-Ethical, Retail, Unprotected Market based on SU, Netherlands Xponent

Inspite of a relatively higher average generic drug price based on the categorization of the various target groups of consumers, an adequate cushion on the prices of medicinal products exists due to the consolidated Dutch health insurance facility. However, since all individuals are given equal treatment within the health sector, it has caused prominent obstacles in the introduction of user charges. Due to the huge cost-containment pressure, Netherlands underwent a massive change in their healthcare policy in 2006, introducing the Health Insurance Act, where every citizen of Netherlands is entitled to purchase a public health insurance policy and the government only provides the guidelines. The innate characteristics of this health care measure have been discussed in subsection 1.2.1. This reform policy on managed competition has effectively increased the buying power of the insurers, who are given the freedom to choose their own insurance agency. But this may put the generic industry at a disadvantageous position due to the cost-containment preferencing. On the other hand, the preponderant shift of power to health insurers puts the long-term sustainability of pharmaceutical welfare at risk.

In addition, apart from price competition and administrative management strategies, the inherent competition in the drug production market and the apparent monopolistic behaviour of the innovative firms to shield off their breakthrough

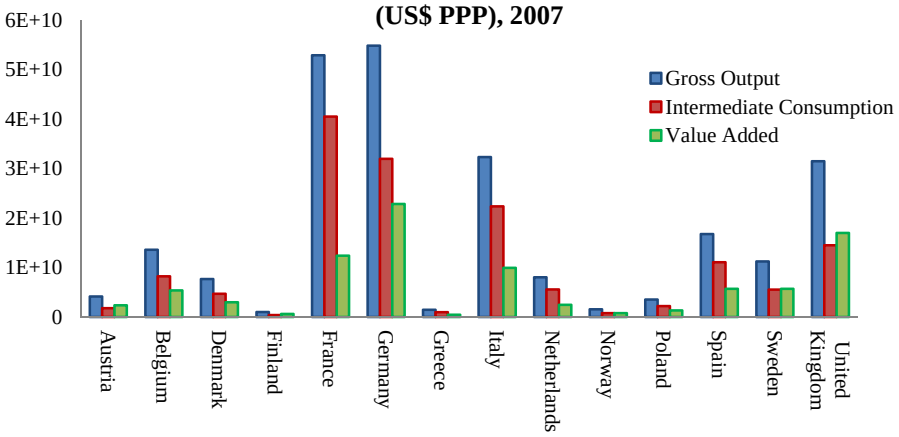
drugs from imitation is an all pervasive phenomena in the drug industry. Therefore, it remains as a challenging task for the policymakers to manage the intricacies of the competition strategies prevalent in this sector.

1.2.4 Firm performance

Although the pharmaceutical industry is a science based and research-oriented industry, it is also involved in large scale production under constant surveillance on accuracy in terms of quality of the final products. The amount of bulk drug production is most compelling for antibiotics, vitamins and aspirins. However the marked ageing population in the Netherlands also calls for a huge demand for cardio-vascular drugs, analgesics and oncologics. Therefore, we briefly highlight on the overall firm performance for the Dutch pharmaceutical firms, based on their productivity.

Pharmaceutical production not only plays a cardinal role to meet the needs of the patients, but it also contributes significantly to the respective country's economy. The diagram below illustrates the gross output level, intermediate consumption and value-added of some countries in Europe. Clearly, countries like Germany, France, Italy and United Kingdom are positioned prominently higher in terms of their demand and supply credentials, in comparison to the other countries including Netherlands. As reported by CBI market survey (2010), Netherlands ranked eighth among the EU countries in pharmaceutical production during 2006 and 2007, with an approximate share of 3.5% of the total EU production in pharmaceuticals. In essence, the pharmaceutical production in Netherlands was not as promising as the total production in EU, which accounted for an annual average increasing rate of 2.3% between 2003-2007. A major reason for the Netherlands' repressing medicinal production status is the inadequate number of multinational pharmaceuticals in this country (Espicom, 2009).

Fig. 1.9: Pharmaceutical production and value-added in some countries of Europe (US$ PPP), 2007



Source: OECD, 2012: StatExtracts

Likewise, in terms of the level of pharmaceutical consumption, the Netherlands was graded as the seventh largest consumer market in EU. However, while production level declined by 0.3% between 2003-2007, consumption level witnessed an upturn with 7.3% annual average rate during the review period. The increased consumption level can be attributed to the Netherlands' ageing population, the introduction of the Dutch Health Insurance Act in 2006 and the expansion of the non-prescription pharmaceutical market. The gap between domestic production and consumption is well compensated by the pharmaceutical imports across the borders. We take up this issue on the Dutch pharmaceutical trade in the next chapter.

Nonetheless, the presence of an efficient workforce helps the Netherlands' pharmaceutical industry to keep pace with the productivity level of the reigning countries specialized in the field of drug production. Additionally, the provisions provided by the health care act of 2006 and the modernisation of technology have also accelerated the productivity level to a substantial extent. However scepticism still pervades on how the long run productivity level can effectively curb the drug prices. Notwithstanding the criticism, the present policy alternatives are likely to

boost the pharmaceutical productivity performance, owing to a more strategic market competition and innovation incentives.

1.3 Conclusion

Along with the uncertainly, risk and cost involved in the pharmaceutical market, the recent economic recession has also played a vital role in stagnating this sector to some extent. However, the drug sector is expected to be resilient to all odds as it is not only the linchpin to human survival, but also a potential channel for job creation, technological advancement and economic growth. With effective governmental aid, research cohesion and its eventual propagation; developed countries like the Netherlands have a wider perspective and scope to its betterment.

But the pharmaceutical status of different geographical territories experience widely varied economic, social and administrative conditions, thereby operating in different strata of the world pharmaceutical activities. While developed countries are more adept to innovating new curatives, the developing countries are deeply engaged in generic drug development. Although the innovation performance of the less developed countries are at the rudimentary level, their openness in the world market has intensified the export and import conditions, along with the occurrence of numerous mergers and acquisitions. Needless to say, the global pharmaceutical structure is undergoing significant alterations. Therefore, in view of global functionalities of our concerned sector, we discuss on the world pharmaceutical performance in our next chapter, before proceeding with our analysis further on the Netherlands pharmaceutical industry.

Chapter 2

THE WORLD PHARMACEUTICAL INDUSTRY: A COMPARATIVE APPROACH

2.1 Where does Europe's pharmaceutical industry stand?

In the first chapter, we have provided a detailed theoretical outline for the Dutch pharmaceutical sector. In consonance with the introductory chapter, chapter 3 onwards primarily deals with an exhaustive micro level investigation of the various firm level determinants for the drug industry in Netherlands. However, before focusing on the predominant area of this thesis, we aim at reviewing the pharmaceutical industry in a global platform. In other words, this chapter essentially deals with the performance and characteristics of the pharmaceutical industries operating in different parts of the world, under varied socio-economic paradigm.

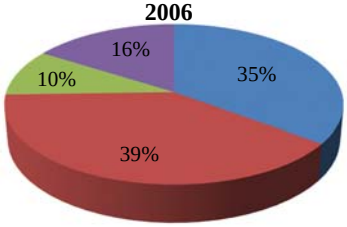
Since the preceding chapter has already taken into account the position of Netherlands with respect to the other neighboring EU countries of significant importance, in this chapter we contemplate and compare the pharmaceutical industry in the whole of Europe and the other nations of the world. To begin with, we first benchmark the performance of the drug sector operating in Europe, in regard to those countries which are operating in the later stage of their development process. Fig. 2.1 testifies the fact that, USA and Japan are the fundamental competitors of Europe's drug market as emphasized by their percentage share of the various components that determine the performance of their respective markets.

In totality, fig. 2.1 determines that the stature of Europe's pharmaceutical market evidently lags behind that of USA.

As has been specified in the EFPIA report (2008), the USA dominates the global pharmaceutical market, possessing the greatest share of pharmaceutical revenues and R&D expenditure. In addition, this report puts forth the data documentations that affirm the supremacy of the US pharmaceutical industry. Based on their study for the period 1990-2007, R&D investment increased by 5.3 times in USA, whereas it increased by only 3.3 times in Europe. Plus, the pharmaceutical sales accounted by the North American pharmaceutical companies measured 45.9%, while Europe experienced a 31.1% of the global pharmaceutical sales in the year 2007. A study by Gambardella *et al.* (2000) reveals that the European pharmaceuticals face the predicament of a comparatively lower growth in the demand for new medicines, which primarily hinders the sales in Europe's drug sector. Therefore, although the European pharmaceuticals possess the ability and the infrastructure to develop breakthrough drugs, the American drug industry excels in a better market strategy to propagate their products at the international forefront.

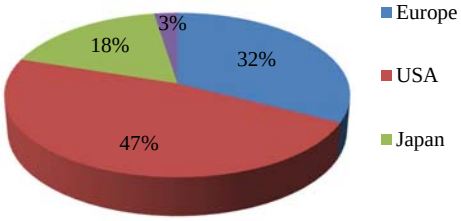
Fig. 2.1: Characteristics of the world pharmaceutical industry

Percentage of world pharmaceutical industry production at ex-factory prices, 2006



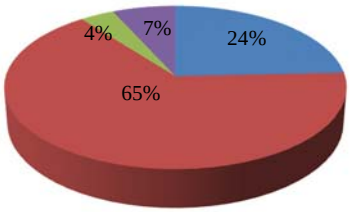
Source: EFPIA member associations, P*h*RMA, JPMA, OECD, IMS Health, EFPIA calculations.

Percentage of firm origin of top 40 pharmaceutical companies w.r.t R&D investment, 2006



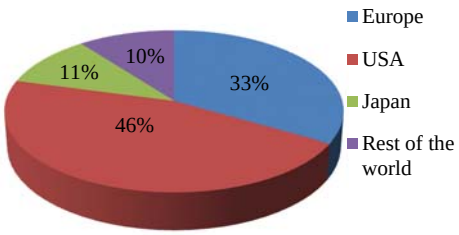
Source: UK Department of Trade and Industry, 2007 R&D Scoreborad, EFPIA Calculations, Own calculations.

Percentage of sales of new medicines, 2002-2007



Source: IMS Health Midas MAT, Dec. 2007

Percentage of New Chemical Entities (NCEs) launched, 2003-2007



Source: SCRIP, 2003-2005, CMR, 2006-2007, EFPIA calculations, Own calculations.

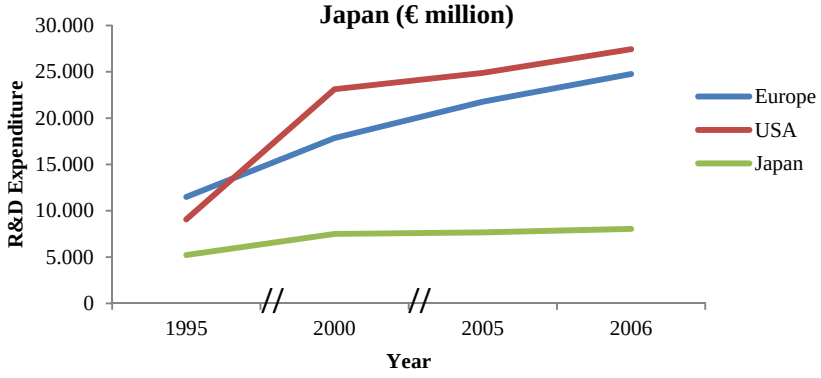
Note: Europe includes non-EU member countries and CIS surveys

Further, the above charts illustrate that the pharmaceutical industry in Japan scores after USA and Europe, making a niche for itself as one of the top ranking countries in terms of performance and growth in this sector. Taking into account the fact that Europe is a congregation of a number of countries, the comparison put forward between Europe, USA and Japan does not precisely map their performances based on country profiles. Thus, considering their performance based on the geographical

territories that demarcates the different nations, the pharmaceutical market in Japan ranks next to USA, amounting to US$50 billion of sales value in 2003, which was equivalent to the combined sales of France, Germany, UK and Australia (Mahlich, 2007). Even though the sales value of Japan is far lower than that of US, due consideration should be given to the fact that Japan has a geographical area of only one twenty-fifth of U.S.A, with around 40% of total US population and one-third of its GNP.

Considering the pharmaceutical innovation pipeline, US also enjoys a preeminence over Europe, with an estimated growth in R&D investment between 1990-2007 to be 5.2 times, as compared to a 3.3 times growth rate in Europe (EFPIA Report, 2008). Regarding the Japanese pharmaceuticals, Mahlich (2007) expounded on the rise in its in-house R&D activities after the implementation of the deregulation policies in the 1990s. Fig. 2.2 illustrates the trend in R&D expenditure for the pharmaceutical industry in Europe, US and Japan for 1995, 2000, 2005 and 2006. Clearly, the pharmaceutical R&D expenditure of US played a second fiddle to that of Europe during 1995, but progressed persistently in the following years and surpassed European pharmaceuticals with a consistently higher R&D expenditure.

Fig. 2.2: Pharmaceutical R&D expenditure in Europe, US and Japan (€ million)



Source: EFPIA member associations, PhRMA, JPMA

Note: The figures are converted into a common currency (€) using the current exchange rates

Apart from those regions which have been able to establish a profound global footprint in the pharmaceutical arena, countries like India and China are marked by bulk production of generic drugs and have a huge market demand, owing to its enormity in population. Moreover, the current trend in establishing pharmaceutical R&D centers in the emerging economies is increasing at a faster pace. Besides the Asiatic nations, the Australian pharmaceutical sector experiences a sluggish innovation performance and performs distinctly lower than the pioneering countries. However, considering the fact that Australia contains only about 0.3% of the world population which consumes only 1% of the global pharmaceutical sales, it is for obvious reasons that its impact is miniscule in the global context. Nevertheless, a report issued by the Australian government on pharmaceutical industry profiles states that Australia ranked twelfth in world pharmaceutical sales in 2009. Concerning the rest of the world, there are still numerous countries which are still struggling to flourish in their indigenous drug production. The African countries are amongst those countries, which are still in the budding phase of their pharmaceutical development, having low production in drugs and hence, mostly resorting to medicinal imports. Comtrade (2012), providing international import and export data, estimated that India and China accounted for 17.7% and 4.1 % of imports in the pharmaceutical sector in Africa in 2011, as compared to 8.5% and 2.0% consecutively in 2002. However the African pharmaceuticals are evolving fast and they are providing rich opportunities for the multinational companies to set up their businesses in Africa; such that the potentials that these African countries have in store, are effectively utilized.

2.2 Trade relations of the Netherlands with other countries

In this section, we resort back to our country of interest i.e. Netherlands, where we investigate on its leading suppliers for pharmaceutical products. In a way, we attempt to take a note of those countries that channelize their pharmaceutical trade in the Dutch economy and their trends thereof. According to the CBI market survey

(2010), the most important trading partners for the Netherlands mostly comprise of the EU countries, constituting 57% of its imports in 2008. The primary EU suppliers to Netherlands are Germany, France, Belgium, Italy and UK; who contributed to a 5% rise in the Dutch pharmaceutical imports between 2004 and 2008.

Several countries outside the EU also have a substantial share of Dutch imports, of which USA plays a leading role. In the year 2008, the medicinal imports from USA amounted to 22% of the total Dutch pharmaceutical imports, while Singapore and Switzerland also export substantially to the Netherlands, with a share of 7.6% and 5.1% in the same year.

However, the imports of medicinal products from developing countries amounted to a meager 0.4% in 2008. But the most significant supplier amongst them was India, which marked a 0.3% of medicinal export share to Netherlands. Although the developing countries accounted for a very small share of imports, there was a perpetual increase in their average annual rate of imports by 29% from 2004 to 2008. Table 2.1 demonstrates the percentage share of pharmaceutical imports to the Netherlands from the different regions of the world pharmaceutical market.

Table 2.1: Percentage of pharmaceutical imports to Netherlands

	2004	2006	2008
Europe	53.59%	47.53%	57.13%
Outside Europe (excluding developing countries)	46.24%	52.26%	42.46%
Developing Countries	0.17%	0.21%	0.42%

Source: CBI, 2010; Own Calculations

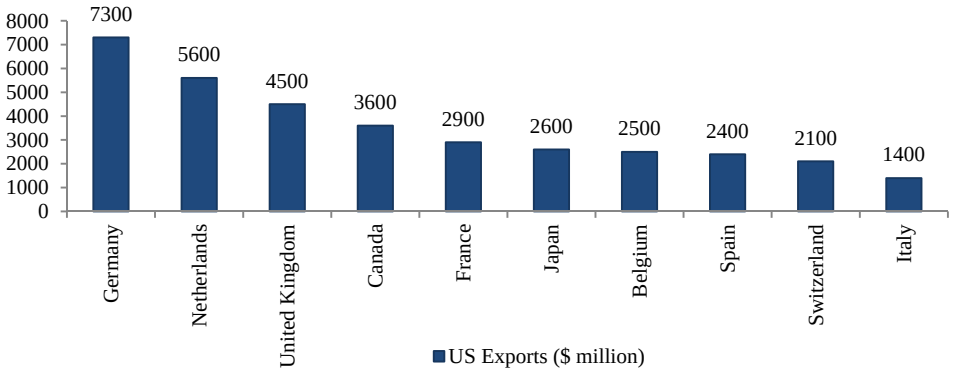
In view of the pharmaceutical sectors of those countries which have disseminated their trading activities with the Netherlands, we elucidate further on the US and the Indian pharmaceutical industry in greater details. The reason behind reviewing exclusively these two countries' pharmaceutical regime is based on the fact that, they are the most prominent non-European countries from the group of developed and developing nations respectively, that export their medicinal products primarily

to Netherlands. Therefore, it is worthwhile to explore the nuances of their pharmaceutical performances that define their trends and characteristics, in the context of a developed and a developing economy.

2.3 The pharmaceutical industry in USA

The US pharmaceutical industry undeniably dominates the world pharmaceutical arena, in terms of both commercial prowess and innovation excellence. The Economic Census (2007) states that, there exists 1,552 pharmaceutical companies in U.S., with an estimated shipment of products valued at $195 billion in 2008. As has been enumerated before, the U.S. pharmaceuticals comprise of the maximum share in the world market, propagating their trade in every corner of the world. The diagram below demonstrates the ten most prominent countries which import US pharmaceutical drugs. The figures represent the export values in the year 2009.

Fig. 2.3: Export by U.S. pharmaceuticals in the top 10 destination countries in 2009



Source: U.S. Trade Patterns for Pharmaceutical Medicines and Manufacturing Sector (NAICS 3254): International trade Administration (http://trade.gov)

As can be noticed from the diagram, Netherlands constitutes as one of the topmost nations that import US drugs. However the trade statistics asserts that the total medicinal import is considerably greater than the exports in U.S., leading to a negative trade balance. This emanates from the fact that the U.S. pharmaceutical

companies have established their manufacturing units in other countries to have more proximity to their target markets and also evade the taxes and tariffs involved in exportation. Also these U.S. pharmaceuticals situated in foreign countries export manufactured drugs to their own country. Therefore the export-import statistics do not give a true picture of the market value of U.S. pharmaceuticals that have spread across the globe, with an annual turnover of $300 billion in 2009 (IMS health, 2010).

Important policy changes in health that affected the U.S. pharmaceutical industry

The U.S. pharmaceutical industry has the largest prescription drug market, in spite of having no national coverage for health. However the health care reform in 2010 with the introduction of Patient Protection and Affordable Care Act (ACA) is expected to have its effect in the pharmaceutical pricing and its subsequent profitability. Concurrently, the equality to medicinal access entailed the pharmaceuticals to lower their drug prices. However the bigwigs in this industry tried to safeguard their pricing strategy, citing the cost-containment by medicinal intake on time to prevent hospitalization, which would be more cost adverse for the patients. A stronger argument as was put forth earlier by Abelson (1993) was the consequent hindrance in new drug development with the price reduction for existing drugs. In the light of this debate, we introspect on the several changes in the erstwhile U.S. public health policies over time.

- 1965: Introduction of Medicare and Medicaid to ascertain medical care for people above 65 years irrespective of their economic and health status. However Medicare was slightly altered in 1972 to accommodate only those citizens above 65 having health disabilities. The Medicaid law additionally provided basic medical access to all Americans below a certain income level.
- 1988: Deductibles ascertained for certain prescription drugs under the Medicare Catastrophic Coverage Act.

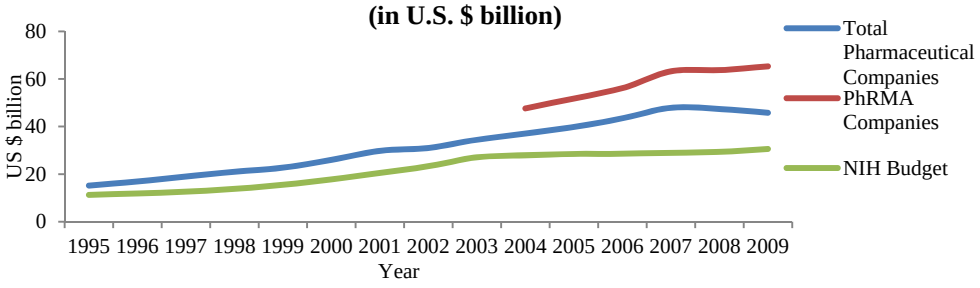
- 2010: With the rising of the prescription drug costs, Medicare policy was altered, where the beneficiaries were entitled to pay up to $310 as the baseline for prescription drugs.
- 2010: Enforcement of the Patient Protection and Affordable Care Act that provided health care insurance for young adults until the age of 26 years.

In the course of the implementation of ACA, the pharmaceutical companies experienced its share of impact which involved an increased sale for prescription drugs. Additionally, with the latest insurance status, it is expected that U.S. will experience a demographic shift in the future with a greater life expectancy and increased sale of medicinal products.

Innovation, Competition and Firm performance: From the American pharmaceutical aspect

The U.S. pharmaceutical industry is one of the most crucial and competitive sector with a very robust innovation system and infrastructure. Needless to say, the U.S pharmaceuticals are a store house of innovation potential, expending 70% of the world pharmaceutical R&D expenditure in 2009 (PhRMA, 2010). Although the pharmaceutical industry has experienced a perpetual rise in the R&D expenditure over time (as can be seen in fig. 2.4), studies like Cockburn *et al.* (2004) has emphasized on the inadequate utilization of R&D investment in the development and discovery of breakthrough drugs.

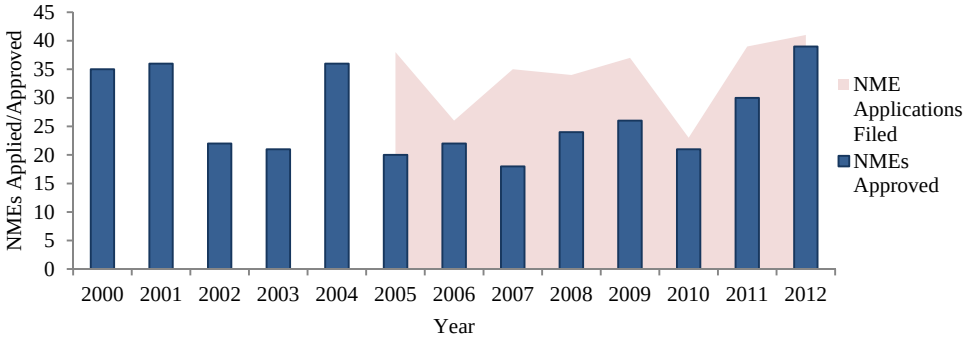
Fig. 2.4: R&D expenditure in the U.S. Pharmaceuticals, 1995-2009 (in U.S. $ billion)



Source: PhRMA, Burill Company, National Institute of Health (NIH) Office of Budget.

Emphasizing on the innovation output, Cockburn *et al.* (2004) asserts that the New Molecular Entities (NMEs) is extremely fluctuating over short time periods and does not give a true picture of the actual yield in pharmaceutical research. Since NMEs do not consider vaccines, large molecule medicines and other biologics; NME count cannot be used as a yardstick to measure the quality of innovation output.

Fig. 2.5: NMEs applications and approvals for the U.S. pharmaceutical industry



Source: Cockburn (2004),

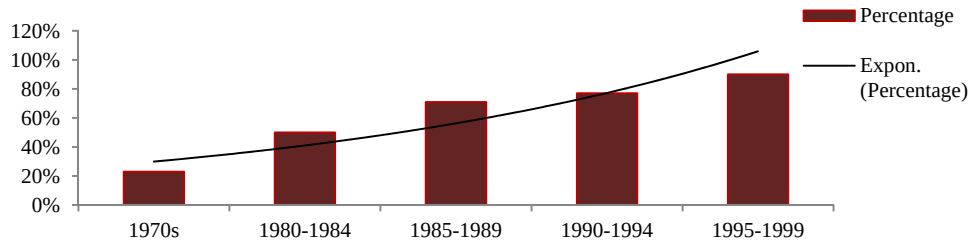
Food and Drug Administration, U.S. department of Health and human Services, Jan. 2013.

However we cannot negate the fact that, even though the U.S. has the largest pharmaceutical research base and a strong governmental funding for innovative research, the rate of pharmaceutical innovation output seems to have stagnated to a certain extent not only in U.S. but in the entire world pharmaceutical market, regardless of a consistent increase in the R&D expenditure (Pammolli *et al.*, 2011). As Scherer (2007) scrutinizes on the intricate drug discovery process, it is undeniable that the entire process is a daunting task with huge time consumption. In U.S., the average time involved to pass from phase I through phase III is a span of six to seven years. Concurrently, the regulatory process for approval consumes an additional one to two years. Plus, it involves a complicated pathway of several

additional years before a drug is used for commercial purposes, after taking into account all the precautionary measures for human consumption.

Furthermore, the innovation activities are juxtaposed by competition, that sometimes affect the innovation productivity in this sector. The intensity of competition goes beyond the generic drug market and is even becoming more conspicuous in the invention of new drugs. The diagram below depicts the percentage of new drugs which already had their competitor drugs in phase II clinical testing during the time of approval. As can be seen in the diagram, nearly all drugs in the approval stage had an alternate competing drug at the phase II of clinical trials by 1995. This suggests an increase in competition within the therapeutic categories, wherein the average time for effective market exclusivity for first approved innovative drugs has decreased abysmally over time (DiMasi and Paquette, 2004).

Fig. 2.6: Percentage of first in class medicines having a competitor already in phase II clinical testing at the time of approval



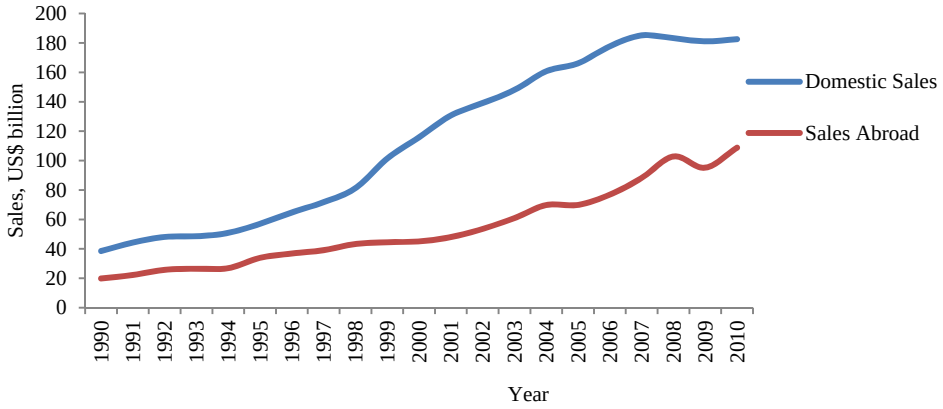
Source: Di Masi, J. and L. Faden (2009), "Follow-on drug R&D: new data on trends in entry rates and thwe timing of development", Tufts Centre for the Study of Drug Development, Working paper.

Since generic infiltration in the market as well as competition within the therapeutic class is increasing in the drug sector, a corpus of vertical competition governs this industry's activities. According to Cockburn (2004), these vertical competition may escalate the level of innovation and productivity, but can also result in wastage of resources and repressing innovation science. Nonetheless, the U.S. pharmaceutical companies located in or outside USA have deftly managed to

collaborate biotechnology units, research hospitals and leading universities. Their symbiotic innovation strategies have brought into effect several effective therapeutic treatments. The prime concern for the recent innovation activities involves the target drugs for the treatment of cancer, autoimmune and infectious diseases, for which rigorous innovation activities to achieve effective medication are in progress.

However, there is no dearth in the production of already existing therapeutic drugs. Both branded and generic drugs have their individual markets and consumers. With the American drug companies spreading their business across the world, they experience the fruits of profit from a global network. A significant rise in domestic as well as foreign sales is seen in fig. 2.7, where the former portrays a higher rate of increase than the latter.

Fig. 2.7: Domestic sales and sales abroad: PhRMA companies, 1990-2010



Source: PhRMA Annual member Survey, 2011

A prognosis on the total factor productivity in this sector reveals that the U.S. pharmaceutical production are more dependent on capital and R&D inputs, rather than labor inputs (Gambardella *et al.*, 2000). The total share of value-added over production value is more pronounced in U.S. than the other developed countries. Based on the Eurostat data, Europe's pharmaceutical industry is found to be more

labor intensive than that of USA and Japan. However the share of value-added net of labor cost on the production value is evidently higher in USA and Japan than in Europe. The increased pharmaceutical innovation activities in the U.S. compared to Europe from 1990s has culminated to a more prominent production value for the US drug industry.

Fig. 2.8: Components of pharmaceutical growth, U.S. pharmaceutical industry, 1992-1997 & 1986-1991

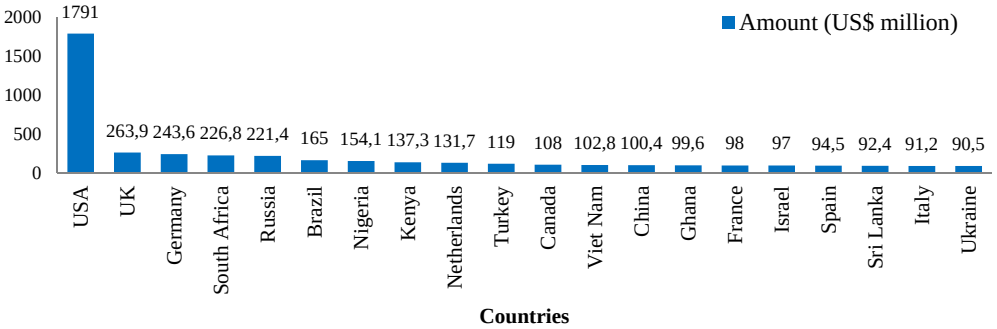
Source: Gambardella *et al.* (2000) calculations

Eurostat data

2.4 The pharmaceutical industry in India

In conjunction with the above deliberation, we continue with our discussion on the Indian pharmaceuticals by first investigating in fig. 2.9, the top twenty countries where Indian drugs are exported. As evidenced by the diagram, Netherlands scores considerable high in terms of importing Indian drugs.

Fig. 2.9: The amount of export of Indian pharma products in the top 20 destination countries in 2010



Source: Directorate General of Commercial Intelligence and Statistics (DGCIS), Kolkata

The considerable thrust for export promotion by the Pharma Export Promotion Council (Pharmexcil) has led to a perpetual improvement in this sector. Furthermore, low cost of scientific pool or innovation, along with well managed regulatory compliance have contributed to its current status. Thus far, the Indian pharmaceutical industry is in the third position in terms of volume of drug production (10% of the global share), and the fourteenth in regard to the drug value (1.5% of global share). According to the Ministry of Foreign Affairs report of 2010, the major reason for the lower ranking in value of Indian medicinal products is due to their exorbitantly low price.

Nonetheless, there exists an array of critical issues and multiple challenges that dominate this sector's functioning. Some major policy issues and the concurrent structural scheme provides the prima facie evidence of the nature of functioning of this novel but complex field of drug manufacturing in the context of the Indian scenario. We therefore aim at discussing on the very nature of the Indian pharmaceutical industry, in order to bring to the fore the divergence in the way this industry operates in developing countries, in comparison to the more developed territories.

Important Policy changes as witnessed in the Indian pharmaceutical industry:

As mentioned in the study by Ghosh and Chakraborty (2012), India witnessed a series of policy alterations that affected the economic, structural and social framework of its pharmaceutical industries:

- 1900-70: dominance of multinationals, having 68% of the share in 1970.
- 1970-1990: The Patent Act of 1970 lead to the amendment of the Patent Act of 1972, which resulted in a bulk production of drugs within the country. During this time, the Indian pharmaceutical

industry developed to become a leading global player in terms of generic drug production.

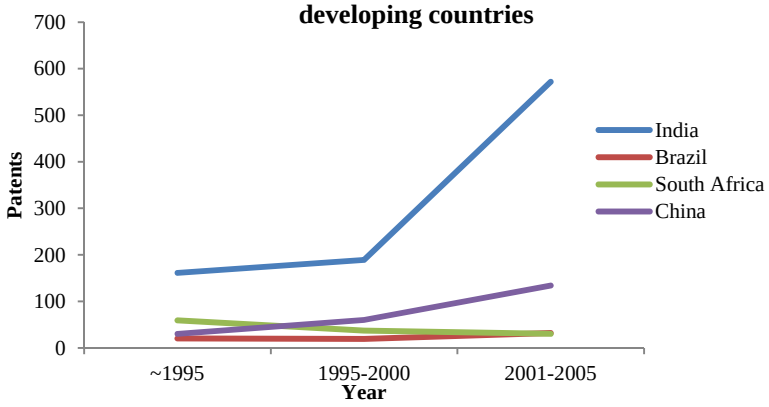
- 1990-2005: Introduction of trade liberalization measures. The TRIPS agreement signed by the Indian government in 1994 caused a prominent increase in production due to an increased competition level between the domestic and foreign competitors.
- From 2005: Conforming to the trade liberalization scheme, the amendment of the Patent Act of 1970 lead to the introduction of patents in the pharmaceutical production in India, along with the compulsory implementation of the Good Manufacturing Practice (GMP).

In the purview of the above historical digression, we attempt to cast some light on how these policy changes have further affected the Indian pharmaceutical sector's innovation and performance based on its transitional firm characteristics.

Innovation, Competition and Firm performance: From the Indian pharmaceutical aspect

The fundamental change that the Indian pharmaceutical industry witnessed was the introduction of intellectual property regime under the influence of WTO in 2005. The amendment of the patent Act of 1970 was completed in 2005, whose foremost criterion was the introduction of patents in pharmaceutical production. This resulted in the emergence of research centers for the pharmaceutical industries in India, along with the outsourcing of clinical trials to the Indian pharmaceutical companies and emergence of other international business opportunities. As posited by Kamble *et al.* (2012), the shift in the focus for Indian pharmaceuticals from generic production to NCEs has resulted in an appreciable patenting performance.

Fig. 2.10: Pharmaceutical patents (USPTO) of some developing countries



Source: www.uspto.gov

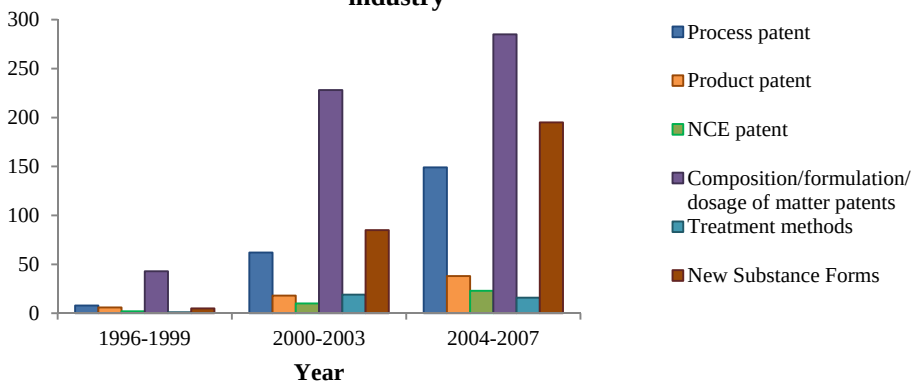
Chaudhuri,S. (2007), "Is product patent Protection Necessary in developing countries for innovation? R&D by Indian Pharmaceutical Companies after TRIPS", IIM Calcutta Working Paper

Fig. 2.10 provides evidence of an encouraging increase in the number of patents by Indian pharmaceutical industry in the U.S. patent office, especially after the year 2000. Besides, the number of Indian patents is distinctly higher than other developing countries which includes Brazil, South Africa and China.

However, to examine in details the ramifications in the innovation process by the Indian pharmaceutical industry over time, we present in figure 2.11 a comprehensive overview of the innovation patterns practiced by the pharmaceutical companies in India and filed at the US Patents and Trademark office (USPTO). Although the diagram bears testimony to the increased research activities in the post- TRIPS period, it is seen that the Indian pharmaceuticals have given maximum emphasis to the identification and reduction of impurity levels, formulations and appropriate dosage depictions of medicinal products. Furthermore, the number of process patents has increased with a greater momentum than the product patents over time. Therefore, the emerging pattern provides a compelling evidence of innovation performance in the modifications of

existing drugs, rather than invention of novel therapeutic products. Nonetheless, a significant overall improvement in the foray of innovation competency is observed after the introduction of the TRIPs policies.

Fig. 2.11:Pattern of innovation in the Indian pharmaceutical industry



Source: www.uspto.gov

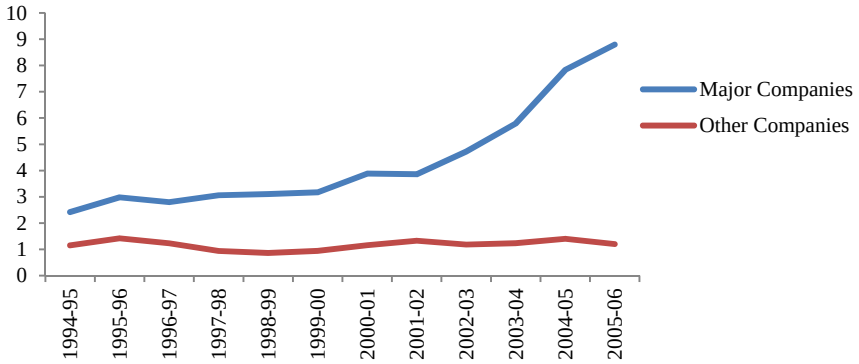
But the downside of these policy implementations was the economic scarcity for small pharmaceutical industries, who could not afford to cater to the new regulations in spite of governmental aid. The upsurge of the new patent law has caused a huge blow to the generic drug supply, leading to the closure of numerous pharmaceutical plants which are not licensed to manufacture patented medicines. In addition, the inability to comply with the Good Manufacturing Practice (GMP) of WHO by several small pharmaceutical plants, led to their exit from the market or had their licenses suspended. As a result, several Indian companies were taken over by their foreign counterparts. This brings forth the question of whether patent protection jeopardizes or preserves the growth in the pharmaceutical sector of a developing economy.

In a study by Siebeck (1990), it has been asserted that, significant patent protection is redundant until a country reaches a level of stability in its development process. A later study by Maskus (2000) also contends that the costs for stronger patent

policies in developing countries can predominantly offset its benefits, and may even result in curbing its technical changes in the course of development. But, in stark contrast with his previous viewpoint, Maskus *et al.* (2005) empirically established that there might exist positive relationship between stronger patent protection and economic development in a study on China. This empirical evidence is contrary to a study on Japan by Sakakibara and Branstetter (2001), who found that stronger intellectual property rights does not corroborate to higher innovation. In spite of the various conjectures, it is indisputable that the effect of strong patent protection is specific to countries and industries. However, it is beyond the scope of our study to empirically investigate on whether the introduction of patent protection has been beneficial to the Indian pharmaceutical industry. Therefore, the rationale for the introduction of patent system in developing countries like India, where many small industries lack infrastructure and adequate financial foothold, remains an open question.

Nevertheless, as pointed by Sharma (2011), the changes in the TRIPS policies have undoubtedly invoked the development and formulation of new chemicals entities (NCEs) in the Indian drug sector. As depicted in figure 2.12, the R&D expenditures for the major pharmaceutical firms in India have soared constantly higher over time, with the increase being more prominent after 2002. This noticeable improvement in the Indian pharmaceutical R&D is a desirable outcome for the overall development of the economy. However no increasing trend was found for the other lesser important pharmaceutical companies. In actuality, their R&D activities remained unaffected by the policy changes that were underway. Thus, R&D investment has been diligently practiced by the most prominent pharmaceutical minorities, while the majority of the Indian pharmaceutical firms did not experience much innovation activities at least until 2006.

Fig. 2.12: R&D expenditure of the Indian pharmaceutical industry



Source: PROWESS database, Centre for Monitoring Indian Economy.

This is mainly due to the fact that, most of the Indian pharmaceutical firms have inadequate funds and infrastructure for undertaking R&D investments independently. Hence most of these companies are engaged in bulk production of generic drugs. On the other hand, the R&D performing firms are, in essence, involved in the modification of existing chemical entities and development of generics. Hence, they are yet to be fully involved in the formulation of new medicinal products, although a few engage themselves in the development of NCEs. It is of utmost importance for the Indian pharmaceuticals to be motivated in basic research for new drugs, especially for those diseases that solely exist in developing countries and neglected at large. Despite the need to promote target identification of new drug development, the Indian pharmaceuticals face the predicament of huge cost and uncertainty that is involved in its later stage of development process. Therefore, the encouragement for more R&D activities through financial support and joint ventures is perhaps a better stepping stone rather than the compulsory entrance of patent systems.

Undoubtedly, the Indian pharmaceutical is undergoing a transition process and consequently, the disparities in opinions and perspectives are evident. However, in the context of economic competitiveness, it is rightly articulated by Abrol *et al.*

(2011) that, the fundamental goal of the developing countries in instrumenting their competition policies, is to pursue sustainable development. Thus, the provisions for competition need to be fine-tuned in order to attain social, political and economic development in this industry, rather than exclusively increasing their exports.

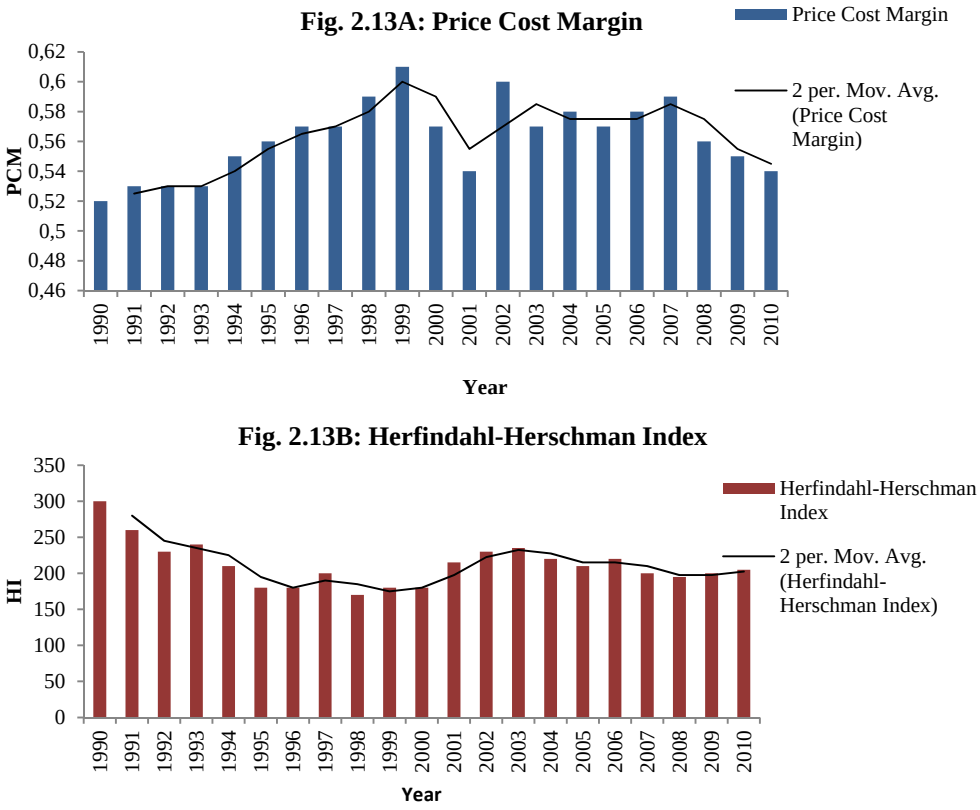
Although the introduction of product patent under the WTO TRIPS agreement intended to equalize the pharmaceutical policies of all countries alike, it should be stressed in this regard that the pharmaceutical industry in India cater to the needs of those people whose majority are below or at the edge of poverty line. Moreover no public insurance policies exist for medicinal or therapeutic products. Hence, the entire medicinal cost is borne by the people which makes it even more crucial for keeping the drug prices at an affordable range. Bearing in mind the absence of adequate health security from the Indian government, the need of the hour is to formulate effective public policies to regulate the level of competition in the pharmaceutical sector.

With the Indian pharmaceutical companies entering the international market, it is anticipated that they become more competitive in the recent years. Besides, the Competition Act of 2002 which came into effect in 2009, aims at minimizing the restrictive trade policies. Despite the Indian pharmaceuticals are gaining better footholds in the international market with their export practices, the high price of the imported branded drugs are mostly unaffordable for the poorer section of the Indian economy. Another imperative argument is whether the latest mergers and acquisitions by foreign pharmaceutical firms are inadvertently affecting the competition in the Indian economy. Recently, the National Pharmaceutical Pricing Policy (NPPP) has been introduced which intend to promulgate a market based approach rather than a price based approach for pricing the medicines. However, it is crucial and strategically superlative for the Indian pharmaceuticals to engage in a more prolific competitive regime within the boundaries of the country. In addition, governmental procurement, support and distribution of drugs is perhaps a

beneficial step to catalyze competition, control drug prices and improve its quality, along with providing greater availability of medications to all the people dwelling in the different strata of the economy.

In a comprehensive study by Bhattacharjea and Sindhwani (2014), an in-depth analytical deliberation concerning the pharmaceutical competition in the Indian economy has been put forth. In this study, the competition within the Indian market is evaluated using the Price Cost Margin (PCM) and the Herfindahl-Herschman Index (HI) indicators. For this, firm level data has been accrued from the PROWESS database and applied for the calculation of PCM and the HI for the Indian pharmaceutical industry. Fig. 2.13A reveals that the PCM exhibits a decreasing trend in the recent years, starting from 2007. As conceived by Bhattacharjea and Sindhwani (2014), the plausible reason for this decline in PCM, indicating an increase in competition, is due to the influx of newer firms, increase in imports or changes in the regulatory policies. The moving average trend line represents an increasing trend in the initial years, until 1998. Post 1998, the PCM witnesses a sharp decline, which is in congruence with the latest years.

Fig. 2.13: Competition in the Indian pharmaceutical industry



Source: PROWESS database

Bhattacharjea, A. and F. Sindhwani (2014), "Competition Issues in the Indian Pharmaceutical Sector", Centre for development Economics, Delhi School of Economics.

Regarding the HI in fig. 2.13B, the concentration measure shows an approximate downturn in the later years. However a more prominent down sliding of firm concentration is viewed for the years after 1990. In both the cases, an increased level of competition can be inferred. But the period between 2000-2003 marks an increase in concentration which is presumably due to more merger activities and restricted entries. This finding is homologous to Kaur (2012). However, no clear comparability can be drawn from the two different measures for calibrating the

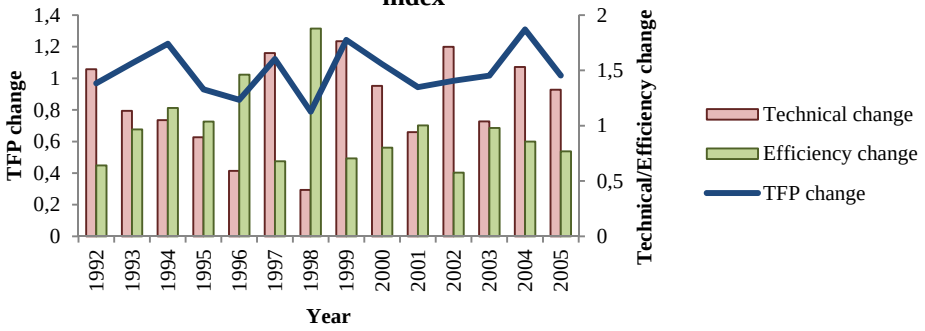
changes in competition over time; as they reflect little or no congruence to each other. Nonetheless, the latest increase in competition is a desirable phenomenon and it is expected that the public policies harmonize with the current needs of the economy to promote healthy competitive actions, within the complexity of the pharmaceutical industry.

Finally, we briefly discuss on the productivity and performance of the Indian drug industry. By focusing on the productivity frontier approach to justify firm performance, Mazumdar (2013) has reviewed several literatures that have conjectured on the productivity of the Indian manufacturing firms after the enforcement of the liberalization policies. Studies like Balakrishnan *et al.* (2000) and Srivastava (2001) have purported that the productivity enhancement in the post liberalization era is largely by the multinational companies which have advanced technological amenities at their disposal. Likewise, Saranga and Banker (2010) appraises the Indian pharmaceutical industry and arrives at a similar inference, as it detects better and more developed product portfolios for the more innovative and technologically advanced multinational pharmaceuticals. As a result, the domestic firms that practice imitation of generic drugs are less technologically efficient to upgrade their productivity performance, in spite of having large scale production.

In the study by Mazumdar (2013), the total factor productivity (TFP henceforward) of the pharmaceuticals are gauged by the non-parametric Malmquist productivity index by applying the meta-frontier technique that was introduced by Battese and Rao (2002). In this analysis, the productivity index is decomposed into efficiency and technical change components, as indicated in fig. 2.14. Within the thirteen years' time line for productivity changes, the author maps a technological progress in general, which notably experiences a slack and diminishes during 1995, 1996 and 1998. However the change in the efficiency level has regressed after 1998 which seemingly experiences a counter effect with respect to technical improvement. This provides the argument that the industry, as a whole, could not

extract the technological progress through efficiency gain, that the outward production frontier has offered.

Fig. 2.14: TFP change calculated from Malmquist productivity index

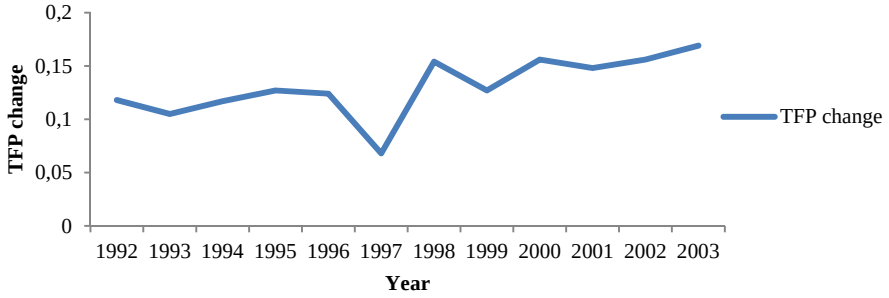


Source: Mainak Mazumdar (2013), "Performance of pharmaceutical companies in India: A critical analysis of industrial structure, firm specific resources, and emerging strategies", Physica Verlag Heidelberg: Springer Publication.

Therefore, the effect of globalization or procurement of new policy regimes have benefitted a handful of the Indian pharmaceutical companies which are technically superior and exist at the margin of the production frontier. In totality, the TFP change is observed to have a slightly upward trend, with intermittent productivity booms and busts along the time frame.

In another contemporary study by Ghose and Chakraborty (2012), the TFP change for the Indian pharmaceuticals is quantified using a translog production function. The diagram below depicts the finding from this study, where no identical resemblance is seen with fig. 2.14. Nonetheless, a prominent increasing trend in productivity is observed over the period of 1992-2003, albeit a stark depression in the year 1997. The most conceivable reasoning for this sudden drop in TFP might be due to the policy changes after the Indian accession with WTO in 1995. However the most logical apprehension is the drawback in the data set.

Fig. 2.15: TFP change calculated from a translog production function



Source: Ghose, A. and C. Chakraborty (2012).

In a nutshell, the complex and diverse nature of the pharmaceutical industry, as captured in the context of the Indian economy, has undergone several changes over the past years, owing to the alterations in the policy norms. Undoubtedly, the introduction of process and product patents have put forth this industry to involve in serious quality control of their final clinical products. Further, the ambitious steps towards a more liberalized environment have triggered higher exports, influx of more multinational pharmaceuticals and greater mergers and acquisitions. However, seconding the proposition by Corrales-Leal (2007), the dire need of this sector is to develop and promote the local pharmaceutical firms, through governmental support, such that they are able to enhance the quality of the complementary drugs and add greater value to exports.

2.5 Conclusion

On analyzing the pharmaceutical activities of different regions of the world, it is evidently clear that there exists huge bifurcations in the way this industry operates, based on the different strata of their development process. Since product patent was introduced much later in the developing nations, the magnitude of drug innovation activities does not commensurate the developing nations. The disparity in both tangible and intangible capital resources plays a key role in their varied progress

level. However imitative drug production is also a prerogative for providing low cost drugs to the vast mass of population that have modest survival amenities at their disposal.

However, the world pharmaceutical scenario is undergoing unprecedented changes with the recent surge for mergers and acquisitions, under a globalized environment. Yet, individual nations have their intrinsic pharmaceutical market dynamics, with varied resources and competition benchmark. Therefore, it is of utmost importance to develop country-tailored policy approaches for a more efficient synergy between administrative and economic policies, such that medicinal facilities can potentially reach the target population.

The most developed countries primarily resort to high level innovation for the invention of new blockbuster drugs. Albeit the colossal expenditure on research and development of new therapeutic products, there are still several challenges for the medical world to provide effective treatment for diseases like cancer, schizophrenia, Parkinson's or even diabetes. Cockburn *et al.* (2004) emphasizes that there has been no new antibiotics in the market for the last forty years. Nevertheless, in spite of apprehensions regarding the performance of the drug industry and its rate of invention, it is obviously encouraging to witness a world statistics of an average decline in mortality rate and improvement of health conditions over the past decade.

This concerted discussion on the world pharmaceutical locus provides an overall coherent knowledge of its functioning in general, which gives a more clear perspective when we concentrate our focus on a micro level analysis for the Netherlands drug sector. As has been mentioned in the first chapter, the Dutch pharmaceutical industry with its strategic location in Europe, scores fairly well in terms of trade as well as on innovative and generic productions. A substantial level of import and export activities in the Dutch pharmaceutical enterprise, with a

positive trade balance and ongoing parallel imports, bears testimony to its globally connected drug operations, with international standards and strategies.

Undoubtedly, its robust research framework, rigorous innovation practices and enviable competitive strategies allows it to excel as one of the most successful pharmaceutical industries in Europe. Therefore, we further proceed with our study on the Dutch pharmaceutical sector and aim at establishing the inherent relation between innovation, competition and firm performance, within the periphery of its pharmaceutical industry.

Chapter 3

DATA CONSTRUCTION

3.1 Data source

From this chapter forward, our study is solely based on the Netherlands' pharmaceutical industry. Therefore, the empirical analyses undertaken in the successive chapters are essentially at the micro-level, using an unbalanced panel data set for the period 1996-2006. In general, micro level data are likely to provide more appropriate level of analysis, devoid of any aggregation problem that plagues macroeconomic investigations. However the prominent setback of firm level data is the lack of enough variables relevant to a specific circumstance for each firm. In addition, the micro level variables that are available at hand are likely to possess higher relative error rate, compared to macro level data where the error rate is averaged out to a large extent (Griliches and Mairesse, 1991).

In spite of the plausible hindrances in micro-level data analysis, we have been able to compile a rich panel data by using several data sources. The respective databases employed to accrue our variables of interest 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). For our firm-level study, we obtained 673 Dutch pharmaceutical firms that were extracted from the Statistics on Financial Enterprises provided by the Central Bureau of Statistics (CBS) and the REACH database (Manufacturing of Pharmaceutical products and Pharmaceutical preparations, NACE Rev.2 Code 21).

Further, we have analyzed the ownership structure of the firms and considered only the ultimate parent firm, in order to avoid double counting of patents or other concerned variables. Since firms register patents or report R&D expenditures under different names, we used the Algemeen Bedrijven Register (ABR=general business register) data, issued yearly by Statistics Netherlands on firms' ownership structure, to find the names and the direct ownership (expressed in percentage) of all their subsidiaries, holding units and their shareholders. We manually assigned a Chamber of Commerce (KvK) code to each firm. Each KvK code was then electronically matched with a Statistics Netherlands internal code in order to obtain the entire ownership structure for each of the firms. By this selection, the number of pharmaceutical firms gets reduced to 520. The sample of firms, so obtained, are the possible (not necessarily ultimate) parent firm, which is necessarily located in the Netherlands and their data on input and output variables is available. To identify the parent firm, CBS takes into consideration a direct and indirect ownership of over 50%. It is observed that a considerable number of subsidiaries (daughters) are completely owned by a parent.

For each of the 520 firms, we looked at their entire ownership structure, including all possible subsidiaries (through an extensive manual search from the ABR), and subsequently matching them with patent counts from the patent database that has been made available by the Dutch Patent Office (Octrooicentrum, Netherlands). The accrued patent database comprises of all patents from the United States (issued by the USPTO) and Europe (issued by the EPO). The usefulness of this database is that, it eliminates any double counting of USPTO and EPO patents.

It may be noted that the patent data set from the Dutch patent office also gives us additional information about the application year, the patent owner (name of the firm), patent title, name of the inventor, publication year, and location. Although these supplementary information on the patents has not been used in our analysis, it

enriches us with a more in-depth knowledge on the patents that are applied by the pharmaceuticals in the Netherlands.[1]

Also, we have used a complementary database of the total population of European patents (issued by the EPO) for the period 2000-2006, that was partially made available from Statistics Netherlands. With this complementary database, we were able to double check the EPO patent counts for our firms with those that were obtained using the first data source. Subsequently, the citations for each patent were extracted from the PATSTAT database and matched with the individual patents. Even though, both forward and backward citations for each patent were extracted from the PATSTAT, only the forward citations have been used in our analysis for quantifying the citation-weighted patent counts.

Regarding the extraction of R&D expenditure data, it should be mentioned that a complete dataset of this variable is difficult to achieve, as many firms do not report their R&D expenditure. In spite of the difficulties, we have tried to retrieve as many R&D observations as possible, by using different databases. The first database that was consulted was the REACH database, which provided information on the R&D expenditure for the top 5000 largest Dutch firms (>100 employees) along with its detailed financial data and ownership structure. However, we find that in this 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 intangible fixed assets or operational costs. Additionally, the Dutch law that obligates firms to publish financial details

[1] In this regard, we emphasize on the fact that, the patent data considered in our analysis only takes into account the patent applications from the pharmaceutical firms that are situated within the geographical boundary of the Netherlands. Hence the patent activities of any of their subsidiaries, partner firms or their own establishments, that are stationed outside the Netherlands have not been considered in this compiled patent dataset.

(balance, profit and loss accounts, annual reports, ownership information), including their R&D expenditures, is applied to only the top 5000 firms.[2]

Therefore, we used two other 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 extracted R&D data from the Community Innovation Surveys (CIS2, CIS2.5, CIS3, CIS3.5 and CIS4) 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. From the surveys we complemented R&D data for the 520 Dutch pharmaceutical firms.

Besides, the innovation indicators, information on output, value-added, net tangible fixed capital assets, sales, depreciation, and wages, all expressed in thousands of euro, as well as the number of employees have been extracted from the production statistics database of Statistics Netherlands (CBS). By using these core variables, the competition measures and other explanatory variables are calculated.

3.2 Variables constructed

In this section, we discuss on the different variables that have been calculated or estimated for the compilation of our final database from the raw data obtained. However the analytic data descriptive of our concerned variables and their corresponding inter-linkage has not been discussed in this chapter. It is taken up further in the individual chapters that follow, such that the data analysis on respective topics can be focused upon separately. In the following subsections, we classify our data construction on the basis of innovation, competition, productivity and other control variables.

[2] A possible solution to the missing R&D investment data is to use the wages of the R&D employees as an alternative metric. In addition, the latter may have its applicability for sensitivity analysis. However, due to the unavailability of data for the wages of R&D employees this alternative measure could not be used. Also, the information on the wages of R&D employees are likely to have a similar problem of paucity of adequate observations at the firm-level. However, the wages for employees engaged in R&D is a representative variable, and does not reflect the overall investment in innovation.

3.2.1 Innovation indicators

To define quantitatively the amount of innovation by the Dutch pharmaceutical industry, R&D intensity and patents (or citation-weighted patents) are used. It can be mentioned in this context that, while the former acts as the barometer to explain research investments (or innovation input), the latter reflects the innovation performance (or innovation output) in the course of drug discovery. Since the pharmaceutical sector is heavily involved in R&D activities and has a very high propensity to patent their innovations, it is completely pertinent to choose these innovation measures for our study. However, it may be noted that although our patent data is complete with no missing values, the data on R&D expenditure has the shortcoming of many missing observations.

R&D intensity

For our empirical analyses, we have used R&D intensity of the firms instead of their R&D expenditure, for quantifying innovation investment. For calculating the R&D intensity, we use the ratio of R&D expenditure to total number of employees. Using total number of workers in the denominator instead of total sales in this case is more pertinent, as the former provides identification of changes over a limited time period compared to the latter. Therefore,

$$R\ \&\ D\ intensity = \frac{R\ \&\ D\ expenditure}{Number\ of\ \ employees} \tag{3.1}$$

In our regression models, log-transformed R&D intensity has been used. Further, we have deflated the nominal values of R&D intensity by using gross fixed capital formation price index from the EU KLEMS database (where 1995 is considered as the base year).

Patent and Citation-weighted patent counts

The innovation performance of the firms is indicated by patent counts and citation-weighted patent counts. In a way, these two variables essentially calibrate the innovation output of the research-intensive pharmaceutical sector, where the latter emphasizes on the quality of the innovation output.

The patent count is denoted by the total number of patents applied in the European Patent Office (EPO) and the US Patent Office (USPTO). The extraction of the patent data has involved an extensive manual search through the parent firms and their corresponding subsidiaries, and immaculately checked to avoid any double counting between them.

As mentioned before in section 3.1, the citation counts are correspondingly matched for each patent in our database. The citation counts (or citation-weighted patents) are the total number of forward citations of each patent in the concerned year. Hence according to this measure, the value of each patent is reflected by the number of citations it receives.

3.2.2 Competition indicators

Product market competition plays an intrinsic role in the pharmaceutical industry and is one of the most crucial explanatory variables in our research investigations. However, competition is a multidimensional concept which cannot be restricted to one specific indicator as there exists no single benchmark to measure competition. Therefore, we have used alternative indicators to capture and quantify competition in the successive chapters. In this regard, chapter 6 needs special mention, as it provides the largest spectrum of competition indicators, with a more in-depth analytic digression of these measures.

Concentration index

Traditional concentration measures like the Herfindahl-Hirschman index (HI hereafter) are important in case of mergers (Bishop and Walker, 2002), for measuring the level of product market competition. The idea behind the H-index is that, low level of concentration implies higher competition because many firms are operating and active in the market. Hence a fall in H is interpreted as an increase in competition and vice versa. In other words, more intense competition is caused due to more entry of firms in the industry. Therefore more firms entering a market tends to lower the concentration and consequently increases competition. Likewise, more concentration in an industry implies lesser competition. In the arithmetic notation, HI can be represented as

$$HI = \sum_{i=1}^{n} \left(\frac{p_{it} y_{it}}{\sum_{i=1}^{n} p_{it} y_{it}} \right)^2 \tag{3.2}$$

where $p_{it} y_{it}$ represents the sales of firm *i* at time *t*.

Therefore, HI is the summation of the squares of market share of all firms in an industry. Hence, equation (3.2) can be re-written as,

$$HI = \sum_{i=1}^{n} S_{it}^2 \tag{3.2}'$$

where, S_{it} is the market share of firm *i* at time *t*, and *n* is the number of firms.

But the HI index is a measure at the industry level as it sums over the square of market share of each firm in the industry. Hence for our firm level empirical analysis, we consider the market share of each firm as the determinant for competition. However, we have used this industry level measure of HI for the data descriptive in chapter 6, whereby HI is measured for the pharmaceutical industry

over time. Specifically, the level of competition is measured for each period of time, within our time framework of 1996-2006 for the descriptive statistics.

In our analyses, the market share of each firm 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. To sum, we calculate market share as the average turnover of a firm at time *t*, multiplied by the total number of firms chosen in the base year 1996.

$$Market\ Share_{it} = \frac{Sales_{it}}{\sum_{i=1}^{n} Sales_{it}} \times n_{1996} \tag{3.3}$$

It can be noted that, although we have approximated market share by the turnover of the firms (as in Nickel, 1996), an alternative metric for market share measure is supplemented in chapter 4. According to these alternative concentration measure, the number of employees has been used instead of sales. The reason behind the incorporation of this alternate measure is the fact that, the observations for the number of employees is substantially greater than that for sales. Hence, in view of a better estimation result, the former is applied as a sensitivity analysis. This measure of employment concentration to gauge competition and agglomeration is found in the study by Martin *et al.* (2010).

Price Cost Margin with constant returns to scale

As an alternative measure of competition, the price cost margin (PCM) is widely used in the papers of Nickell (1996), Nevo (2001) and Aghion *et al.* (2005). This competition measure, also known as the Lerner index, relies on the assumption of constant returns to scale and is an approximation of the mark-up. While the market share measures the concentration of a firm within a market, the LI measures the profitability of a firm, 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, LI falls in the range of no competition to perfect competition. It can be defined as,

$$LI_{it} \equiv \frac{\mu_{it} - 1}{\mu_{it}} \tag{3.4}$$

where $\mu_{it} \geq 1$ or $0 \leq l_{it} \leq 1$.

Thus, it basically measures the market power of the firm, as the negative inverse of demand elasticity, and ranges from 0 to 1 (with 0 denoting 'no competition' and 1 as 'perfect competition'). The classical method of computing the PCM is by calculating the difference between output price and marginal costs, weighted by the output price. It can be defined as,

$$PCM_{it} = \frac{\sum_{i=1}^{n} (p_{it} y_{it} - c_{it} x_{it})}{\sum_{i=1}^{n} p_{it} y_{it}} \tag{3.5}$$

where $p_{it} y_{it}$ is the sales and $c_{it} x_{it}$ is the marginal cost of firm *i* in year *t*. A greater difference between price and marginal cost indicates a larger distance between price and competitive price of each firm, which articulates their market power. However, owing to the lack of data on price and marginal cost at the firm level, we compute the index following Aghion *et al.*, (2005). According to Aghion *et al.* (2005), LI can be directly retrieved from balance sheet data by a simple ratio between profits and sales of each firm defined as,

$$LI_{it} = \frac{Value\ Added_{it} - Wages_{it}}{Sales_{it}} \tag{3.6}$$

Therefore, the LI used in our analyses is obtained as the difference between value-added and total wages divided by total sales.

Price Cost margin with adjustment for economies of scale

Although the traditional measure of PCM does not account for economies of scale, it is important to clarify the existence of scale and scope economies. The presence of non-exhausted scale economies is an indication that the ability of the firms to reduce costs has not been dealt with. Hence, this alternative technique to measure PCM, which is adjusted for economies of scale, can be seen as an indirect indicator of the lack in competition (Bikker and Van Leuvensteijn, 2008). In this case, the mark-up is measured as the scale elasticity times the ratio of outputs over multiple inputs. The scale elasticity is derived from a translog production function.

In line with previous research (for example, Amoroso *et al.*, 2011; Basu and Fernald, 2002; De Loecker and Warzynski, 2012 and Vancauteren, 2012) we have computed our variable of interest, whereby a production or dual cost function is used to obtain the same. This calculated mark-up is an input-dependent, time-varying elasticity of scale which is multiplied by an output-input share at the firm level. Therefore,

$$Mark\ up_{it} \equiv \frac{ValueAdded_{it}}{(wages_{it} + capital\ user\ costs_{it})}\theta_{it} \tag{3.7}$$

where θ_{it} is the elasticity of scale. While the output elasticities are required to be estimated, the ratio of output over input is directly calculated using the firm-level production data. The firm-level scale elasticities are estimated using semi-parametric techniques. Relying on the theoretical models of innovation (for example, Melitz and Ottaviano, 2008) with heterogeneous producers, our assumption of a time-varying and firm-specific mark-up remains consistent. In addition, the time-varying characteristics of mark-up is closely associated with the cyclical pattern changes in market share and the price elasticity of demand. For a

more detailed information, the survey by Jaimovich and Floetotto (2008) may be consulted.

For estimating the output elasticities, we consider a production function with a translog approximation (as in, De Loecker and Warzynski, 2010). A more flexible structure on the data is imposed with the assumption of a translog function. Therefore, with this assumption of a flexible functional form, the pattern of production for heterogenous firms with prominently varying sizes can be revealed. In other words, the heterogeneity in firm size and additional firm characteristics that explains variations in the input-output structure is taken into account by the estimates. This is due to the fact that, the first derivative of the translog production function differs by firm-year observation. The empirical notation for translog production function can be presented as follows,

$$\ln(y_{it}) = \beta_0 + \beta_k \ln(k_{it}) + \beta_1 \ln(l_{it}) + \beta_{kk} \ln(k_{it})^2 + \beta_{ll} \ln(l_{it})^2 + \beta_{kl} \ln(l_{it}) \ln(k_{it}) + \varepsilon_{it} \qquad (3.8)$$

where y_{it} denotes the value-added of firm *i* in year *t*, l_{it} is employment, k_{it} is capital, and ε_{it} is the error term. In case of a translog production function, the elasticities of scale θ_{it} is indicated by the sum of the output elasticities: $\theta_{ikt} + \theta_{ilt} = \theta_{it}$, which can be represented as,

$$\theta_{ikt} = \frac{\partial \ln(y_{it})}{\partial \ln(k_{it})} = \beta_k + 2\beta_{kk} \ln(k_{it}) + \beta_{kl} \ln(l_{it}) \qquad (3.9.1)$$

and, $$\theta_{ilt} = \frac{\partial \ln(y_{it})}{\partial \ln(l_{it})} = \beta_l + 2\beta_{ll} \ln(l_{it}) + \beta_{kl} \ln(k_{it}) \qquad (3.9.2)$$

The firm-level mark-up is thus generated by using the firm heterogenous estimated elasticities of scale.

However, the estimation of the production function raises questions about simultaneity bias. Since firms choose input and output simultaneously, unobserved

firm-level characteristics may cause the error term to be correlated with the input factors of the production function. As a result, this violates the orthogonality of the error term, so standard Ordinary Least Squares (OLS) techniques will be biased and inconsistent. Hence, we consider the semi-parametric methods developed by Levinsohn and Petrin (2003) to control for simultaneity. This method solves the simultaneity problem by using intermediate inputs to proxy unobserved productivity shocks (assumed to follow a first-order Markov process) that are observed by the firm.

Profit elasticity

The fourth and last measure that we take into account is the profit elasticity, as introduced by Boone (2000). The profit elasticity or the Boone indicator is the estimation of the impact of efficiency on the performance of the firms. According to this indicator, competition increases the performance of efficient firms and impairs the performance of the inefficient firms. This measure is calculated from the relative profit differences (RPD), as the percentage fall in profits caused by a percentage increase in the marginal cost (refer Boone, 2008). Based on a general consensus, an increase in cost reduces a firm's profit. But, in case of a more competitive market, the same percentage level of increase in cost will lead to a bigger fall in profits.

The Boone indicator, estimated by using firm-level data, is measured over time and across industries. In other words, it measures the degree of competition at the industry level. For our data of single sector, we calculate the Boone indicator for each period of time, from 1996 to 2006. The methodology of Boone *et al*. (2007) suggests the following two-way fixed-effect regression:

$$\ln(\pi_{it}) = \alpha_i + \alpha_t - \beta_t \ln(c_{it}) + \varepsilon_{it} \tag{3.10}$$

where α_i is the firm fixed effect, α_t is the time fixed effect, π_{it} denotes the profit of firm *i* in year *t*, c_{it} denotes the marginal cost, ε_{it} is the error term and β_t is the profit elasticity (or the Boone indicator).

Boone (2008) focuses mainly on the change in β_t over time within a given industry. Also, since we calculate the profit elasticity of a single industry for each time period *t*, we omit the time fixed effect α_t in our analysis. Hence the profit elasticity in our model gets reduced to,

$$\ln(\pi_{it}) = \alpha_i - \beta_t \ln(c_{it}) + \varepsilon_{it} \tag{3.11}$$

The profit elasticity measure is based on firm-level panel data, and hence could be estimated using our dataset. The marginal cost is approximated by the ratio of variable cost and operating revenue (or market share). The variable cost is obtained as the sum of material cost and employees' salaries. It can also be approximated by the cost of intermediate inputs. Profit (π_{it}) is defined as the difference between a firm's revenue and its variable cost. We have used a time trend and an interaction term of marginal cost and year dummies as the independent variables. The Boone indicator explains that, more negative the value of β_t, the greater is the level of competition. Therefore, an absolute value of this indicator also indicates a positive relation with competition. Hence we use the absolute value of the estimated profit elasticities to quantify competition.

Finally, as we wrap up with the construction of the competition measures, we would like to bring to notice that, the last two competition measures discussed have been applied only in case of chapter 6 and not in the rest of the chapters.

3.2.3 Output and productivity

The total production of the Dutch pharmaceutical industry and its corresponding level of productivity reflect the overall performance of the firms. These variables

are exclusively used in chapter 7, which focuses on the characteristic driving force that enhances the productivity and thereby the performance of the pharmaceuticals.

Output

Regarding the production output measure, we use the log of value-added. The value-added variable is extracted from the Statistics Netherlands data source, similar to the other control variables. Further, we have used the value-added price deflator from OECD STAN database (considering 2005 as the base year) to deflate its nominal value.

Total factor productivity

We estimate the Cobb-Douglas production function using the semi-parametric Levinsohn-Petrin (2003) technique in the following form.

$$\ln(Y_{it}) = \eta_0 + \eta_1 \ln(L_{it}) + \eta_2 \ln(K_{it}) + \lambda_{it} + \omega_{it} \qquad (3.12)$$

where Y, L and K denote the value-added, labour and capital of firm *i* in year *t* in their logarithmic forms. It is to be mentioned that L represents the number of employees, as data on the labour working hours of each firm is not available. In this methodology the error term can be decomposed into two components, where λ denotes the transmitted productivity component and ω is an unobserved error term which is uncorrelated with the corresponding inputs.

In this estimation technique in Stata, the original dataset is estimated first, which is followed by the storing of sample moments by global macros. Finally a difference between their values and the bootstrapped sample moments is obtained, while the objective function of the bootstrapped sample is minimized. In our estimation we have specified 250 bootstrap replications. The proxy variable that we have used as the material input is the intermediate consumption. From the estimation results, the predicted value of TFP is obtained, which is enumerated in chapter 7.

Labour productivity

In our analysis, we have incorporated labour productivity, which is expressed as the total output on a per worker basis. The labour productivity is a widely used productivity measure (as in Crepon *et al.*, 1998 and, Raymond *et al.*, 2013 among others). Although papers like Griffith *et al.* (2006), Janz *et al.* (2003) and many others use log of sales per employee as the measure of output; there are also many evidences of other prominent research works like Crepon *et al.* (1998), Benavante (2006), Loof and Heshmati (2006) among others who have proxied output using log of value-added per employee. For our analysis, we have considered the latter measure and thus, the demand shifting effect of innovation is not captured in our study. Therefore, the labour productivity in our study can be defined as,

$$Labour\ productivity = \frac{Value\ Added}{Number\ of\ employees} \tag{3.13}$$

3.2.4 Dummy variables

The dummy variables that have been incorporated in our study includes the patent dummies and the entry-exit dummies. While the former accounts for the dynamics in the innovation process, the latter gauges the turbulence level in the market.

Patent dummy

In order to identify the drivers of persistence in our model we use a lagged patent dummy, which is 1 if any of the previous years (or year) have patents for a particular firm, and zero otherwise. This assumption is based on the fact that, a firm invests in innovation in the period *t* if it has innovation output (as measured by the number of patents) in its past years, in the sample data period of 1996-2006.

Other than the lagged patent dummy, lagged patents have also been used as an alternative measure to establish if the persistence in the innovation process is maintained. It can be noted, in this context, that both firm size and entry-exit are

supposed to play a major role in the innovation persistence of firms. Various empirical studies like Geroski *et al.* (1997) and, Duguet and Monjon (2004) stresses on the fact that, innovation persistence is influenced by size of firms. In addition, competitive turbulence, as defined by the entry-exit or survival of the firms, is significant for dynamics in innovation (e.g. Antonelli *et al.*, 2010 and; Malerba and Orsenigo, 1996).

Entry Exit dummy

The entry and exit dummies reflect the turbulence created by an increase in the number of incumbents or exit in the existing firms, which in a way, affect the competitive conduct of a firm in the market.

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.

3.2.5 Other control variables

Besides the competition indicators and the dummy variables, we have used additional explanatory variables in our analyses that characterize the innovation or productivity of the firms. From our extracted raw data, the respective control variables are calculated, which includes the size of firms, capital intensity, firm age, entry-exit dummies and patent dummies. While we briefly describe the first three mentioned variables here, the dummy variables are already discussed in the previous subsection.

Firm size

The logarithmic value of the number of employees has been used as a proxy for firm size in our analysis. The data on employment are extracted from the general business register (in Dutch, “Algemeen Bedrijven Register” or abbreviated as

"ABR"). According to Mairesse and Mohnen (2002), firm size measured by the number of employees reflects access to better financing.

Capital intensity

Following Konings *et al.* (2001), the capital stock was approximated by net fixed tangible assets valued at book level, i.e., tangible fixed assets at historic costs minus depreciation. These data for calculating capital stock has been extracted from the database "Statistics of Finance of Enterprises", accrued from Statistics Netherlands. The capital intensity is calculated as the ratio of capital stock and the number of employees. Hence, it is measured on per worker basis.

Similar to R&D intensity, the nominal value of capital intensity is converted to its real terms by using gross fixed capital formation price index from the EU KLEMS database (where the base year is 1995).

Firm Age

The age of firms is measured as the difference between the entry and the exit years for each firm in the successive years. In our estimation analyses, its logarithmic value is used, similar to most of the other variables, such that a better conformity to a normal distribution is maintained. Bearing in mind that the life span of firms, along with their employment and capital base, play an important role in determining innovation and productivity growth, these variables can be slated as important determinants in our empirical estimations.

Regulation indicator

In order to control for the plausible endogeneity on the causal relationship between competition and innovation, the regulation impact indicator (REGIMPACT) is incorporated in the empirical analysis of chapter 6. The REGIMPACT data is procured from the OECD database, and is the only data file that provides a wide coverage of countries and sectors. Additionally, it also spans across several years

which includes our concerned time frame of 1996-2006. Developed by Conway and Nicolletti (2006), this regulation indicator gauges the repercussion effect of regulation in the non-manufacturing industries over all industries. The values for this indicator has been normalized within the range of 0 to 1, ranging from the least restrictive to the most restrictive industrial regime. For a detailed elucidation of this indicator, Braila *et al.* (2010) may be consulted.

To sum up, this chapter specifically dealt with the data that has been used in our research analyses in the next chapters. Since most of the variables used in the different chapters and their data source bear congruence, we aimed at confining them in one single chapter to prevent repetition. However, as has been mentioned before, we do not incorporate the data descriptive in this chapter and elucidate them with each concerned topic, such that the essence of the interrelationships of the different variables is relevant with the contents of the specific topics.

Chapter 4

DETERMINANTS OF R&D INVESTMENT

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 chapter empirically investigates on how a firm's competitiveness, size, capital intensity and age affect 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, that 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 are not available. The maximum likelihood (ML) approach following Wooldridge (2005) is applied for handling the individual effects. In order to identify whether a 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 level determinants, Competition, Lagged patent, Tobit II

4.1 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 chapter 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.

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 chapter 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 chapter provides a quick review of the literature dealing with R&D and its determinants. Section 4.3 describes the data used for the analysis, and section 4.4 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 4.5. Finally, section 4.6 concludes.

4.2 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. Although the costs on R&D expenditure have surged over the last decade (DiMasi *et al*., 2003), the pharmaceutical industry undoubtedly remains as one of the most R&D intensive sectors (OECD, 2003). There exists a significant amount of research work on the valuation and drivers of R&D expenditure. Here we review

some of these past literatures that deal with the various firm characteristics which affect their R&D investment.

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 chapter 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 on 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, it has 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

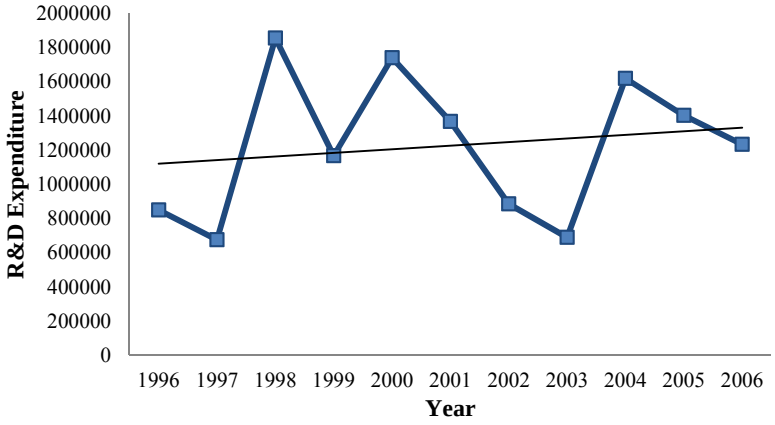
asserted by Taylor and Silberson (1973), Levin *et al*. (1987) and Cohen *et al.* (1997). Hence, firms are prone to invest in R&D and subsequently patent their innovations, which in turn augments R&D investment in order to transform the patent into a more commercial innovation. This propensity is cyclical, 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.

4.3 Data descriptive

As previously elucidated in chapter 3, we employ a purpose built dataset based on a panel of firms located in the Netherlands, with annual data from 1996 through 2006, that includes financial determinants of R&D expenditures. The inclusion of the plausible parent firm establishes the groundwork for preventing repetition of innovation observations. In particular, firms may register patents or report R&D under different names and this may result in double counting. Hence, the selection of the ultimate parent firms are made, where the patent counts of their subsidiaries are incorporated as the observations of the individual parent firm.

From the compiled data, the trend in R&D expenditure over the period 1996-2006 is represented in fig. 4.1. 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.

Fig. 4.1: Trend in R&D expenditure in the Netherlands' pharmaceuticals



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 have 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.[3]

The various determinants that influence the level of R&D investment and have 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 patent dummy variable (p_k).

Table 4.1 captures the statistical summary of the variables used in this chapter. It may be noted that, we have incorporated the R&D intensity of the firms instead of

[3] Based on the diagrammatic representation of fig. 4.1, it is observed that the R&D expenditure trend exhibits crests mostly for the even years and troughs for the odd years. This might plausibly occur due to a lack of data for the R&D surveys (since R&D surveys account for data during the odd years and CIS provides the same for the even years). For this reason, a comparison with the OECD STAN database (ISIC Rev. 3) has been undertaken for the Dutch pharmaceutical industry. It is observed that the overall trend in the R&D expenditure over the concerned time frame is analogous to our compiled R&D dataset.

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.

Table 4.1: 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 have 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 atleast 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 4.2 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 resemble those pharmaceuticals that has missing or (/and) zero R&D data for the entire period under consideration.

Table 4.2: 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 4.2 is merely a representation of the data at hand and is not conclusive.

On competition and innovation

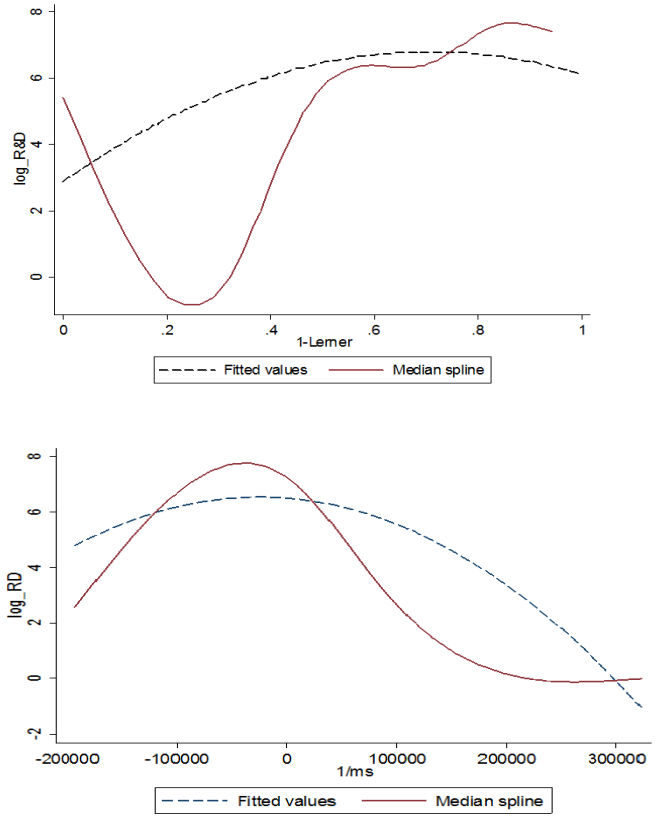
Competition in the pharmaceutical industry prevails not only on the introduction of newly invented drugs into the market, but also on the imitative drug therapies. Within the periphery of a given therapeutic class, the family of medicines goes through a well-delineated life cycle. Amongst the inventive 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 more wider perspective, where the geographic and product boundaries of a market for the operating firm do not bear any significance. The pros and cons of the different competition measures, along with their respective variable construction mechanisms are provided in chapter 3.

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 figure 4.2. 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 may be noted in this context that, the competition measures are transformed in order to reflect the direct relation between competition and innovation (as in Aghion *et al.*,

2005). It is observed that, in both the cases, a non-linear inverted U-shaped curve is achieved.

Fig. 4.2: 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.

4.4 Empirical methodology

In the section on data analysis, it has been observed that 86% of data on R&D intensity is missing with only 191 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 firms 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:

$$\begin{aligned} R\,\&\,D_{it} &= R\,\&\,D_{it}^{*} \text{ if } R\,\&\,D_{it}^{*} = \alpha_{1i} + \beta_1' X_{1it} + \varepsilon_{1it} \geq 0, \\ &= 0 \text{ otherwise} \end{aligned} \tag{4.1}$$

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_{1it} 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

$$\begin{aligned} z_{it} &= 1 \; if \;\; z_{it}^{*} = a_{2i} + \beta_2^{'} X_{2it} + \varepsilon_{2it} \geq 0, \\ &= 0 \; otherwise, \end{aligned} \tag{4.2}$$

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,

$$a_{1i} = \alpha_{10} + \delta_{10} R\ \&\ D^{*}_{i0} + \delta_1' \overline{X}_{i1} + \xi_{1i}, \tag{4.3A}$$

$$a_{2i} = \alpha_{20} + \delta_{20} z_{i0} + \delta_2' \overline{X}_{i2} + \xi_{2i} \tag{4.3B}$$

where α_{10} and α_{20} are constants, $\overline{X}_{i1}, \overline{X}_{i2}$ are the vectors which includes the time averages of the explanatory variables $(e_{it}, me_{it}\ or\ ms_{it}, l_{it}, k_{it}, a_{it})'$, $R\ \&\ D^{*}_{i0}$ 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 $\xi_{1i} \mid X_{i1} \sim N(0, \sigma^2_{\xi 1})$ and $\xi_{2i} \mid X_{i2} \sim N(0, \sigma^2_{\xi 2})$. However in this case, ξ_{1i} and ξ_{2i} are assumed to be equal to zero.

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 (4.3A) and (4.3B) can be re-written as,

$$a_{1i} = \alpha_{10} + \delta_{10} R \,\&\, D^*_{i0} + \delta'_1 \overline{X}'_{i1} + \xi_{1i}, \qquad (4.3A)'$$

$$a_{2i} = \alpha_{20} + \delta_{20} z_{i0} + \delta'_2 \overline{X}'_{i2} + \xi_{2i} \qquad (4.3B)'$$

where, $\overline{X}'_i = \frac{1}{T-1} \sum_{t=2}^{T} X_{it}$

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 (1995) and Klepper (1996) provide theoretical insights into the nature of the dynamics of innovation process. As mentioned in section 4.3, 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. It is noteworthy that we have considered lagged patent dummy over lagged patents due to the skewness in the patent data, such that the patent error is minimized.

4.5 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.

4.5.1 Results of the static model

Table 4.3 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 variables 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 2, 3 and 4, 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 two regression results, market share using sales is solely used as the competition measure. However in model 3 and 4, both market share and Lerner index have been used as competition measures. But in column 4, 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 have 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 4.3.

From the results obtained in table 4.3, it is seen that the coefficient of firm size has a systematically negative and significant effect on R&D intensity.[4] 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, Crepon *et al*., 1988, Klomp and Van Leeuwen, 2006 and Benavente, 2006). Although there exists an extensive empirical work on innovativeness and firm size, no conclusive inference has 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

[4] Since employment (e_{it}) has been used to calculate the log of R&D intensity $\left(\log\frac{R\&D_{it}}{e_{it}}\right)$, there can be possibility of biasness in the regression results, when the former is used as an independent variable and the latter as the dependent variable. However on running a regression for these variables, it was tested that the coefficient on the $\log(e_{it})$ is statistically less than 1. This implies that the impact of $\log(e_{it})$ on $\log(R\&D_{it})$ is negative. Therefore, our empirical result is justifiable.

(2006) has asserted that the pharmaceutical industry in Germany, France and the 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 may 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 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 Crepon *et al.* (1998), the signs for 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 Crepon *et al.* (1998) study.

Concerning the age of the Dutch pharmaceutical firms, a positive and significant effect on R&D intensity is obtained 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 2. But it loses its significance when Lerner index is incorporated as an additional regressor in the last two regressions (column 3 and 4).

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 4.3. 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 3 and 4). The Lerner index exhibits a significantly positive effect on R&D intensity in the regression result of column 3. 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 4.3 suggest an inverted U relationship between them, the finding was only suggestive, as no consideration was made for the non-reporting R&D firms. Nonetheless, a more detailed analysis on competition and innovation activities of the Dutch pharmaceuticals is carried out in chapter 6.

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 4.3: 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

4.5.2 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.* (2010) have investigated the dynamic relation between patent and R&D in reference to the Netherlands. Both studies confirm a persistence in innovation over time. In the work of Raymond *et al*. (2010), 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[5] in the innovation process. Hence

[5] It may be noted in this context that the term 'dynamics' coined for this particular analysis denotes the cyclicality in the behavior of past innovation performance affecting the present innovation investment. Hence the persistence in innovation is perceived between two different measures of innovation, and not on the plausible dynamic effect of the same innovation measure at different periods of time. Hence, the impact of past patenting on the current R&D investment justifies whether continuity in the overall innovation process exists. This phenomenon has been referred as the dynamics in innovation for this empirical investigation.

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 4.4. 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 4.4. 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 effects remain the same on R&D intensity and they do not exhibit any prominent divergence from the results of the static model. Summarizing the effects 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 4.4.

Table 4.4: 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.

4.5.3 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 4.5. The first regression in table 4.5 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 (also for our subsequent chapters).

Table 4.5: Tobit I estimations

Dependent variable	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]
Time Dummies	YES	YES
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 4.5 yields similar results as the Tobit II estimations in table 4.3 and 4.4, 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, to test for the presence of endogeneity between product market competition and R&D intensity, we use the structural model approach (as in Cameron and Trivedi, 2009). The endogeneity test has been described in details in the appendix. It is observed from the analysis that endogeneity is rejected at 5% level. Hence we can consider our estimates as robust.

4.6 Conclusion

The prime objective of this chapter is to empirically investigate and elucidate on 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, due to 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 have 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. Our empirical findings put forth a number of policy implications that can ratify and accentuate the R&D performance of this particular Dutch sector, which we discuss in chapter 8.

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.

A.4. APPENDIX

A.4.1: Testing endogeneity between competition and R&D intensity

In our analysis, the regressors are assumed to be exogenous variables. But product market competition, as depicted by market share and Lerner index, can be endogenous with R&D intensity. Therefore we use the structural model approach (Cameron and Trivedi, 2009) to check if there exists a potential reverse causality between competition and R&D intensity. Other possible reasons for endogeneity may be due to omitted variable bias or data measurement error.

We consider the wages of employees and depreciation as the instrumental variables (IV), which is excluded from our basic type-2 Tobit estimations. The condition for robust identification is that, there must exist atleast one valid instrumental variable. We use a two-step estimator in STATA (as in Cameron and Trivedi, 2009), since maximum likelihood method is difficult to implement. The first stage linear regression with instrumental variable (IV) using panel data is represented in table A.4.1.

Table A.4.1: First stage linear regression

Dependent Variable	Wages	Depreciation	Patent Dummy	Log (Employees)	Log (Age)	Log (Capital Intensity)	Entry	Exit
Market share	5.43e-06*** [2.66e-07]	2.59e-06*** [9.01e-07]	-0.059* [0.035]	0.045*** [0.009]	-0.013 [0.009]	0.031*** [0.006]	-0.087 0.064	-0.004 [0.057]
Log(LI)	4.95e-07* [2.66e-07]	1.88e-06** [8.79e-07]	0.278* [0.144]	-0.069*** [0.017]	-0.003 [0.029]	0.043** [0.0175]	-0.029 [0.101]	-0.219*** [0.074]

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

Two separate linear regressions were performed using market share and log(LI) as the dependent variables. From the panel regression results, so obtained, it is found that the coefficients of the IVs are significant in both the cases, with a positive sign (although the values of the coefficients are extremely low).

In the second stage, we fit the dynamic tobit-2 model with random effects, to the regressors and the predicted residual that we obtain in the first stage. In the first

regression, where we have used market share as the proxy for competition, the z-statistics of the predicted residual for the selection equation and the outcome equation has a p-value of 0.079 and 0.775 respectively. In case of the second regression model where log of Lerner index is used as the competition measure, the z-statistics of the predicted residual has a p value of 0.662 for the selection equation and 0.050 for the outcome equation.

Table A.4.2: Second stage Tobit II results with residuals

Dependent variable	Probit (R&D=0/1)	P>|z|	Log of R&D per employee	P>|z|	Probit (R&D=0/1)	P>|z|	Log of R&D per employee	P>|z|
Log(Employees)	0.435	*0.000*	-0.592	*0.000*	0.539	*0.000*	0.002	*0.982*
	[0.053]		[0.104]		[0.046]		[0.082]	
Log(Age)	-0.214	*0.004*	-0.166	*0.126*	-0.160	*0.040*	-0.116	*0.337*
	[0.074]		[0.108]		[0.078]		[0.121]	
Log(Capital Intensity)	0.050	*0.103*	0.190	*0.014*	0.049		0.370	*0.000*
	[0.031]		[0.077]		[0.043]		[0.048]	
Market Share Using Sales	-0.045	*0.644*	0.535	*0.000*				
	[0.098]		[0.118]					
Log(Lerner Index)					0.207	*0.065*	0.221	*0.190*
					[0.112]		[0.169]	
Entry	0.101	*0.555*			0.046	*0.815*		
	[0.170]				[0.196]			
Exit	-0.504	*0.001*			-0.471	*0.002*		
	[0.148]				[0.155]			
Patent Dummy	0.220	*0.230*	0.227	*0.213*	-0.379	*0.059*	0.261	*0.174*
	[0.183]		[0.182]		[0.201]		[0.192]	
Predicted Residual	0.393	***0.079***	-0.032	***0.775***	0.034	***0.662***	-0.242	***0.050***
	[0.224]		[0.111]		[0.077]		[0.121]	
Intercept	-2.403	*0.000*	3.938	*0.000*	-2.067	*0.000*	0.216	*0.755*
	[0.258]		[0.719]		[0.289]		[0.692]	
Initial[Log(R&D/Employee)]			0.424	*0.000*			0.364	*0.000*
			[0.050]				[0.040]	
..Initial[R&DProbability]	0.912	*0.000*			0.884	*0.000*		
	[0.126]				[0.136]			
..Random Effects	YES				YES			

The z-statistics for the coefficient of the predicted residual provides for the robust Wald test of the null hypothesis of exogeneity, where the null hypothesis $H_0 : \rho = 0$. It is observed that the z-statistics has a p-value greater than (or equal to) 0.05 in both the regression analyses, leading to the non-rejection of the null hypothesis at 5% significance level. Since ρ=0, the competition measures can be treated as exogenous. Therefore in our analysis, endogeneity is rejected at 5% level, and we may consider our empirical estimations to be robust.

However we observe that the outcome equation in the second regression analysis (when Lerner index is used as the competition measure) is marginally exogenous with a p-value of 0.05. Also, the above endogeneity tests undertaken is a two-sided test and one-sided tests are not attempted. Hence, there exists a scope to further improvise our empirical research by using lagged values of competition measures that can further mitigate any problem of endogeneity that may exist.

Chapter 5

THE EFFECT OF R&D INVESTMENT ON PATENTING

Abstract

This chapter analyzes the effect of R&D intensity and other economic determinants on the innovation output of the Dutch pharmaceutical industry. A dynamic count data model is developed and applied, dealing with the initial value problem following Wooldridge (2005), in the context of a panel data framework. Our model incorporates fitted R&D intensity (obtained from Tobit II estimation), and other firm characteristics as explanatory variables. Concerning the dependent variable, we have considered both patent counts and citation counts, also for EPO and USPTO patents individually, as the innovation output indicator. The reason for using the alternative dependent variables is due to the fact that, although both patent counts and citation-weighted patents can be viewed as indicators of technological impact and information flow, the latter reflects the 'quality' of the patents. However, analogous empirical findings were obtained for both patents and citation-weighted patents, which directed towards similar inferences. The estimated results provide a consistent evidence of R&D intensity to be a major determinant for generating new patents. In addition, our results further suggest that firms with a higher employment base, have a greater propensity to persistently innovate, under prominent barriers to entry and exit.

Keywords: Patents, Citation-weighted patents, Fitted R&D intensity, Zero inflated negative binomial

5.1 Introduction

Although, both R&D and patents are used as indicators of technological capacity of firms, it has often been recognized that these measures capture different aspects of the innovation process. While R&D expenditure can be viewed as a measure of the resources devoted to innovation, patents reflect the results of the innovation processes. Although patent is a widely used proxy for innovation output, other indicators for quantifying innovation output include innovative sales, innovation counts or product information. However we have used patent data in our analysis as it is more appropriate for our study, which is based on the innovation intensive pharmaceutical industry. Moreover, the quality and availability of the data on R&D and patents has improved and refined in the recent years. Computerization of patent offices and regular surveys of R&D activities allow researchers to perform detailed analysis of patent-R&D relations. Therefore, we attempt to analytically and quantitatively clarify the contemporaneous relation between patenting and R&D expenditures at the firm level using a panel data framework.

Patent has always been recognized as a rich and potentially fruitful source of data for the study of innovation and technical change. Patent data is particularly pertinent for studying pharmaceuticals because drugs are one category of innovation where the incentive-giving role of patents works best, given the considerable investments they require. The pharmaceutical industry is intensively research oriented, performing various innovation activities consistently. Patenting in the pharmaceutical industry plays a pivotal role to accrue the huge cost involved in R&D activities of discovering, developing and gaining regulatory approvals for new drugs. Levin *et al.* (1987) showed that a patent is the most effective method to appropriate returns in industries with chemical base, such as pharmaceuticals. This in turn enables them to recover the R&D investment. Patent protection or equivalent barriers allows the innovative firms to guard their innovations from the

imitators, who could free-ride and duplicate the invented drug using a very small fraction of the innovator's cost.

But innovations vary extensively in their technological and economic importance and significance. Moreover the distribution of such "importance" or "value" is highly skewed. Allison *et al.* (2009) posits that, less than 10 percent of patents are worthy of the cost involved in securing them. Hence, simple patent counts as the measure of innovation output has often been criticized for not upholding the value of the innovations that are patented. Schankerman and Pakes (1986) and Pakes and Simpson (1991) clearly revealed the drawback of simple patent count data, and used patent renewal data instead. Since patents with greater value are renewed for more number of times, patent renewal data can act as a weighted patent, precisely indicating the value of patents.

However, patents weighted by forward citations is a more widely used indicator of the "importance" of individual patents in order to capture the enormous heterogeneity in their respective values (as in Hall *et al*., 2000; and Jaffe *et al*., 2000). Earlier studies (like Harhoff *et al*., 1999; Lanjouw and Schankerman, 2001; and Trajtenberg, 1990) reveal that the forward citations are positively correlated with the monetary value of the patent. This clearly reveals the fact that forward citations act as a barometer for determining the worth of the patents in terms of its "originality" and "generality". A "generality" score determines whether the patents have a widespread impact, influencing subsequent innovations in a variety of fields. On the other hand, "originality" reflects how authentic or novel the patented innovation is. If a patent cites previous patents that belong to a narrow set of technologies, then the originality score will be low (Trajtenberg *et al*., 1997). In additional, patent citations can also be used for studying spillovers and knowledge flows. Based on the study by Hall *et al.* (2005), the pharmaceutical sector has distinct characteristics of discrete product technologies where patents perform the traditional role of exclusion, and citations measure their value on an individual

basis. In our analysis, we focus and deal with the citation-weighted patents, in addition with simple patent counts, as the indicators to measure the output of innovation.

For analyzing the relationship between patents and R&D expenditures, statistical models of counts (non-negative integers) in the context of panel data framework has been implemented. The model used is an application and generalization of the Poisson distribution to allow for independent variables, persistent individuals (fixed or random effects) and noise or randomness in the Poisson probability function. In addition, our panel data allows us to analyze the relation between past innovation activities to current innovation activities. Consequently, this helps us to comprehend if there exists a persistence in innovation at the firm level. In general, micro level studies that look at the dynamics of patent-R&D relationship show evidence of the persistence in innovation (for example, Van Leeuwen, 2002).

As posited by Peters (2007), a couple of reasons can be cited for firms to innovate persistently. Firstly, the dynamics of a firm's innovation behavior is an essential assumption for endogenous growth models, that rationalize the idea of intertemporal complementarity in innovation. Secondly, the so-called "success breeds success hypothesis" assumes that firms become more prosperous through successful innovation, due to broader technological opportunities. Finally, some theoretical explanations consider the sunk costs in R&D investments as an important source of persistence since they create barriers to entry, causing engagements to continue innovation. It is observed that the pharmaceutical sector, which is primarily based on knowledge, is more susceptible to technological accumulation and pioneering in the persistence of innovation process, compared to other industries. Also, the innovative pharmaceutical firms have the tendency to patent their inventions steadily, even by marginally changing their past innovations, so that they can ward off unwanted competitors or imitators.

Therefore, apart from identifying the relation between R&D expenditure and innovation output in the Dutch pharmaceutical sector, the contribution of the study is two-fold. Firstly, our panel data allows us to analyze the dynamics of the innovation process. In other words, it enables us to find whether past innovation activities affect current innovation activities. Secondly, our study pioneers in providing empirical investigation using both patents and citation-weighted patents, at a detailed and comprehensive level. Also, our intensive dataset provides us with information on whether the patents are applied at the US or European patent offices. This allows us to draw inferences on national and international patenting activities.

The remainder of chapter 5 proceeds as follows. In section 5.2 the relevant literatures concerning R&D-patent relationship, along with their corresponding determinants and dynamics, has been reviewed. Section 5.3 describes the data used in our model and section 5.4 offers a brief overview of the empirical methodology. The empirical findings of different versions of the model explaining innovation activities are then discussed and contrasted in section 5.5. Finally, section 5.6 provides with the concluding remarks

5.2 Literature review

Both patent information and R&D figures are commonly used economic indicators to analyze technical change. It is a general consensus that the estimated elasticity of patents with relation to R&D is positive and significant, but the amplitude of the elasticity differs based on the level of analysis and the econometric specification used. At the firm level, earlier studies include Pakes and Griliches (1980), Hausman *et al.* (1984) and Cincera (1997) among others. With the increase in the availability and quality of the innovation data, these indicators are widely used in recent literatures to capture different aspects of the innovation process. Recent studies at the firm level include Licht and Zoz (2000), Blundell *et al.* (2002) and

Czarnitzki *et al.* (2009). Special emphasis on the innovation in the pharmaceutical sector is found in the studies by Nesta and Saviotti (2005) and Duflos (2006).

The econometric techniques for measuring the relation between innovation expenditures and innovation output were first developed by Griliches (1979) and Crepon *et al*. (1998). In the work of Griliches (1979), innovation performance relation was divided into three equations, where the second equation, that is, the knowledge production function, relates innovation inputs to innovation output. According to Klomp and Van Leeuwen (1999), firms that perform R&D on a continuous basis shows a significantly higher innovation output. Lööf and Heshmati (2000), while focusing on the relation between expenditures on innovation input and its effect on innovation output, found that the most important source of knowledge comes from within the firm, whereas competitors are the most important external source of knowledge. Mairesse and Mohnen (2005) found that innovation output is generally more sensitive to R&D in low-tech sectors than in high-tech sectors.

Since patent data consists of discrete and non-negative integers, most research on the effect of R&D intensity on patents is based on the count data modeling framework. Count data models are applied to the patent-R&D relationship by many earlier researchers, which include Bound *et al*. (1984) and Hausman *et al.* (1984). Recent empirical studies include Griffith *et al.* (2006), Mohnen *et al*. (2006), Hall *et al*. (2000) and Raymond *et al.* (2009). As stressed by Mairesse and Mohnen (2010), the R&D-innovation framework has been extended in various directions as the use of innovation expenditures rather than the use of R&D expenditures (Janz *et al*., 2004, and Lööf and Heshmati, 2006), by including a demand shifting effect of innovation output (Klomp and van Leeuwen, 2006), making a distinction between new-to-firm versus new-to-market innovations (Duguet, 2006), and using other determinants along with R&D as innovation inputs (physical capital

investment for process innovation in Parisi *et al*., 2006, and Hall *et al*., 2009, and ICT in Polder *et al*., 2009).

Besides R&D expenditures, several studies have also indicated that patenting activity is prominently affected by other firm-level determinants as well. Regarding the size of firms, earlier finding of Geroski *et al*. (1997) asserted that larger firms innovate steadily over a period of time, as it reflects access to better financing (Mairesse and Mohnen, 2002). But this happens till a threshold level, beyond which firms fail to innovate persistently. However, as pointed out by Cefis and Orsenigo (2001), although innovation persistence seems to increase with firm size, the relation is rather sector specific and country specific. Nevertheless, among other determinants, firm size may affect the marginal costs of patent application. The cost per patent application for small firms are expected to be higher than large firms since most of the small firms neither have a specialized unit dealing with patents nor property rights. Also they do not have detailed prior information about the patent system. In addition, it is argued that small firms hesitate to apply for patents because of the large patent litigation cost (Cohen and Klepper, 1996). Nesta and Saviottti (2005) also confirmed size advantage of firms for their performance in patenting in the U.S. pharmaceutical industry, which however had a declining effect over time. But, in sharp contrast to the above findings, empirical studies like, Acs and Audretsch (1991) and Pavitt *et al*., (1987) have found that, small firms tend to innovate comparatively more.

In addition, many studies point at the substantial differences in which innovation activities evolve based on the entrants, surviving and incumbent firms. Papers like Audretsch (1995) and Klepper (1996) provide theoretical insights into the nature of this dynamics of innovation due to the difference in the age of the firms. An empirical study by Balasubramanian and Lee (2008) exhibited a significantly negative relation between the age of firms and their patenting activities. The fundamental reasoning given for this outcome is the plausible existence of inertia

due to ageing, which deters the innovation productivity of the firms. Additionally, the drugs and medical sector was found to possess the highest inertial effect, having the largest negative coefficient on age with respect to patenting. Also, the turbulence in a sector, as measured by the entry and exit of firms can occur in the absence of innovativeness (Malerba and Orsenigo, 1996). Hence entry and exit of firms reflect on their survival mechanisms, that consequently affect the heterogeneous processes of innovation and growth.

Innovation is an inherently dynamic process between heterogeneous firms. But most empirical studies conclude that there is no strong and clear cut evidence of persistence in innovation activities. Montalvo (1997) referred to possible simultaneity problems in the relationship between patents and R&D. The previously employed count models were based on strict exogeneity of the expenditure in R&D with respect to patents. However, once a patent is granted, the firms may need to invest in R&D in order to transform the patent into a more commercial innovation for obtaining benefits. From this viewpoint R&D is used as a predetermined variable rather than being strictly exogenous. Hence, it is worthwhile to investigate if R&D affects the effective patenting in the pharmaceutical industry, which boosts further R&D expenditures in order to propagate greater patenting activities, so that the persistence in the innovation process is maintained.

But Peters (2009) finds a strong persistence in innovation input, both in terms of R&D or non-R&D innovation expenditure, as well as in terms of new products or processes in the market. Also Peters (2007) infers that success breeds success, as the past share of innovative sales influences positively the probability of innovating in the future. Based on the work of Duguet and Monjon (2004), there exists a strong persistence of innovation at the firm level, provided that the theoretical modeling is based on the firm size. Both Roper and Dundas (2008) and Antonelli *et*

al. (2010) confirmed on the persistence of innovation, focusing on the Irish Innovative Panel and the Italian manufacturing firms respectively.

In the context of pharmaceutical industry where patents are important, persistence in innovation is presumed to occur, as patent protection covers the costs of introducing and developing new innovations. Patents create a dynamic benefit on the pharmaceuticals, which outdo and dampen the short-run inefficiencies created by patenting (Philipson and Mechoulan, 2003). Since the R&D expenditure in the pharmaceutical industry is extremely high, patent protection plays a pivotal role in this sector compared to other high-technology sectors. But the pharmaceutical innovation activity is a complex phenomenon with the problem of moral hazards and adverse selection. However, based on a study by Hubbard and Love (2004), lower costs for medicines and lesser expenditure on pharmaceutical R&D can be maintained without patenting, but by using other alternative techniques that they promulgate. Nonetheless, this sector provides prominent benefits in improving health conditions by effectively reducing the burden of human diseases through their unique innovation mechanisms (Lichtenberg, 2004).

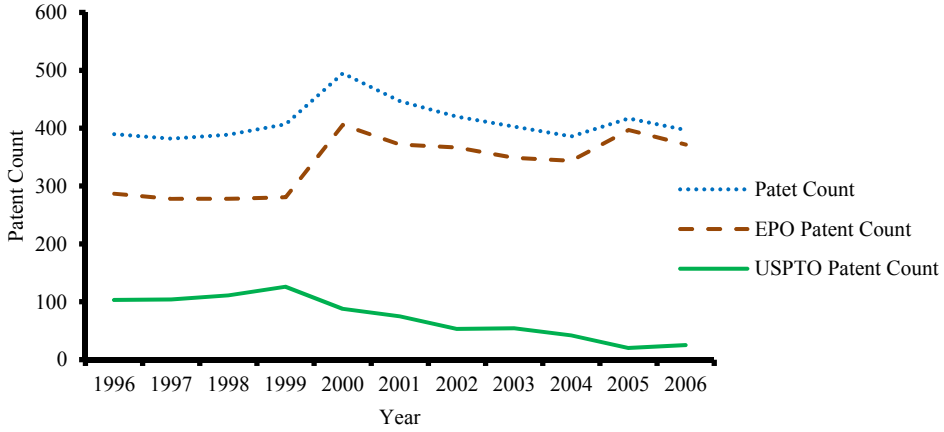
5.3 Data descriptive

We continue with our analysis using our compiled unbalanced panel data set for the period 1996-2006. With the determination of the ownership structure of the Dutch pharmaceuticals and using the ultimate parent firms, we construct our unique database using various data sources. A detailed elaboration on the data source and the subsequent variable constructions are provided in chapter 3. Therefore we proceed with the descriptive analysis of our data, which is contextual to this study.

As has been discussed before, the innovative performance of the firms is indicated by patent counts and citation-weighted patents, which acts as the innovation output indicator in the research intensive Dutch pharmaceutical sector. Due to the

availability of detailed information in our data set, it was possible to perform analysis on not only the overall patents, but also on EPO and USPTO patents individually. As depicted in figure 5.1, we find that there is a trend of gradual increase in the EPO patents over the concerned time period.

Fig. 5.1: Patent Counts for 1996-2006



It is evident from the diagram that the highest number of total patent counts is in the year 2000, taking a downward trend for the next four years. But again after 2004, there is an increase in the number of patents for a year, reaches another peak in 2005 until it takes a downturn again. The EPO patent counts show similar trends, due to the fact that, Dutch firms apply for patents mostly in the European patent office, and not in the US patent office (also evident in table 5.1). In year 1999, we find the highest peak for the trend in USPTO patent application, whereas a dent is observed for the EPO in the same year. In the latter years, the USPTO patents exhibit a gradual decline.

But the number of patents applied in the US patent office is found to be remarkably lower than the patents applied in the EPO throughout the considered time period. A lucid explanation for this may be the geographical proximity of the European

patent office as compared to the US one. Additionally, as pointed by Quillen *et al.* (2002), the rigorous EPO examination system reduces the cost of post-grant litigation compared to USPTO patent examination. Another imperative factor to be considered for the statistically lower number of patents applied to the US patent office is that, until 2000, the USPTO patent application were not published before the patent was granted. A study by Graham *et al.* (2002) asserts that the likelihood of an EPO patent having an equivalent USPTO patent is higher in case of software or semi-conductor industries than in the biotechnology or pharmaceuticals, which reflects that the patent applicants in the latter industries have a greater tendency to opt for national instead of a global intellectual property protection strategies. Moreover, the presence of non-industrial patent assignees like government laboratories and universities in these sectors are more likely to pursue less for global patent applications.

Fig. 5.2: Citation counts for 1996-2006

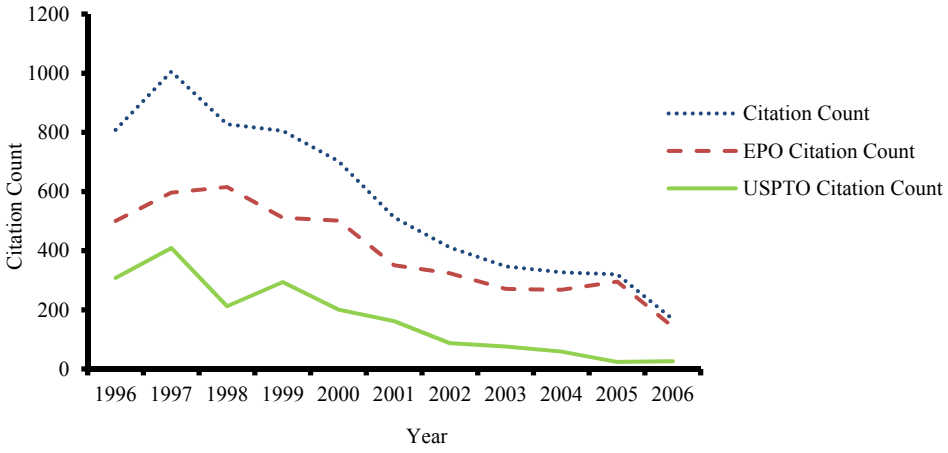


Figure 5.2 illustrates the trends in citation counts. Since citations counts (or citation-weighted patents) is the number of citations that each granted patents receive, it thereby reflects the innovation quality or performance of the pharmaceutical firms at large. From the above figure, the overall citation counts

reaches its peak in 1997, after which it shows a downward trend in general. Based on a study by Cockburn (2007), although a huge investment is made in R&D, the rate at which drugs are introduced is suffering from severe decline. A survey by *The Economist* asserts similar trends in the worldwide data, reporting estimates of global industry R&D expenditure increasing from $30 bn per year in 1994 to $54 bn in 2004, with global drug launches decreasing from 40 per year to 26 per year over the same period of time. The downturn in overall patenting and citation counts after 2005 may also be attributed to the structural Healthcare Reform of 2006 for the merging of the sickness funds and private insurers, which caused a prominent transformation in the regulatory policies to a considerable extent. It induced managed competition in the pharmaceutical arena which might already have an impact in the innovation output of the Netherlands pharmaceuticals. However the possibility of truncation error cannot be completely neglected, since 2006 was the last possible year for which patent information was available.

Our approach to use both R&D intensity and patent data allows us to exploit differences between innovators and non-innovators, both at the level of R&D expenditure and patent activities. A descriptive statistics on R&D and patent behavior of sample firms is reported in table 5.1.

Table 5.1: Innovation data sample

	R&D Reported	R&D Not Reported	Total
All Firms	191	329	520
Patenting Firms	44	28	72
Only EPO Firms	19	15	34
Only USPTO Firms	2	3	5
Both EPO and USPTO Firms	23	10	33
USPTO Patent Counts	613	188	801
EPO Patent Counts	3192	539	3731

Our innovation data consists of 520 firms for every year during the period 1996-2006, after selecting the possible ultimate parent firm operating in the industry. Among the 520 firms, 191 firms reported R&D for minimum one year. Similar statistics is carried out for all patenting firms, which includes the firms having EPO

and (/or) USPTO patents. It is evident from the table that, the total number of patents over the period 1996-2006 is 4532, with 3731 EPO patents and only 801 USPTO patents. Hence, an overwhelming majority of the Dutch patenting firms have used the European patent office, and not the patent office in US. We have already discussed on the lower number of patent applications in the US patent office while interpreting figure 5.1. In addition, appendix (A.1) provides a brief description on the general differences between the EPO and the USPTO patenting systems.

We also find that the total number of firms that patents is only 72 out of the 520 firms, wherein 44 patenting firms report R&D and 28 patenting firms do not report R&D. Therefore, a large group of pharmaceutical firms are not engaged in patent activities, perhaps due to the probable uncertainty and high costs that patenting incurs. Also, a majority of these firms can also be classified as non-R&D firms. As pointed by Licht and Zoz (2000), a large share of patents is applied by only a small number of firms and therefore the distribution of patent applications among firms is highly skewed. Similarly, the number of pharmaceutical firms that reports R&D are much lesser than the number of firms that does not report R&D. Many firms avoid reporting their R&D activities simply because of strategic reasons. On the other hand, lower propensity to patenting or R&D activities among the pharmaceuticals may be because, innovation activities involves huge volatile cost and time lapse in discovering and developing new ethical drugs. Hence more firms resort to generic drug productions, as the cost of generic drug production is prominently less than the original drug that it copies.

However it is surprising to note that, firms with patents sometimes do not report R&D. But this ambiguity can occur due to certain criterion followed while constructing the data file. Firstly companies sometimes report only the “material” R&D expenditure, and so the CIS waves or R&D survey may report R&D as zero (but not necessarily) if R&D expenditure is non-material. Alternatively, companies

may say nothing about their R&D and keep their R&D expenditure as confidential. In such cases, R&D is reported as 'not available'. It is also likely that companies reported as "not available" include some which are randomly missing, that is, a company performs material R&D, but for some reason Statistics Netherlands could not accrue the data for a particular year or a given period.

Besides R&D, the other explanatory variables included in our static model are size of firms (e_{it}), age of the firms (a_{it}) and entry-exit barriers (β_k, γ_k). For estimation purpose, a log transformation has been used in order to allow for the skewness of the distribution. It can be noted that we do not include competition measures in this chapter, as we aim at focusing primarily on the effect of the elemental firm-level determinants on the innovation performance of the Dutch pharmaceuticals, whereby the efficacy of the results obtained is checked using different innovation output measure. However a detailed analysis using various competition measures have been adopted in chapter 6. Nonetheless, the entry-exit dummies considerably represent the competitive conduct of the firms, whereby competition in a market is prominently affected by the influx and efflux of firms.

Table 5.2 represents the summary statistics of the variables used in the static model, and obtained from our unbalanced panel dataset, consisting of 520 pharmaceutical firms for the period 1996-2006.

Table 5.2: Summary statistics

				<----------Quantiles---------->				
	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
Patent Counts	5720	0.79	9.3	0	0	0	0	210
Citation-weighted patents	5718	1.09	15.51	0	0	0	0	564
Number of employees	3880	814.09	3308.31	1	1	7	66	26575
Entry	5676	0.45	0.5	0	0	0	1	1
Exit	5676	0.46	0.5	0	0	0	1	1

Focusing solely on the innovation data, it is seen that the patent counts are reported for the entire sample of the data (since the total number of observations in our

database in 5720). Concurrently, the citation-weighted patent variable has an almost complete set of observations. However, as we have noticed before, R&D intensity is not reported by most of the firms. For the other control variables used, we have 3880 observations for employment, and hence there exists many missing observations for the size of firms. Entry-exit dummies and age variable is almost complete, having observations for most of the firms in the concerned time frame.

Finally, we have tried to capture the dynamics of the innovation process by incorporating lagged patents (y_{it-n}) and lagged patent dummies (*ydummy*) as the explanatory variables. Innovation persistence is an important determinant for the concentration of innovation activities of firms. It can be noted, in this context, that both firm size and entry-exit are supposed to play a major role in the innovation persistence of firms. Various empirical studies like Geroski *et al.* (1997) and, Duguet and Monjon (2004) stress on the fact that, innovation persistence is influenced by the size of firms. In addition, competitive turbulence, as defined by the entry-exit or survival of the firms, is significant for dynamics in innovation (e.g. Antonelli *et al.*, 2010 and; Malerba and Orsenigo, 1996).

5.4 Empirical methodology

In this study, we focus on adopting statistical models of counts (non-negative integers) in the context of panel data and using them to analyze the relationship between patents and R&D expenditures. A panel data analysis of knowledge production function was initiated by Pakes and Griliches (1980), who defined a theoretical model relating innovation input to innovation output. They derived a distributed lag regression, where the number of patents was regressed on current and five lags of R&D and firm individual effects. In their specification they ignored the discreteness of the patent data and used the 'within' estimator to account for individual effects. Pointing out the limitation of this study, Hausman *et al*. (1984) proposed a number of panel data models in order to estimate the patent-

R&D relationship that took into account the discreteness of the patents, namely the fixed effect and the random effect Poisson and negative binomial regressions.

The discreteness of patent data motivates us to use the count model, similar to Nesta and Saviotti (2005). Our model is an application and generalization of the Poisson distribution to allow for independent variables, persistent individuals (fixed and random effects) and noise or randomness in the Poisson probability function. A Poisson distribution is the foundational framework of a count data model. For the univariate Poisson distribution (Poisson $(y_{it} \mid \lambda_{it})$), the occurrence of *y* over an exposure time *t*, has the probability function as:

$$P_0(Y = y_{it}) = \frac{\exp(-\lambda_{it})\lambda_{it}^{y}}{y_{it}!} = P_0(y_{it}, \lambda_{it}) \qquad \text{where } y = \{0,1,2,\ldots\} \tag{5.1}$$

In equation (5.1), Y is the number of patents, λ_{it} is the Poisson distribution parameter, where the first two moments are, $E(Y) = \lambda_{it}$ and $Var(Y) = \lambda_{it}$. The equality of the conditional mean and the conditional variance is due to the equidispersion property of the Poisson distribution.

Count outcomes are often characterized by a large proportion of zeroes which is also evident in patent data, since relatively few firms have patents. Most firms decide not to patent due to the huge cost incurred, time lapse, market uncertainty or just for strategic reasons. Patent statistics does not cover the incremental or imitative innovations. Hence, there are many firms in our data which are never granted any patent for the entire sample period and consequently, there are several zero patent counts in our patent data. Although linear and logistic models have often been used to analyze count outcomes, the large number of zero values are likely to cause the resulting estimates to be inefficient, inconsistent and biased. Owing to the huge number of zero patent counts in our dataset, we estimate the

innovation output using a zero-inflated count model. Zero-inflated count model has been used in the works of Hall (2000) and Min and Agresti (2005), allowing for unobserved heterogeneity by means of random effects. To model this zero inflation, we consider $(1-\gamma_{it})$ of extra zeroes and $\gamma_{it}\exp(-\lambda_{it})$ as the Poisson distribution.

Hence the zero-inflated Poisson (ZIP) density function can be written as,

$$\Pr(y_{it} \mid \gamma_{it}, \lambda_{it}) = \begin{cases} (1-\gamma_{it}) + \gamma_{it}\exp(-\lambda_{it}) & \text{if } y_{it} = 0 \\ \gamma_{it}\, \mathrm{P}_0(y_{it}, \lambda_{it}) & \text{if } y_{it} > 0 \end{cases}$$

This can be re-written as,

$$\Pr(y_{it} \mid \gamma_{it}, \lambda_{it}) = (1-\gamma_{it})P_0(y_{it}, 0) + \gamma_{it}P_0(y_{it}, \lambda_{it}) \tag{5.2}$$

where $(1-\gamma_{it})$ is the probability of extra zeroes. In the context of the Poisson density function, the Poisson distribution has two possible data generating process. For the first process with probability of γ_{it}, only zero counts are generated. For the second process with a probability of $(1-\gamma_{it})$, counts are generated from the Poisson model. γ_{it} is expressed as,

$$\ln(\gamma_{it}(1-\gamma_{it})^{-1}) = z_{it}'\delta \tag{5.3}$$

where z_{it} represents the vector of the zero-inflated covariates and δ is the vector of zero-inflated coefficients to be estimated.

Subsequently, we model $\ln\lambda_{it}$ as,

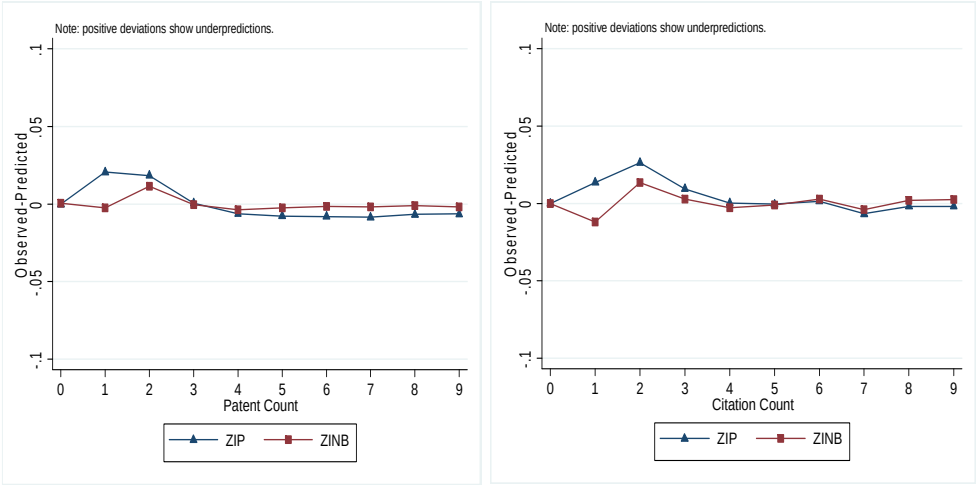
$$\ln\lambda_{it} = (\theta_i + \delta X_{it}) \tag{5.4}$$

where θ_i is a time-invariant unobserved firm effect and, X_{it} is the vector of independent variables that includes the log of fitted R&D intensity ($R\&D_{it}$), the log of the number of employees (e_{it}), age of the firms (a_{it}) , time dummies (α_k), entry dummies (β_k) and exit dummies (γ_k) .

But the Poisson model is restricted in the sense that, the conditional variance always equals to the conditional mean. However in most real-life data, conditional variance exceeds the conditional mean, causing overdispersion. The negative binomial model, which is a generalization of the Poisson model allows for overdispersion by introducing unobserved heterogeneity for each observation. Hence, in our analysis, we test the statistical properties of various count data models and adopt the zero-inflated negative binomial model (ZINB) instead of ZIP, as it takes into account the unobserved heterogeneity with respect to the propensity to patent and the ability of firms to generate inventions (Cincera, 1997).

The diagram below represents the residuals from the tested ZIP and the ZINB models using patent counts and citation counts as the dependent variables in the first and the second graph respectively. Smaller residuals indicate a better fitting model. It is observed that in both the graphs, the ZINB model is more closer to the zero value than the ZIP, which indicates that the former has a smaller residual than the ZIP value. Hence, it suggests that the ZINB model is preferred over the ZIP model for a given set of data. In addition, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) implicate that ZINB is the preferred model.

Fig. 5.3: Comparison of ZIP and ZINB Model



In case of the ZINB distribution the probability function is,

$$P_0(y_{it},\lambda_{it}) = \frac{\Gamma(\phi^{-1}+y_{it})}{\Gamma(\phi^{-1})+\Gamma(1+y_{it})}\left(\frac{\phi^{-1}}{\phi^{-1}+\lambda_{it}}\right)^{\alpha^{-1}}\left(\frac{\lambda_{it}}{\phi^{-1}+\lambda_{it}}\right)^{y} \tag{5.5}$$

where $\Gamma(.)$ denotes the gamma function. It reduces to ZIP when $\phi = 0$.

In our random effect zero inflated count models, the random effects are assumed to be standard normal variables multiplied by standard normal probability density function that enters the log-likelihood function. The log-likelihood for the zero inflated count model with random effects is given by,

$$\log L = \sum_i \log \rho(b_i) + \sum_i \sum_t I_{it} \log(\gamma_{it}) + (1-I_{it})\log(1-\gamma_{it}) + I_{it}\log[P_0\{y_{it},\exp(\theta_{3i}+\delta X_{it})\}] \tag{5.6}$$

where ρ is the standard normal probability density function and, I_{it} is an indicator variable which is equal to 1 if $y_{it} > 0$, and 0 if $y_{it} = 0$.

The patent count data is fully observed in our sample and consists of patents from United States (USPTO= U.S. Patent and Trademark office) and Europe (EPO=European Patent Office). However the R&D data is conspicuously inadequate. Nevertheless, the model used in this chapter is adapted to a panel data framework, where R&D availability is not necessarily a prerequisite. The insufficient number of R&D observations necessitates us to calculate a fitted value for R&D intensity by applying Heckman's Tobit II estimation technique. Therefore, in case of firm's unobserved $R\&D_{it}$, we consider its predicted values from the Tobit II estimation, which is applied for regressing R&D intensity on the other firm level determinants (the Tobit II estimation result is present in appendix A.5.2). In this method for computing predicted R&D, the selection criterion for the panel data is such that we use data on the firms that report R&D and compute the predicted R&D for those firms which do not report their R&D effort. Examples of empirical studies that uses similar R&D selection criterion, in a cross-sectional dimension, are given by Griffith *et al*. (2006), Klomp and Van Leeuwen (2006) and Hall *et al*. (2009). A detailed Tobit II estimation methodology can be found in section 4.4 of chapter 4.

Subsequently, we calculate the effect of R&D intensity, so obtained, along with the other determinants on the patenting activity for all firms in our dataset. 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. However, it should be noted that the R&D intensity variable, after merging with the accrued fitted R&D observations do not account for a complete set of observations. The reason behind the missing fitted R&D values is the presence of missing observations for the independent variables that are applied in the Tobit II estimations, and thereby influencing the obtained predicted R&D.

5.4.1 The unobserved heterogeneity issue

An important feature in the panel data application is the unobserved heterogeneity. Although our preference for Negative Binomial over Poisson model to account for overdispersion also takes into consideration random effects, we additionally incorporate the maximum likelihood (ML) technique in our model which is similar to chapter 4. Therefore, this ML method is applied to further develop our model, following the approach proposed by Wooldridge (2005) in a dynamic panel data context. This approach is suitable for handling the individual effects, conditioned on the initial values and the within-means of the time-variant independent variables. It deals with the plausible correlation between the unobserved individual effects and the exogenous variables, similar to the approach proposed by Chamberlain (1984).

In this case, the distribution of the unobserved effects (a_i) are modeled as,

$$a_i = \alpha_0 + \alpha_1 y_{i1} + \alpha_2 \overline{X}_i + \xi_i \tag{5.7}$$

where α_0 is a constant, $\overline{X}_i$ is the vector which includes the time averages of the explanatory variables ($e_{it}, a_{it}, R\&D_{it}$), y_{i1} is the initial value of patents, α_1 and α_2 are the corresponding coefficients (vectors) to be estimated, and ξ_i is the unobserved individual effect which is assumed to be independent following normal distribution $\xi_i \mid x_i \sim N(0,\sigma_\xi^2)$; $a_i \mid \overline{x}_i, y_{i1} \sim N[\alpha_0 + \alpha_1 y_{i1} + \alpha_2 \overline{x}_i, \sigma_\varepsilon^2]$. Hence, a conditional likelihood is obtained, similar to standard static random effect, where the joint distribution of the observations are conditioned on the initial values. In our regression analysis, ξ_i is assumed to be zero.

However, while calculating the time averages $\overline{X_i}$, we have omitted the initial period explanatory variables, as in Hesketh and Skrondal (2013), in order to

prevent biased conditions. Since the constrained version of Wooldridge (2005) includes the explanatory variables of all concerned periods (including the initial period), the within means in this approach can be significantly biased. It is because, the conditional distribution of the unobserved effects depend directly on the initial period explanatory variables instead of the explanatory variables of all concerned periods. At times, the dependence occurs only with the initial period explanatory variables and the initial dependent variables, resulting in a serious problem of biased results. Recent studies that pinpoints on this drawback includes Conti and Pudney (2011), Akay (2012) and, Hesketh and Skrondal (2013) among others. Hesketh and Skondral (2013) provides two plausible solutions to this issue, which includes, either incorporating the initial period explanatory variables as regressors along with their within means of all periods, or excluding the initial period explanatory variables from the within means. We have considered the latter solution for our empirical estimations. Therefore, we can re-write equation (5.7) as,

$$a_i = \alpha_0 + \alpha_1 y_{i1} + \alpha_2 \overline{X}_i^{'} + \xi_i \qquad (5.7)'$$

where, $\overline{X}_i^{'} = \frac{1}{T-1}\sum_{t=2}^{T} X_{it}$

This method to control for unobserved heterogeneity has also been elucidated previously in chapter 4. We repeat on its methodology again in the aforementioned section of this chapter, in order to emphasize on the contextual difference of its application (viz., in the context of count data model, rather than the two-step Tobit II estimation).

5.4.2 Extension to dynamics

Due to the richness of our panel data , we extend our empirical model to a dynamic framework. Thus, we try to find the plausible dynamics in the innovation process, i.e., whether the propensity to patent in the current year depends on the past history of patenting by the individual firms.

With specific reference to Netherlands, existing studies that have investigated the dynamics of the relationship between R&D and patenting activity include Van Leeuwen (2002) and Raymond *et al*. (2009). Both studies confirm persistence of innovation. Firms may innovate persistently for a number of reasons. Past innovation performance provides a broader technological opportunities for firms, such that continuity in their innovation process is maintained. In other words, accumulation of knowledge would induce state dependent invention flows and hence, persistence of innovation. Another theoretical reasoning considers the sunk costs in R&D investments as a predominant source for steady innovation as they create entry barriers and hence, engagements to continue innovation.

In our model, we try to investigate whether firms exhibit persistence in innovation output by using lagged patents and patent dummies for the past years, within the concerned time frame. Using patent lags and lagged patent dummies might throw some light on individual firm's propensity to patent. The requirement to allow for such individual effects eliminates much of the variance in the available short time series framework.

Hence, our basic model (eq. 5.3) gets transformed to the following two equations:

$$\ln \lambda_{it} = (\theta_i + \delta x_{it} + y_{it-n}) \tag{5.3A}$$

$$\ln \lambda_{it} = (\theta_i + \delta x_{it} + ydummy) \tag{5.3B}$$

where, y_{it-n} denotes the patents of the previous years until year n (in our model n=10), and *ydummy* denotes the lagged patent dummy. The lagged values of patents and patent dummies are used interchangeably in different models in our analysis (and, not together in any of the models) to avoid the problem of endogeneity. It can be noted that, there might exist a autocorrelation between the initial values of the dependent variable and the lagged patents when we allow for random effect into our analysis. Additionally, due to the skewness in the patent

data, the lagged patent dummy may better reflect the dynamics in innovation performance, by minimizing any error in the patent data that may exist. Hence, we would prefer the lagged patent dummy as a better indicator of persistence in innovation output, in the Dutch pharmaceutical industry.

5.5 Empirical estimation

5.5.1 Estimations using simple patent counts as the dependent variable

Our basic model incorporates the R&D intensity variable, taken as its predicted value from the preferred Tobit II equation (annexed in the appendix section A.5.2). The other independent variable that we consider in the basic model is the log of the number of employees (as a proxy for the size of firms). We further use the log of firm age and entry-exit dummies as additional regressors. Finally, dynamics is incorporated in the model by using a lagged patent dummy and lagged patent counts interchangeably.

In this section, we use simple patent counts (overall patents, only EPO patents and only USPTO patents) as our dependent variable. To overcome the problem of excess zeroes, we have used the zero inflated negative binomial model. A Vuong test (Vuong, 1989) for each of the estimations is applied, in order to discriminate between negative binomial (NB) and zero-inflated negative binomial (ZINB) models. This test corrects for the complication that ZINB reduces to NB only at the boundary of the parameter space.

Table 5.3 provides the estimates of the patent equation. From this table, we can find that model 1 and 2 are the basic models where no random effect is assumed. However from model 3 onwards we have introduced initial conditions and averages based on Wooldridge (2005), for handling individual effects. Accounting for random effects, conditioned on the initial values of the dependent variable and the exogenous variables is pertinent to our panel data model due to plausible

selectivity bias. But model 3 assumes random effects by excluding the initial conditions and considering only the averages of the explanatory variables, while model 4, 5 and 6 takes into account full random effect. A key point to note in this context is, we lose 1996 data in our analysis in order to alleviate the initial condition problem. Therefore, to keep parity amongst all the regression results, we have considered the data from 1997 onwards for all the regression models (also for the models where we do not account for random effects). We incorporate dynamics by introducing lagged patents and lagged patent dummy in models 5 and 6 respectively, reflecting whether each successive patenting builds on its predecessors in an effective way. In the light of the above regression techniques it can be noted that, the same analyses have been performed in the subsequent estimations using different regressands, as reported in table 5.4, 5.5 and 5.6 (which are discussed in the subsequent sections).

From table 5.3, it is observed that R&D intensity (as the fitted value) has positive and highly significant effect on patenting, as observed in all six models, confirming results from past literatures which unanimously asserted the same (as in Pakes and Griliches, 1980; Licht and Zoz, 2000; and Nesta and Saviotti, 2005). This explains that R&D intensity is an important determinant in generating new knowledge. However, the coefficients are less than unity, suggesting that an increase in R&D expenditure causes a less than proportionate increase in the expected innovation output. This confirms the assertion by Acs and Audretsch (1989) that not all inventions are patented.

Table 5.3: ML-regression results for the patent equation

Dependent Variable	Patent Counts	Patent Counts	Patent Counts	Patent Counts	Patent Counts	Patent Counts
	ZINB Model 1	ZINB Model 2	ZINB Model 3	ZINB Model 4	ZINB Model 5	ZINB Model 6
Log(R&D per employee)	0.391***	0.549***	0.327***	0.201***	0.198***	0.205***
	[0.075]	[0.063]	[0.062]	[0.052]	[0.050]	[0.051]
Log(Employees)	0.529***	0.967***	0.314**	0.277***	0.254***	0.234***
	[0.053]	[0.065]	[0.129]	[0.085]	[0.084]	[0.084]
Log(Age)		-2.248***	-0.286	-0.113	-0.041	-0.078
		[0.285]	[0.266]	[0.165]	[0.153]	[0.150]
Entry		-4.229***	-4.371***	-3.079***	-2.482***	-2.688***
		[0.727]	[0.622]	[0.539]	[0.481]	[0.504]
Exit		-1.001***	-1.558***	-0.752***	-0.816***	-0.707***
		[0.264]	[0.284]	[0.220]	[0.213]	[0.217]
Lag(Patent)					0.025***	
					[0.006]	
Dummy(Patent)						0.940***
						[0.221]
Intercept	-2.743***	1.023	1.617**	2.391***	1.836***	1.439***
	[0.561]	[0.732]	[0.656]	[0.429]	[0.408]	[0.452]
Initial(Patent)				0.015***	-0.006	0.015***
				[0.002]	[0.006]	[0.002]
Random Effect	NO	NO	YES	YES	YES	YES
Log likelihood	-1060.548	-1016.291	-989.014	-873.894	-784.906	-789.757
N Observations	1197	1197	1197	1197	1197	1197
Nonzero observations	189	189	189	189	189	189
Zero observations	1008	1008	1008	1008	1008	1008

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

For the first regression model, we have used only fitted value for R&D intensity and firm size as the explanatory variables. In the latter models, additional variables are introduced. Nevertheless, coefficients for both R&D intensity and the size of firms are found to be systematically positive and significant in all the regression results, ruling out the possibility of any omitted variable bias in model 1. The categorically consistent result for big pharmaceutical firms patenting more actively bears similarity with the earlier study by Nesta and Saviotti (2005), among others. This confirms, a priori, that larger pharmaceutical firms have a comparative advantage over their smaller counterparts in terms of generating patents. Our inference on firm size is based on logical grounds, as it stems from the fact that,

although firms may profit from research activities being performed elsewhere, but the production of patents is dependent on their internal competencies with a better reserve for capital and human resources.

However, concerning the role played by the age of firms on patenting, our results suggest a systematic negative effect in all the regression models. This corroborates that, younger firms are more innovation prone than the older firms, as their involvement in innovation is crucial for their survival and growth in the market (Audretsch, 1995). In the work by Balasubramanian and Lee (2008), age is found to have a significantly negative effect on patenting due to the presence of firm inertia that has the tendency to mitigate the inventive productivity with the increase in firm age. But contrary to our findings, they also purported a negative effect of firm size on patenting as well. Nonetheless, a robust positive effect on the size of Dutch pharmaceutical firms in our analyses leads us to the consensus that young firms, which essentially create more employment, are more susceptible to patenting. However, the negative effect is found to be significant only when no random effect is allowed (in model 2), suggesting that the results might be biased. In the consecutive models where random effect is assumed (model 3, 4, 5 and 6), the coefficient for firm age becomes insignificant, thereby depicting a non-deterministic and inconsequential relation. Hence, the negative effect of age on patenting is not confirmatory and the insignificant results suggest that age of firms may not affect the patenting activities of the pharmaceutical firms substantially.

For the entry and exit dummies, a consistently systematic negative and significant effect is observed, which indicates that a non-competitive environment is prerequisite for patenting in the Dutch pharmaceutical industry (confirming the traditional Schumpeterian viewpoint). The coefficient for entry dummy is found to maintain a negative coefficient with high magnitude, implying that an increase in entry barriers causes a more than proportionate increase in patent output. Likewise, the exit dummy also reflects a robust negative relation with patenting, which

however falls below unity when full random effect is assumed in the last three estimation results. Therefore, our persistent empirical result for entry-exit dummies suggests that the inflow and outflow of firms, causing turbulence in the pharmaceutical industry, can be a hindrance to its knowledge output.

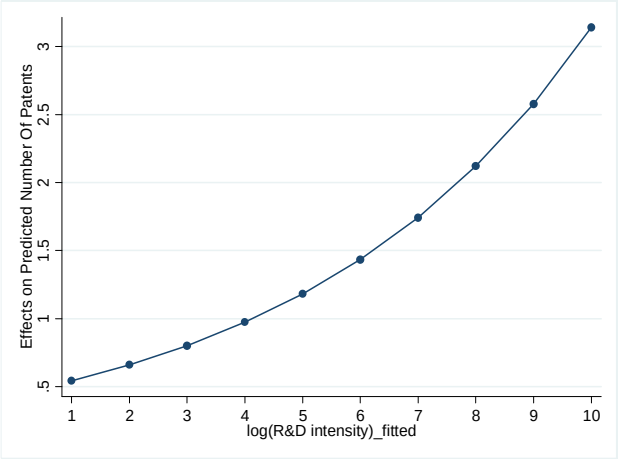
The positive and significant values for the lagged patents and lagged patent dummies (in model 5 and 6 respectively) provide evidence of a persistence in innovation output among the pharmaceutical firms. But the coefficient for patent dummy has a greater magnitude than the coefficient for lagged patent. Nonetheless, a highly significant coefficient in both the cases proves that firms that patents in the past years have a strong tendency to patent in the following years, which in a way insinuates a condition of sequential innovation in terms of patenting.

Among all the regression models in table 5.3, only model 5 and 6 considers additional characteristics of full random effects and the presence of dynamics. However in case of model 5, coefficient for the initial patent is not found to be significant, which may be due to the probable autocorrelation between the lagged patent values and the initial patents. Hence, as mentioned in section 5.3, it is worthwhile to consider lagged patent dummy as a better indicator to capture dynamics when we allow for full random effects as compared to lagged patents. Additionally, most of the coefficients are significant with expected sign in model 6. Therefore, based on the empirical digression, model 6 is found to be statistically superior to the other models and so we consider it to be our preferred model.

However, the sign of the coefficient provides the direction of the effect, but not the marginal effect. Therefore we try to take a closer look on how an unit change in R&D intensity affects the conditional mean of patents. But since the consideration of mean marginal effect is subject to several criticisms (Bartus, 2005), we consider the average marginal effect whereby the computed marginal effects are averaged. Fig. 5.4 represents the average marginal effect of fitted log(R&D) on the predicted

number of patents (the marginal effect of representative values are considered), which is found to be increasing in function.

Fig. 5.4: Average marginal effect of fitted log(R&D intensity) on the predicted number of patents



Note: Calculations are based on the regression estimates in Model 6 of table 5.3

5.5.2 Estimations using EPO and USPTO patents as the dependent variables

Due to our detailed data structure, it has been possible to demarcate the EPO patents from the USPTO ones. Since the EPO and USPTO patent applications are subject to divergent structural and institutional policies (detailed discussion available in the appendix A.5.1), it is of analytical interest to empirically investigate on how the various economic determinants affect them, when considered separately. Hence, table 5.4 provides the estimation results when EPO and USPTO patent counts are considered as the regressands, instead of total number of patents.

However the Netherlands pharmaceuticals seldom apply for USPTO patents, as it is more strategically and institutionally convenient for them to apply for patents within Europe. As seen in the innovation data sample of table 5.1, the number of

USPTO patents that report their R&D expenditure are significantly less than the EPO patents, and hence they do not capture the true picture of innovation output of our concerned sector in the Netherlands.

Owing to the paucity of non-zero USPTO patent data, it is difficult to achieve convergence in the ZINB estimations. Hence, in order to achieve convergence using the USPTO data, a number variables which are used in the main equation are omitted in the inflation equation. Since the estimation technique is not completely robust in this case, we only consider model 5 and 6 for USPTO patents. It can be noted that, model 5 and 6 are our final models, where we allow for full random effect, by including both averages and initial conditions. Furthermore, dynamics is also included in these last two models. Thus, including the estimates with the USPTO patent counts for model 5 and 6 allows us to attempt for a comparative analysis with the estimates obtained using EPO patents.

The estimation results using EPO and USPTO patents in table 5.4 bear similarities to that of table 5.3 where overall patent counts are used. The fitted R&D intensity is found to be perpetually positive and significant, with decreasing returns to scale, for both EPO and USPTO patent data. Similarly, the coefficient for firm size is positive in all the regression models. However it loses its significance for the EPO patent counts in model 3, where only averages are used to take into account random effects. But the effect of firm size remains significant when full random effect is assumed. In case of the regression estimates using USPTO data, the coefficient for firms size exhibits a positive sign, but insignificant. Altogether, it may be reconfirmed that, ceteris paribus, larger firms with a greater employment capacity, have a relative advantage over the smaller firms in terms of patenting. Further, the coefficient for age appears to be negative steadily, but loses its significance when we allow for unobserved heterogeneity in case of EPO patent counts. But the negative value for age shows high level of significance with magnitudes greater than unity for the USPTO patent estimation. This hints towards

newer firms having a greater propensity to patent. The entry-exit dummies appears to be negative and significant in all the models where EPO patent count is used, but loses its significance in case of USPTO data, confirming the traditional theoretical assumption that absence of competition is crucial for the existing firms to patent.

Table 5.4: ML-regression results for the EPO and USPTO Patent equation

Dependent Variable	EPO Patent Counts	EPO Patent Counts	EPO Patent Counts	EPO Patent Counts	EPO Patent Counts	EPO Patent Counts	USPTO Patent Counts	USPTO Patent Counts
	ZINB Model 1	ZINB Model 2	ZINB Model 3	ZINB Model 4	ZINB Model 5	ZINB Model 6	ZINB Model 5	ZINB Model 6
Log(R&D per employee)	0.407***	0.527***	0.324***	0.179***	0.192***	0.173***	0.322***	0.345***
	[0.075]	[0.063]	[0.065]	[0.047]	[0.046]	[0.046]	[0.118]	[0.101]
Log(Employees)	0.547***	0.957***	0.184	0.202**	0.189**	0.178**	0.246	0.129
	[0.052]	[0.066]	[0.170]	[0.086]	[0.085]	[0.086]	[0.152]	[0.138]
Log(Age)		-2.059***	-0.312	-0.072	-0.032	-0.048	-3.209***	-3.323***
		[0.289]	[0.239]	[0.126]	[0.127]	[0.117]	[0.653]	[0.665]
Entry		-3.565***	-3.809***	-2.377***	-2.047***	-1.999***	-44.926	-39.487
		[0.712]	[0.666]	[0.481]	[0.441]	[0.470]	[5.71e+08]	[3.84e+07]
Exit		-1.095***	-1.570***	-0.726***	-0.754***	-0.677)***	-0.021	0.253
		[0.263]	[0.304]	[0.197]	[0.194]	[0.194]	[0.482]	[0.449]
Lag(Patent)					0.019***		0.091***	
					[0.006]		[0.022]	
Dummy(Patent)						1.128***		2.629***
						[0.246]		[0.315]
Intercept	-3.226***	0.328	0.992	1.870***	1.426***	0.800*	1.733	1.844*
	[0.585]	[0.722]	[0.727]	[0.378]	[0.377]	[0.425]	[1.095]	[0.984]
Initial(EPO Patent)				0.024***	0.000	0.023***		
				[0.003]	[0.008]	[0.003]		
Initial(USPTO Patent)							-0.165***	0.049***
							[0.049]	[0.014]
Random Effect	NO	NO	YES	YES	YES	YES	YES	YES
Log likelihood	-982.916	-942.047	-915.066	-777.901	-696.317	-696.478	-419.881	-399.412
N Observations	1197	1197	1197	1197	1197	1197	1197	1197
Nonzero observations	180	180	180	180	180	180	84	84
Zero observations	1017	1017	1017	1017	1017	1017	1113	1113

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

With the extension of our model to a dynamic framework, we find a positive and significant effect of lagged patent values as well as lagged patent dummies, thereby proving again the concept of persistence of innovation at the micro level for the pharmaceutical industry. A strong evidence of persistence in patenting is found for both EPO and USPTO patent counts. This confirms the findings of Van Leeuwen (2002) and Raymond *et al.* (2009), establishing the fact that firms are more prone to innovate if they performed innovation in the past years.

5.5.3 Estimations using citation-weighted patent counts as the dependent variable

Patent quality is proxied by the forward citation counts on each of the patents, to account for the great dispersion in the value of different patents (based on empirical studies like Jaffe *et al*., 2000 and Hall *et al.*, 2005). Citation counts provide a more accurate metric for the innovative performance of the firms, as many less important patents emerge from past inventions which might not actually be put to application. Limited number of patents are actually 'blockbuster' invention; while others are simply of intermediate or negative value that are used as a strategic step to ward off competition. Therefore, in this section, we focus and discuss on how R&D intensity and other determinants affect the quality of patents in the Netherlands pharmaceutical sector.

Although the estimation techniques used are akin to the past estimations, we have introduced year dummies in this case to control for the time effects, in order to circumvent any truncation bias. This is due to the fact that citations acquired by a patent is positively related to time (older patents have more time to receive citations than younger patents). The results for overall citation-weighted patents are enumerated in table 5.5.

Table 5.5: ML-regression results for the Citation-weighted patent equation

Dependent Variable	Forward Citation Counts	Forward Citation Counts	Forward Citation Counts	Forward Citation Counts	Forward Citation Counts	Forward Citation Counts
	ZINB Model 1	ZINB Model 2	ZINB Model 3	ZINB Model 4	ZINB Model 5	ZINB Model 6
Log(R&D per employee)	0.511***	0.636***	0.465***	0.199**	0.189*	0.186*
	[0.084]	[0.068]	[0.077]	[0.101]	[0.102]	[0.100]
Log(Employees)	0.321***	0.928***	0.346*	0.484***	0.398***	0.423***
	[0.060]	[0.070]	[0.180]	[0.092]	[0.100]	[0.102]
Log(Age)		-2.745***	-2.393***	-0.947*	-0.568	-0.065
		[0.306]	[0.494]	[0.549]	[0.459]	[0.431]
Entry		-5.397***	-5.514***	-4.075***	-3.458***	-3.330***
		[0.842]	[0.780]	[0.830]	[0.743]	[0.780]
Exit		0.128	-0.420	-0.310	0.036	-0.547*
		[0.321]	[0.409]	[0.302]	[0.375]	[0.291]
Lag(Patent)					0.012**	
					[0.006]	
Dummy(Patent)						0.655**
						[0.310]
Intercept	-1.05*	2.145***	2.416***	3.637***	2.870***	2.767***
	[0.572]	[0.788]	[0.766]	[0.774]	[0.764]	[0.829]
Initial(Citation)				0.005***	0.002	0.005***
				[0.001]	[0.002]	[0.001]
Random Effect	NO	NO	YES	YES	YES	YES
Time Dummies	YES	YES	YES	YES	YES	YES
Log likelihood	-1040.557	-1007.500	-993.835	-910.140	-834.046	-839.489
N Observations	1196	1196	1196	1196	1196	1196
Nonzero observations	169	169	169	169	169	169
Zero observations	1027	1027	1027	1027	1027	1027

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

A positive and significant relation is observed between fitted R&D intensity and patent citations in all the regression models. This finding is obvious and expected as it is vital for a firm to engage in conspicuous R&D activities for producing unique and good quality patents. However the coefficient for R&D intensity remains below unity, with a decreasing magnitude when random effect is assumed. This implies a decreasing returns to scale, where an increase in R&D causes a less than proportionate increase in citation counts of the patents.

Similar results are observed for the size of firms, which is found to be positive and significant in all the regression models. This reconfirms our past estimation results that bigger firms have higher endowment and capability to engage in producing higher quality patents. Again, affirming our prior results, the coefficient for age is significant with a negative sign when random effect is not assumed. But it loses its significance when average and initial values are added, although remaining negative. This seems to indicate, albeit less robust, that newer firms are more dynamic and motivated to produce genuine knowledge output.

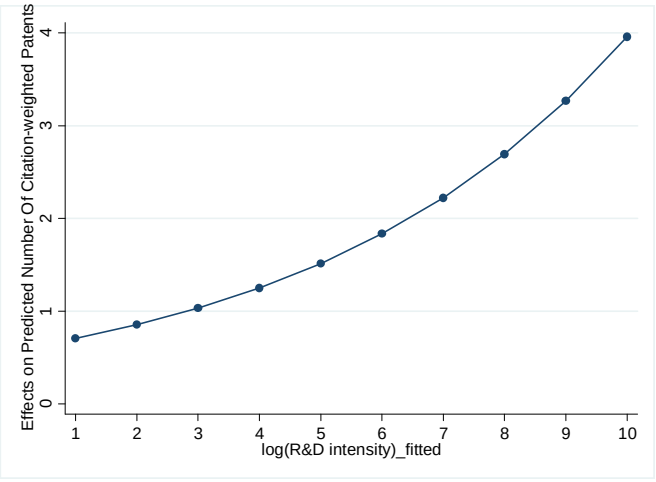
The entry dummy has consistently exhibited significantly negative values, with high magnitudes of the coefficients. This suggests that new entrants are subject to prominent barriers to enter the Dutch pharmaceutical arena. Although this conjecture seems to oppose our finding that newer firms are more innovative, we should also consider the fact that, only those new firms that possess a high level of human resource and strong pharmaceutical pipelines sustains. Moreover, the coefficient for the exit dummy is also found to be negative. But it is significant only in our final model, at 10% significance level. Thus far, the negative effect of entry-exit dummies on citation counts implies that the pharmaceutical firms are more inclined to innovate in a stable and non-competitive regime.

With the incorporation of dynamics, the regression results for both lagged patents and lagged patent dummies provide positive and significant effects. This again proves persistence in generating innovation output. It is noteworthy that the initial citation value loses its significance in model 5 due to the possible existence of autocorrelation with lagged patents (as described in the preceding section). Consequently, lagged patent dummy is a better indicator than lagged patent values for capturing the dynamics of innovation output. Therefore, we adhere to our previous inference of considering model 6 as our most preferred model.

Figure 5.5 represents the average marginal effect of fitted R&D intensity on the predicted citation-weighted patents. It is found to be identical to fig. 5.4, as the

average marginal curve is an increasing probability function with a positive slope. The only noticeable difference in fig. 5.5 compared to fig. 5.4 is that, the intercept of the marginal curve in the former is greater than that in the latter. On the whole, the marginal curve diagrams points to the same direction, that is, the instantaneous rate of change of the predicted innovation output increases with an increase in fitted R&D intensity values, subject to the other explanatory variables.

Fig.5.5: Average marginal effect of fitted log(R&D intensity) on the predicted citation counts



Note: Calculations are based on regression estimates in model 7 of table 5.5

5.5.4 Estimations using EPO and USPTO citations as the dependent variables

Finally, we perform regression on the EPO and USPTO citation-weighted patents. In section 5.5.2 we had attempted to delve into details on how EPO and USPTO patents react to changes in the de facto economic characteristics of the Dutch pharmaceutical firms. On similar grounds, we have tried to check in this section if the quality of EPO and USPTO patent data is affected differently by the various concerned determinants. The corresponding estimation results are summarized in table 5.6.

It is seen that the sign and scope of most of the explanatory variables are similar to that obtained when overall citation-weighted patents have been used; and for that matter, is also largely consistent with the results obtained for EPO and USPTO patent data as well. In line with our previous estimation results, the impact of fitted R&D intensity and size of the firms on both EPO and USPTO citation counts is testified to be positive and significant in most of the regression models. Regarding the age of firms, there seems to exit a negative and significant relation with EPO patent data for models 2 and 3, when full random effect is not allowed. It, however, loses its significance and becomes positive in sign when full random effect is assumed for the citation counts of EPO patent data. Nonetheless, the effect of age on USPTO citations is negative and significant, which again points towards younger firms being more enterprising and innovation prone.

Conforming again to the past results, entry dummy is persistently negative and significant when EPO citations are used as the dependent variable. However, when we allow for unobserved heterogeneity in case of the estimations using USPTO citations, it remains no longer significant. However the large magnitude of the coefficients for entry dummy, when USPTO data is used, is presumably due to anomalies that might occur because of the very few number of USPTO citation data and their subsequent convergence problems. But the exit dummy does not provide a consistent and unidirectional result for the two types of patent citations. Although it remains negative for the estimations dealing with EPO citations, it becomes positive when USPTO citations are used. Plus, it is highly significant in the latter case, while for the former it is significant only in model 4 and 6.

Table 5.6: ML-regression results for the EPO and USPTO citation-weighted patent equation

Dependent Variable	Forward EPO Citation Counts	Forward EPO Citation Counts	Forward EPO Citation Counts	Forward EPO Citation Counts	Forward EPO Citation Counts	Forward EPO Citation Counts	Forward USPTO Citation Counts	Forward USPTO Citation Counts
	ZINB Model 1	ZINB Model 2	ZINB Model 3	ZINB Model 4	ZINB Model 5	ZINB Model 6	ZINB Model 5	ZINB Model 6
Log(R&D per employee)	0.492***	0.567***	0.400***	0.091	0.184**	0.129*	0.825***	0.858***
	[0.077]	[0.067]	[0.078]	[0.098]	[0.074]	[0.070]	[0.160]	[0.160]
Log(Employees)	0.356***	0.922***	0.025	0.386***	0.354***	0.314***	0.764***	0.473***
	[0.057]	[0.071]	[0.256]	[0.092]	[0.095]	[0.084]	[0.144]	[0.137]
Log(Age)		-2.365***	-1.407**	0.177	0.048	0.293	-5.083***	-4.566***
		[0.297]	[0.601]	[0.318]	[0.308]	[0.222]	[0.713]	[0.657]
Entry		-4.128***	-3.731***	-2.338***	-2.367***	-1.816**	-53.309	-50.152
		[0.790]	[0.760]	[0.793]	[0.643]	[0.708]	[1.59e+09]	[5.61e+08]
Exit		-0.022	-0.643	-0.928***	-0.567	-1.001***	1.199**	1.074**
		[0.322]	[0.452]	[0.294]	[0.352]	[0.267]	[0.553]	[0.485]
Lag(Patent)					0.008*		0.032	
					[0.005]		[0.234]	
Dummy(Patent)						1.106***		2.789***
						[0.291]		[0.405]
Intercept	-1.558***	0.871	0.283	2.529***	2.308***	1.970***	6.193***	6.231***
	[0.554]	[0.752]	[0.792]	[0.772]	[0.591]	[0.614]	[1.385]	[1.285]
Time Dummies	YES	YES	YES	YES	YES	YES	YES	YES
Initial(EPO Citation)				0.011***	0.007**	0.011***		
				[0.002]	[0.003]	[0.002]		
Initial(USPTO Citation)							-0.003	0.011***
							[0.015]	[0.004]
Random Effect	NO	NO	YES	YES	YES	YES	YES	YES
Log likelihood	-931.853	-904.971	-891.620	-795.083	-720.308	-716.955	-481.361	-457.646
N Observations	1197	1197	1197	1197	1197	1197	1197	1197
Nonzero observations	157	157	157	157	157	157	83	83
Zero observations	1040	1040	1040	1040	1040	1040	1114	1114

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

Lastly, innovation persistence is confirmed, yet again, by positive and significant coefficients for lagged patents and lagged patent dummies in most of the regression results, using EPO as well as USPTO citation-weighted patents. Only the coefficient for lagged patent is insignificant in model 6 of USPTO citations. The corresponding insignificance in the initial USPTO citation values once more refers to a potential autocorrelation problem between them.

In consideration of our intensive and meticulous empirical investigation, we may infer that our ZINB estimates have unanimously pointed towards the same outcomes. Although patent counts and citation-weighted patents are different measures for calibrating innovation output, they have proven to be complementary by providing similar estimation results. Hence, we may summarize our empirical findings that, younger Dutch pharmaceutical firms, having high employment capacity, are more likely to invest in persistent innovation output, under a non-competitive paradigm. Although the negative coefficient for age is not always found to be significant, the rest of the determinants provided robust significant inferences.

5.5.5 Testing for endogeneity

In our above estimates, we have considered that the regressors are exogenous. But the problem of endogeneity can be crucial if ignored, as the estimator might be inconsistent in that case. Hence to control for the presence of endogeneity, we have used the non-linear instrumental variable approach (as in Cameron and Trivedi, 2009), and have found that there does not exist the possible endogeneity problem in our dataset. Further explanation is provided in Appendix A.5.3.

5.6 Conclusion

Based on a comprehensive empirical study, this chapter revisits at the firm-level the effect of R&D intensity and other determinants on the innovation output of the

firms for the Dutch pharmaceutical industry. The discreteness of patent and citation data entails us to perform count data analysis. In consideration of the excess zero values for patents and consequently, their citation numbers in our dataset, a zero-inflated negative binomial model has been used.

From our empirical analysis, the R&D investment appears to have paid off as it facilitates essential knowledge output. Research involved in discovery of new ethical drugs, including the pre-clinical and clinical phases involves huge costs and effort. Therefore it is likely that the pharmaceutical firms that engage in R&D expenditure have more inclination to patent in order to safeguard and protect their innovations from the generic firms and recoup the high cost involved in R&D. However, the elasticity for R&D intensity being below unity, is ubiquitous in all our regression models. This leads to the underlying fact that not all R&D expenditures lead to possible innovation output, as much of the R&D activities in this industry deals with imitative and incremental innovation (Licht and Zoz, 2000). Further the significant effect of the other determinants confirms the findings of Nesta and Saviotti (2005) that, R&D investment is a necessary but not a sufficient factor for pharmaceutical firms to engage in patenting activities.

Further our analysis suggests that, the bigger sized firms have a greater propensity to engage in innovation output. As posited by Licht and Zoz (2000), larger firms are presumably more aware of the strategic aspect of patenting, cross-licensing agreements or collaborative R&D activities. Also, conventional wisdom suggests that large firms are more innovative as they have greater access to capital stock and human resources to engage in innovation. In the pharmaceutical sector, mergers and acquisitions might tend to increase the knowledge base, thereby amplifying the innovation output.

At the same time, empirical evidence suggests that, young firms are probably more enterprising in terms of innovation performance. Although the effect of age on innovation output is not robust when we allow for unobserved heterogeneity in our

empirical models, a persistent negative sign evokes the possibility of smaller pharmaceuticals firms being more involved in patenting activities. This finding is in line with the economic rationale that connotes to the idea of innovation performance as the crucial survival strategy for new firms in the pharmaceutical market. In addition, it supports the study by Balasubramanian and Lee (2008), who asserted that, the existence of inertia when a firm ages tends to attenuate their innovation productivity.

But, Balasubramanian and Lee (2008) also affirmed a negative relation of firm size with patenting, which is in stark contrast to our results obtained on the same. Moreover, the negative and significant effect of entry and exit dummies in most of our regression models hints at a strong barrier for the firms to enter and exit the Dutch pharmaceutical market, relating to the early Schumpeterian hypothesis of a non-competitive environment. In other words, it suggests that this sector performs better in productive innovation in the absence of any turbulence caused by the entry and exit of firms.

In conjunction with our findings, Pammolli and Riccaboni (2007) asserted that, pharmaceutical industry operate in a regulated environment which is inter-alia connected to the barriers to entry that prevails predominantly in this sector as a result of natural, legal or strategic phenomena. Since our results bear a strong evidence of barrier to entry in the drug market, it is pertinent to consider that, only those young firms which has strong pharmaceutical pipeline and sufficient reserve of capital and other financial resources mainstays. This instigates the idea that the new firms that have emerged from old firms within the industry, as a result of spin-offs or possible mergers and acquisitions might be the ones that play an important role in innovation output. This conceivable argument supports the concept of potential business strategies that these firms embark upon, in order to ward off competition.

In line with this economic rationale, our results further suggest a highly significant persistence in innovation output, which is justifiable as entry-exit barriers propagate persistence in patenting. Innovative pharmaceutical firms will have the tendency to innovate further due to the huge sunk cost that it involves. On the other hand, the decision to innovate in one period itself enhances the tendency to innovate in the next periods, thereby affirming persistence in innovation in the innovative pharmaceutical sector. Nevertheless, this perpetual innovation performance may have emerged from an effective cycle of R&D investment and innovation output and not from sequential innovations, as persistence also seems to prevail when citation counts of the corresponding patents are considered. In any case, persistence in patenting appears to be a pivotal strategic measure that the firms undertake to form barriers in the market and evade competition.

Our analysis is done in-depth, considering the regression results for both EPO and USPTO patents and citation-weighted patents individually, along with overall patents counts and their citation weightage. Since our findings using patent counts are homologous to the ones using their corresponding citations, the former can be used as a reasonably reliable measure of innovation activities in the Netherlands' pharmaceutical industry. Confirmatory conclusion from our results proposes that, large pharmaceutical firms in a closed territory, have a steady innovation performance over time, without hampering the patent quality (in consonance with the findings of Duflos, 2006). Therefore, the inevitable patenting strategies to curb competition can lead to prominent market inefficiencies, that leads us to question whether patenting is really desirable. Hence, further studies should give more impetus to the social drawbacks caused by patents and their possible alternatives, as suggested by Hubbard and Love (2004).

Altogether, the detailed exploration, conditioned on the various determinants and their causal effects on knowledge output, is formidable for the construction of economic policies, and provides important empirical arguments to formulate the

theories based on innovation strategies for the pharmaceutical industry. However it is worthy to note that, our result is specific to the pharmaceutical firms in the Netherlands, and might show variations in divergent medical and scientific landscapes, under different socio-economic conditions. In addition, from the empirical perspective, a caveat of this analysis relates to the data we use to implement the model. One major assumption in our model is that, the effect for non-R&D reporting firms is the same as R&D reporting firms. This may tend to bias our results to a certain extent. Hence, as a prelude to further research, more investigation can be devoted to the characteristics and matching procedures of R&D reporting and R&D non-reporting firms.

A.5 APPENDIX

A.5.1 A brief description of EPO versus USPTO patents

Patents and their citations are largely used to measure knowledge spillover from the R&D activities of the firms. But there lies prominent institutional differences in the process of governing the decision of granting a patent, or including a patent citation in a patent document. Although a few aspects of patent law has been harmonized internationally, there still remains a number of important differences between them. Since, in our analysis, we also consider EPO and USPTO patents individually as the dependent variables, we try to take a closer look at the differences between them.

The first difference between the EPO and USPTO patents are the priorities given when two candidates apply for a patent for the same invention. In case of EPO patents, the only thing that counts is the filing date. Based on the Patent Act 1995, the first candidate to have filed his application will get the patent, even if the second candidate had come up with the invention first. But in the USA, a determination is made as to who invented it first. This usually involves examining laboratory logbooks, establishing dates for prototypes etc. So even if a person filed a patent later but is found to have invented earlier, he may be awarded a USPTO patent. Considering the rigorous clinical trials and laboratory research required to produce novel drugs in the pharmaceutical industry, it is of foremost importance to credentialize the innovator who actually comes up first with the life saving (or disease alleviating) innovation.

The second prominent difference is that, US patent law requires that the inventor include the best way to practice the invention in the patent application, which bars him from keeping essential or advantageous aspect a secret. In contrary, European patent law has no such requirement. It only requires that at least one way of practicing the invention needs to be included in the application. But it does not

focus on the fact whether the invention used is the best way or not. In this context, secrecy of pathological findings is not preferable, as scientists working on human pharmaceutical therapies usually are obliged to share their inventions to their peers, so that they can also accrue knowledge in the field of their research area. However, contrary to the other manufacturing industries, pharmaceutical inventions can be easily replicated with a very low capital investment. This persuades the actual innovators to keep much of their laboratory research in secrecy.

Another important distinction between the two systems is in the grace period. In case of EPO patents, if the invention has become publicly available (like selling the invention, giving a lecture about it, or showing it to an investor without a non-disclosure agreement), the patent application will be rejected. It does not make any difference whether the person making it publicly available is the inventor, one of the inventors or an independent third party. But for USPTO patents, a one year grace period is provided, which implies that the inventor can freely publish his invention without losing the patent rights. Hence, the less flexible nature of European patent office safeguards the invention. But at the same time, it hinders the scientific community, working on human pathology, to benefit from the knowledge of each other completely. In addition, although both EPO and USPTO requires that an invention be novel and requires an inventive step, EPO has a more strict interpretation of this term. A European patent application involves an inventive step if it solves a technical problem in a non-obvious way. This policy undertaken by the EPO serves as an effective instrument to actually demarcate the pharmaceutical innovations which are truly innovative from the ones which partially imitate their counterparts' inventions.

In regards to the institutional framework and geographical constitution, the US patent law is a federal statute. Since a US patent is a property right which is enforceable in the entire territory of the USA, it allows patent holder to prevent anyone from making, using or selling in the USA the patented invention. In

contrast, the European Patent Convention is a treaty signed by the twenty-seven European countries. As a granted European patent under the EPC confers to its owner the same right as a national patent in those EPC countries he elected in the application; a European patent once granted can only be annulled by separate proceedings in each elected country.

Also, there are relevant differences between citation practices in the USPTO and EPO. The US patent office follows the 'duty of candor' rule which imposes all applicants to disclose all the prior art they are aware of. Hence, many citations at the USPTO come directly from inventors and applicants and finally filtered by patent examiners. But the European Patent office follows no such rules. For the European patents, the patent examiners draft their report, trying to include all the technically relevant information within a minimum number of citations (Michel and Bettels, 2001). Hence, EPO patent citations are usually added by the examiners. Consequently, the analysis of diffusion and obsolescence of technological knowledge and knowledge spillovers may reveal different properties according to the used patent dataset.

The final concomitant distinguishing feature between the two kinds of patents is the two-part claims. European patent applications virtually always have a two-part claim. The latter features are those that constitute the invention. The former features are found in the prior art. If an application is filed with one-part claims, the foremost thing that happens is that the examiner identifies the closest prior art and requests that the claim be delimited there from. On the contrary, US patent applications always have one-part claims. If there exists a two-part claim in a US patent, chances are that the patent is owned by a European firm. The pharmaceutical research, involving different pharmaceutical techniques and inventive steps, in the process of developing new drugs (in the form of varying dosages and forms) are common and a significant part of the pool of knowledge.

Hence, in view of the state of the art, there exists only a narrow range of pharmaceutical developments that can be termed as truly inventive.

The main differences between EPO and USPTO patenting system is depicted through a schematic diagram in Graham *et al.* (2002).

A.5.2 Tobit II estimation to obtain fitted (R&D intensity)

For obtaining the fitted R&D intensity, we perform Heckman's Tobit II estimation, assuming full random effects. Thus, we have implicitly used the first two stages of the CDM model (put forth by Crepon *et al*., 1998) where R&D intensity is explained by its various economic determinants and subsequently, its fitted value and other explanatory variables are used to determine the innovation output at the firm level. The regressors that are considered to determine their causal effect on R&D are identical to the ones used for our count data analysis for explaining innovation output. However the entry-exit dummy is solely used in the selection equation of the Tobit II estimation, owing to the validation of the exclusion restriction. The following table elucidates the effect of the various explanatory variables on the R&D intensity in the Dutch pharmaceutical sector. However, a detailed analysis on this topic is done in chapter 4.

Table A.5.1: Static Tobit II estimation

Dependent variable	Probit (R&D =0/1)	Log of R&D per employee
Log(Employees)	0.020	-0.836***
	[0.043]	[0.073]
Log(Age)	0.066	0.297***
	[0.042]	[0.079]
Entry	-0.235*	
	[0.127]	
Exit	-0.492***	
	[0.111]	
Intercept	-2.181***	0.394
	[0.168]	[0.587]
Time Dummies	YES	
Initial[Log(R&D/Employee)]		0.457***
		[0.034]
Initial[R&DProbability]	0.922***	
	[0.096]	
Random Effect	YES	
Log-likelihood	-1757.897	
ρ	0.384***	
	[0.113]	
σ	1.489***	
	[0.056]	
λ (Inverse Mill's Ratio)	0.571***	
	[0.182]	
Estimation Method	**Heckman Tobit II Sample Selection**	

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

A.5.3 Testing for endogeneity by using non-linear instrumental variable method

We test our model for the existence of endogeneity, that is, whether one or more regressors are endogenous. The possible reasons for endogeneity may be due to omitted variable bias, potential causality or data measurement error. Since our data encompasses a fairly long time period of 11 years, there is a higher likeliness for the existence of endogeneity between innovation output and various firm level determinants. However the problem is difficult to determine, as there might be lag

periods between successful production of knowledge output, and the lags vary for individual firms. Again, the causal relationship may not be straightforward and may be beyond the scope of empirical anticipation.

Nonetheless, we attempt to investigate whether a potential causality exists between the size of firms (measured as the log of employment) and the dependent variables (measured as patent counts or citation counts), using the structural model approach as proposed by Cameron and Trivedi (2009).[6] Therefore, firm size is considered as the probable endogenous variable in our model. We consider the rate of depreciation (d_{it}) and the log of capital intensity (c_{it}) as the instrumental variables (IV), which are excluded from our basic ZINB model. We consider that the rate of depreciation and capital intensity has a direct link to the size of firms, but may not have any effect on their innovation output. The condition for robust identification is that, there must exist at least one valid instrumental variable.

In this approach, the dependent variables (y_{it}) and the endogenous regressor (e_{it}) are defined in explicit models. We use the structural equation for the ZINB model, where the mean depends on an endogenous regressor, which is given by,

$$y_{it} \sim ZINB(\eta_{it})$$

where, $\eta_{it} = E(y_{it} \mid e_{it}, \mathrm{x}_{\mathrm{it}}, \xi_{1i}) = \exp(b_1 e_{it} + b_2 x_{it}' + \xi_{1i})$ (A.3.1)

[6] Since we have considered a fitted R&D value from a Tobit II regression estimate for the R&D intensity variable, we omit the endogeneity test for the same, as IV tests might provide spurious conclusions. However, endogeneity between R&D and innovation output is conceptually obvious, as a cycle of R&D and patenting can be ascertained through economic theory. The other continuous variable in our model is the firm age, which is evidently an exogenous regressor, as age is determinate and innovation output cannot have a reverse causality with the same.

In eq. (A.3.1), x_{it} is the vector of exogenous variables in our model and ξ_{1i} is the unobserved heterogeneity correlated with our endogenous regressor e_{it} but uncorrelated with x_{it}. The error term ξ_{1i} is added to allow for endogeneity.

Next, a linear-reduced form equation is specified to show the relation between e_{it} and ξ_{1i}

$$e_{it} = x_{it}'\sigma_1 + i_{it}'\sigma_2 + \vartheta_{1i} \tag{A.3.2}$$

In the above equation, i_{it}' is the vector of instrumental variables, or the excluded exogenous regressors.

The error terms are related as,

$$\xi_{1i} = \rho\vartheta_{1i} + \varpi_i \tag{A.3.3}$$

where, ϖ_i is independent of ϑ_{1i}. The error term ϑ_{1i} is assumed to affect both y_{it} and e_{it}, after controlling the dependence between x_{it} and i_{it}. Hence, if $\rho \neq 0$, then e_{it} is assumed to be endogenous.

For this analysis, we use the two step estimator procedure in STATA. It is to be mentioned in this context that, for the second stage ML-regression, we have used the Wooldridge approximation to deal with the initial condition problem. Hence all variables are considered from 1997, except the initial patent value, for both the first stage and second stage regression models. The following table represents the results as obtained in the first stage linear regression with IV in a panel data framework.

Table A.5.2: First Stage Linear regression

Dependent Variable	Depreciation	Log (Capital intensity)	Log (R&D per employee)	Log (Age)	Entry	Exit	Patent dummy	Intercept
Log (Employees)	0.000011*** [1.33e-06]	-0.308*** [0.021]	-0.228*** [0.021]	0.086*** [0.032]	-1.846*** [0.445]	1.672*** [0.342]	0.094 [0.107]	4.901*** [0.329]

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

From the linear panel regression result, it is found that both depreciation and capital intensity are highly significant, with a positive and negative sign respectively, although the value of the coefficient for depreciation is considerably low.

In the second stage, we fit the ZINB model on the regressors and the predicted residual that we obtain in the first stage, as enumerated in table A.5.3. As a reference for the regressors, model 6 has been applied where lagged patent dummy is used as an additional regressor to capture the dynamics in innovation.

From table A.5.3, it is observed that, the z-statistics for the coefficients of the predicted residual has a p-value of 0.276 and 0.196 when we use patent counts and citation counts respectively, as the dependent variables. In this context, it is to be noted that, the coefficient of the predicted residual provides for the Wald test of the null hypothesis of exogeneity, where the null hypothesis $H_0 : \rho = 0$. It is observed that the z-statistics has a p-value greater than 0.05 for both the estimated models, leading to the non-rejection of the null hypothesis at 5% significance level. Hence, the presence of endogeneity between the size of firms and innovation output is rejected in both the cases.

Table A.5.3: Second stage ML-regression results with residuals

Dependent Variable	Patent Counts	P>|z|	Citation Counts	P>|z|
	Model 6		Model 6	
Log(R&D per employee)	0.136**	*0.049*	0.049	*0.723*
	[0.069]		[0.139]	
Log(Employment)	0.229	*0.172*	0.519***	*0.003*
	[0.168]		[0.177]	
Age	-0.033	*0.838*	0.404	*0.166*
	[0.161]		[0.292]	
Entry	-0.633	*0.240*	-2.681***	*0.006*
	[0.539]		[0.973]	
Exit	-0.117	*0.752*	-0.420	*0.319*
	[0.369]		[0.421]	
Dummy(Patent)	2.549***	*0.000*	2.086***	*0.000*
	[0.279]		[0.373]	
Initial(Patent)	0.018***	*0.000*		
	[0.003]			
Initial(Citation)			0.008***	*0.000*
			[0.001]	
Random Effect	YES		YES	
Time Dummies	NO		NO	
Predicted Residual	-0.144	***0.276***	-0.272	***0.196***
	[0.132]		[0.211]	
Intercept	-1.419**	*0.037*	3.515***	*0.002*
	[0.681]		[1.150]	
N Observations	762		761	
Nonzero observations	118		103	
Zero observations	644		658	

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

However, further investigation on endogeneity can be done using other instrumental variables, in order to check if we arrive at the same definite conclusion.

Chapter 6

COMPETITION ANALYSIS AND ITS EFFECT ON INNOVATION PERFORMANCE

Abstract

In this chapter, we focus on an in-depth theoretical and empirical analysis of the different competition indicators and their corresponding effect on the citation-weighted patent counts of the Dutch pharmaceutical industry. For the purpose of our analysis, four different indicators of competition have been applied, to arrive at a more consistent inference on the competition-innovation relationship. Apart from the market share concentration index, the other metrics indicating competition in this study include, the Lerner index, mark-up with adjustment for economies of scale and profit elasticity. To check for the existence of any non-linear relation between competition and innovation performance, quadratic specifications for the competition measures have been incorporated. In addition, we have implemented three alternate empirical models for our primary estimations, and the best fit model has been used for further analysis. In the context of empirical results, the competition measures mostly indicate a negative relation between competition and innovation performance, with the market share index and the mark-up hinting at a U-shaped relation between competition and innovation output. Apart from the competition measures, the empirically determined effects of the other determinants unanimously signify that bigger sized firms, with higher level of R&D activities have greater propensity to perform knowledge output persistently. Analogues results were observed when identical estimation procedures were undertaken on the firms which were closer to the technological frontier.

Keywords: Competition, Citation-weighted patents, Count models, PPML

6.1 Introduction

In chapter 5, we have dealt with the effect of R&D intensity and other determinants on the innovation output of the Dutch pharmaceutical industry. Although the entry-exit dummies indicated lack of turbulence (hinting at a less competitive regime), a more detailed analysis on how product market competition affects innovation in this particular sector was not embarked upon. In addition, the previous study did not delve into the plausible non-linearity that might exist between product market competition and innovation output (as was suggested by Aghion *et al*., 2005 and many subsequent research works). Hence in this chapter, we provide a key insight to the applicability of various competition measures and their corresponding relationship with innovation output in the Netherlands' pharmaceuticals, considering a nonlinear framework.

Investigations on the inherent relation between competition and innovation came to cognizance several decades ago by Schumpeter (1943), who estimated linear cross sectional relationship and typically found a negative relation between competition and innovation, which was based on theoretical convictions that prevailed at that time. However, Scherer (1967) allowed for additional non-linearities in a cross-sectional analysis and discovered a significant inverted U-shaped relation between them. Likewise, a strong evidence of an inverted U relationship in a panel data framework was subsequently brought forth much later, in the work of Aghion *et al.* (2005). But most of the research since Scherer's findings had resorted back to estimations with linear specifications, in line with Nickell (1996) and Blundell *et al.* (1999), who obtained a positive effect of competition on innovation on a panel data framework. However more recent studies by Hashmi (2013) and Beneito *et al*. (2014) negates the positive relation hypothesis, and asserted a negative relation between competition and innovation output.

Nevertheless, the divergence in outcomes can be plausibly attributed (at least partially) to the industries under consideration and the geographic location in

which they operate. The competitive strategies undertaken by the research-intensive industries play a pivotal role in determining the extent of their innovation output. This is mainly due to the fact that, the innovation output of those firms is not only to recover their R&D investments, but also for strategic reasons. Competition acts as a key factor for keeping down prices and production costs of off-patent drugs. However, the competition advocacy in this sector is not always simple and requires us to take a holistic view. The vagaries of the R&D expenditure and the innovation process, along with the subsequent costs and delays of the drug authorization procedure causes new drug development a risky affair, which involves huge expenditure. It can be noted in this context that, innovation expenditure in the pharmaceutical industry is a high risk venture with high upfront R&D cost. There can be failure of new drugs at any stage of the approval process. Also, the drugs that are approved are likely to face competition from their rival's products. But, successful drugs that are protected from competition by patenting (or in other words, intellectual property rights) can provide huge profits to the firm such that they are able to recoup the R&D costs that the innovation process incurred. The probable high cost for branded medicines, that may be anticipated due to a non-competitive environment in the pharmaceutical industry, can be mitigated to some extent by proper streamlining of regulations. Nonetheless, the existence of ubiquitous health insurance facilities in almost all developed countries insulates the consumers from the prices of the drugs that they consume.

In spite of the considerable profits over and above the cost of innovation by patent protection, it cannot be denied that the branded drugs are marketed with a conspicuously high price. Hence, after a pioneer's patent in the market expires, there is a strong incentive for the influx of generics, resulting in competition. But it cannot be gainsaid that competition caused by generic entry in the pharmaceutical market is economically inefficient, as it simply duplicates an invention without introducing new capabilities for developing new technologies. Also it can be noted that, a pharmaceutical patent does not grant complete monopoly right to the

manufacturer for treating any particular disease. This naturally transcends into product competition within the therapeutic classes. Hence, competition in the pharmaceutical market can also occur within the research-intensive firms, and not necessarily by generic entry. Moreover, sequential innovation also persists in this industry, which results from the internally generated research. Since in this case, each successive innovation builds up on the preceding innovations, it can be an important determinant to generate competition.

In view of the above backdrop, this chapter explores the level of competition using various competition measures in the Dutch pharmaceutical industry from 1996-2006, and subsequently analyzing the intrinsic relation that exists between competition and innovation output in this sector. By focusing on the Dutch pharmaceutical market, it is evidently characterized by high generic uptake, but at the same time, considerable high price margins for drugs. The introduction of the latest healthcare policy in the Netherlands, which aims at buffering the high drug prices and fostering managed competition, might have an influential effect on the pharmaceutical innovation. A detailed elaboration in this context has been already done in chapter 1. However, since our analysis on the Netherlands' pharmaceuticals is based until 2006, it involves policies prior to the enforcement of the Healthcare Reform Act in 2006. Hence the aftermath of this significant institutional change in the Netherlands' health sector on the pharmaceutical industries is beyond the scope of this chapter.

Concerning the structural outline of the remaining part of the chapter, the rest of its contents unfolds as follows. Section 6.2 discusses the literature survey for the subsequent linkage between innovation output and the competition measures. Section 6.3 and 6.4 presents the data descriptive and empirical model respectively. The empirical estimations are discussed in section 6.5, and finally section 6.6 concludes.

6.2 Literature review

The early foundation of the competition-innovation analysis was rooted to the Schumpeterian hypothesis that more concentrated market (and hence, less competition) stimulate innovation. In his seminal contribution, Schumpeter (1942) claimed that imperfectly competitive markets are necessary for achieving innovative progress. Although the Schumpeterian conjecture is vehemently criticized and altercated in many subsequent studies over the past decades, the traditional Schumpeterian viewpoint seems justifiable for the innovation intensive pharmaceutical sector, on the grounds of theoretical reasoning. It is due to the fact that, patents may create temporary monopolies thus increasing the market power. Although patents play a key role in the pharmaceutical sector, enabling the firms to recoup their investments and make a return on their efforts; the patent strategies might block the development of competing products. Furthermore, the sector enquiry identified strategic (and perhaps anti-competitive) behavior might cause market access delay for generic products, like patent clusters. In fact, it is commonly observed that generic entry does not take place as early as it potentially could.

In the later empirical investigations on the relation between competition and innovation, results ranged from supporting the Schumpeterian hypothesis to findings that are completely opposite. Kamien and Schwartz (1982) revealed an inconclusiveness of the relation between market power and innovation activity. Subsequently researchers like Geroski (1990), Nickell (1996) and Blundell *et al.* (1995) obtained disproportionate evidence against the Schumpeterian hypothesis. According to them, controlling for technological opportunities caused their results to be diametrically opposite to that proposed by Schumpeter.

More recent works including Aghion *et al.* (2005), proposes an inverted U-relation between them. According to Aghion *et al*. (2005), competition discourages laggard firms from innovating but encourages neck and neck firms to innovate. Subsequent

research by Askenazy *et al.* (2008) proves that, there exists an inverted U-shaped relationship between innovation and competition, when firm and innovation sizes are controlled for. However, in the work of Tingvall and Poldahl (2006), it is asserted that the inverted U-shaped curve is sensitive to the competition measure used, where he finds a significant inverted U relation between competition and innovation using Herfindahl concentration index, but not with price cost margin.

A subsequent study by Correa (2011) has proved that the inverted U-shaped relationship as obtained by Aghion *et al.* (2005), is solely due to the structural break in the data in the early 1980s. The structural break plausibly occurred due to the formation of the United States Court of Appeals for the Federal Circuit in 1982. Based on this study, a positive relationship between competition and innovation for the period 1973-1982 is observed, while no statistically significant relation could be inferred for the period 1983-1994. Hence the study by Correa claims that the inverted U- relation as proposed by Aghion *et al.* (2005) does not hold true. It is noteworthy that Aghion *et al.* (2005) himself does not find a statistically significant inverted U curve when he uses R&D expenditure instead of citation-weighted patents, as the indicator for innovation.

For capturing accurately the innovation output, many empirical literatures have sagely designated the citation-weighted patents as the most effective measure. Gayle (2003) points out that citation-weighted patent counts more accurately measure innovation than simple patent counts. He argued that simple patent counts consider technologies covered by patents as equal in terms of their economic and social value. Hence citation weighted patent counts account for the heterogeneity of technologies covered by patents. A measure of innovative activity which accounts for the heterogeneity of technologies that are covered by patents, is not distorted by minor patenting and consequently, they are more reflective of true innovation.

Based on the work of Gayle (2003), the relationship between industry competition (as measured by concentration) and innovation output (as measured by the citation-weighted patent counts) is indirect. He finds that, the amount of R&D expenditure in an industry increases directly with the industry's concentration. Subsequently, the amount of industry's patenting increases directly with the level of R&D expenditure. However, he posited that, given a constant level of R&D, patent activity is independent of industry concentration. Nevertheless, this indirect measure suggests that there exists a positive relationship between citation-weighted patent counts and industry concentration, through R&D investment. Therefore, Gayle's research discards the positive effect of competition on innovation, due to the prevalence of strategic patenting.

Besides Gayle (2003), some latest research (like Hashmi, 2013 and Beneito *et al.*, 2014) have also corroborated to a negative link between competition and innovation output. Hashmi (2013) articulated a negative relation between competition and citation-weighted patent counts for manufacturing firms in the U.S., that involves in public trade. However on tallying his theoretical model to that of Aghion *et al.*, (2005), both negative and inverted U-relation was established. Another very recent paper by Beneito *et al.* (2014), asserts that patenting lowers competition for the manufacturing firms in Spain.

According to the European Commission report (2009), the pharmaceutical sector suffers from substantial lack of competition, which is inter alia linked to the existence of natural, legal and strategic entry barriers. Consequently the concentration ratio of the pharmaceutical sector is moderately high. Moreover, the sector is characterized by the dominance of relatively small group of big pharmaceutical companies which represents a significant part in the annual turnover.

There exists a plethora of academic research on the interaction of competition and innovation. However, no clear-cut consensus could be arrived at, which is mainly

due to the variation in the samples undertaken. In addition, the results obtained are also sensitive to the different metrics used to identify competition and innovation activities. A brief underpinning of the past empirical literatures that relate to the different competition measures used has been integrated in the following section.

6.3 Data descriptive

For our micro-level study of the competition-innovation interaction within the Netherlands' pharmaceutical industry, we attempt to identify our primary determinants with precision, such that we arrive at a robust and accurate inference. To serve this purpose, four alternate competition measures have been applied to our framework. Concurrently, innovation performance by the firms is indicated by the citation-weighted patent counts. Along with the competition measures, several control variables have been applied in our empirical analysis.

The compilation of the unbalanced panel data for the Dutch pharmaceutical industry has already been elaborated in chapter 3. Therefore we refer to section 3.1 for reviewing the data sources that have been used and the corresponding procedures applied to consolidate our data set of 520 Dutch pharmaceutical firms. In addition, the construction of the different variables is incorporated in section 3.2, and hence, we do not recapitulate on them in this chapter. However, in the following subsections, we briefly portray the different variables that have been used for our subsequent empirical estimations and their corresponding descriptive statistics.

6.3.1 Innovation indicators

Citation-weighted patents

Synonymous to chapter 5, our dependent variable is the innovation performance (or innovation output) of the Netherlands' pharmaceutical firms. However unlike the empirical investigation in chapter 5, we only apply the citation-weighted patents in

this study and not the patent counts. We preferred to use the citation-weighted patents over simple patents counts because, the former might reflect a more accurate causality with competition, that enables us to bypass any sequential or incremental innovation output that simple patent counts might have allowed. Hence, to focus not only on the quantity of innovation output, but also on its quality, we apply the forward citation-weighted patent counts (following the method of Hall *et al.*, 2000). The citation-weighted patents is the total number of cites of each patent in the concerned year. Hence, according to this measure, the value of each patent is reflected by the number of citations it receives.

R&D expenditure

Since R&D expenditures of the firms reflect the resources invested in order to fructify innovation, it is viewed as the innovation input indicator. Consequently, it enters our model as an important determinant that can affect the innovation performance of the firms, along with the competition measures. The R&D intensity is calculated as the ratio of R&D investment over the number of employees for each firm at the individual time period, and subsequently deflated to real terms. However many firms do not report their R&D expenditure data which resulted in many missing observations. So, in congruence with the model proposed by Crepon *et al.* (1998), the R&D intensity variable in our model is calculated as the predicted value that is obtained from a Tobit II estimation (estimation results annexed in the appendix A.5.2 of chapter 5).

Patent dummy

Similar to chapter 4 and 5, lagged patent dummy has also been introduced as a control variable. In a way, this measure indicates the dynamics of the innovation performance. More specifically, it provides with the plausible relation between past patenting propensity of the pharmaceuticals in the past years to its patent values of the current years. Therefore, this measure hints at the effective inter-linkage

between past patents and current patent values, thereby establishing the validity of persistence in innovation output.

6.3.2 Different measures of competition

The measurement and quantification of pharmaceutical competition is a complex task and calls for an in-depth analysis. Therefore, a concomitant feature of this study is to consider different competition measures from our available Dutch pharmaceutical data, such that their individual interaction with citation-weighted patents is examined. Early empirical literature, inspired by Schumpeter (1943) and many subsequent works used market concentration as the measure of competition among the firms. However, Caves and Porter (1978) and Sakakibara and Porter (2001) resuscitated the concentration measure by introducing the market share instability as a proxy for competition. In the subsequent work by Aghion (2005), he used the approach of price cost margin (PCM hereafter) or the Lerner index, similar to Hall (1988) for measuring competition. But this method of estimating mark-up (or PCM) rests on the strong assumption of constant returns to scale. Further refinement of the PCM was brought forth in the works of Badinger (2007), Amoroso *et al.* (2010), De Loecker and Warzynski (2012) among others, wherein the scale economies are clarified and incorporated into the model. The most recent of all competition indicators is the Boone indicator (or relative profitability measure), as proposed by Boone (2000) and Boone *et al.* (2007). This measure claims to circumvent the problems of 'reallocation effect' and 'selection effect' arising in the concentration measures or the PCM.

In view of the past literatures on different expressions for competition, we incorporate four alternate measures to quantify the competitive conduct of the Dutch pharmaceutical firms. This includes the market share concentration index, the Lerner index, mark-up with adjustment for scale elasticity and the profit elasticity. Due to the intricate nature of competition in the pharmaceutical industry, it seems more pertinent to advocate more than one competition measure, so that we

can investigate in details how different competition measures determines the innovation intensity in our concerned sector.

Since the computation of the different competition measures has already been covered in chapter 3 (subsection 3.2.2), we conduct a comparative data digression of the various competition measures in the following part of this section. However, before we elucidate on the different measures of competition, it is important to note that the first three competition indicators (that is, market share, Lerner index and scale adjusted mark-up) are inversely related to competition. This connotes to the fact that a rise in these measures indicates a fall in competition. The only exception is the profit elasticity. Since we have considered the slope which determines the percentage fall in profit due to one percent rise in marginal cost, the negative slopes so obtained is converted to their absolute values in order to measure competition. Therefore, the profit elasticity measure used in our analysis exemplifies the direct quantification of competition.

Concentration Index

The most formative method for measuring competition has always been the application and usage of market share to indicate the level of concentration in the market, which dates back to the works of Schumpeter (Schumpeter, 1942). A firm's market share can be viewed as a measure of dominance and hence, should affect a firm's innovative performance. Therefore the underlying assumption for more concentration reflects barrier to entry caused by the market power of the incumbents and consequently, less competition. Crepon *et al.* (1998) and Blundell *et al.* (1999) used firm's market share as a measure of competition, and found it to be an important determinant for measuring innovation of the firms.

As has been mentioned earlier in chapter 3, since our estimations are based on firm level analysis, we have used the domestic market share of each firm, and not the Herfindahl index (HI) which is essentially an industry level measure. However the

industry aggregate HI has been calculated over time (within the time frame of 1996-2006) for the sake of the diagrammatic representations in the descriptive statistics.

Although we have already used the market share measure as the indicator to determine competition in chapter 4, this measure may not fully capture the concept of competition in case of the selection and reallocation effect of competition. As noted by Tirole (1988), if intense competition driven by more aggressive interaction of the firms removes inefficient firms from the market, then market concentration rises. In such a situation, it wrongly signals at a decreased competition, while in actuality, more intense competition is the reason for the rise in the level of concentration. Moreover, fierce competition among the firms also reallocates revenues, where the market share of the efficient firms increases at the cost of inefficient firms. This wrongly indicates a decrease in competition due to an increase in market share. Hence the concentration measure is not appropriate in the context of output reallocation effect, as it basically adheres to the Cournot model with symmetric firms. This ambiguity in concentration measures prompts us to use other measures of competition as well, as the market share concentration index is not a predominant criterion to base economic policies on.

Also, in consideration of the pharmaceutical sector operation over time, a high market share in a single period does not necessarily indicate that there exists a low level of effective competition, when the market share can be quickly diminished by new entrants (either new patented drugs or the generics). New entrants require a period of testing time and approval process in the pharmaceutical industry. Since it is possible to figure out the likely new entrant in the short and medium term, an analysis of competition requires both the present market share and the probable consequence of new entry from the generics and the drugs that are in the approval process. Hence, in order to guard against the possibility of using poorly estimated

and erroneous measures of competition, we aim at a rigorous and in-depth analysis of different measures of competition for the Dutch pharmaceutical sector.

Price Cost Margin with Constant Returns to Scale

The second measure of competition that we take into account is the Price Cost Margin (PCM) with constant returns to scale. It is also known as the Lerner Index (LI) as it was first formalized by Abba Lerner in 1934. It has been applied as an indicator of competition in several papers since then. The most prominent contemporary research using the Lerner index includes Nickel (1996) and Aghion *et al.* (2005). This measure indicates the divergence from optimal allocation of resources. Stating differently, it indicates the profitability of the firms, which marks its ability to set its price over the marginal cost, thereby exhibiting the extent of market power it possesses. Therefore, in the presence of competition, prices gets reduced until the marginal cost; while in the absence of competition, firms can set their prices high such that they accrue a larger share of profit. However due to the unavailability of data on price and marginal cost, we have adopted the methodology used by Aghion *et al.* (2005), for measuring this indicator of competition. An elaboration of the calculation method can be found in chapter 3.

The LI exhibits several advantages over the market share concentration index, as the latter relies precisely on the geographic location and product markets. In the context of the output reallocation effect, the LI provides a more accurate inference than the concentration indices. Rojas (2011) asserted that it 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.

Nevertheless, the LI is also fraught with several theoretical problems. Based on the study by Stiglitz (1989), there can be an increase in profit per unit sales during the period of recession. This results in an increase in the value of LI (since LI can be

expressed as, [(Price-Marginal Cost)/Price)], reflecting a decrease in competition. On the contrary, an increase in competition is expected during recession. In essence, like the HI, another potential source of error in case of the LI is the reallocation effect. Another problem with the LI is that, with the decrease in firm's cost over time, PCM seems to increase, indicating a fall in competition. It is true that, with respect to firm's costs, a high PCM indicates greater market power and thereby less competition. However, conditional on price, a high PCM indicates efficiency. In the latter case, the indication of a low level of competition cannot be justified.

In addition, this framework, which is inspired by Hall (1988) in estimating PCM, rests on the strong assumption of constant returns to scale. This circumvents the need for estimating scale elasticities as these are unobservable from the data. While this measure has a simple computational advantage, it does not correspond to a more theoretical mark-up that needs to be adjusted for economies of scale. Therefore, we introduce another supplementary measure of PCM with adjustments for economies of scale, which we will refer to as the mark-up henceforward.

Price Cost Margin with adjustment for economies of scale (Mark-up)

This approximation of the PCM encapsulates the existence of scale and scope economies. According to Bikker and Van Leuvensteijn (2008), this measure rests on the idea that non-exhausted scale economies implies an incomplete utilization of the cost reduction potentiality. Being denoted by the ratio of output over multiple inputs times the scale elasticity, a higher mark-up is associated with an anti-competitive market. Papers like Badinger (2007) and Vancauteren (2012) have effectively employed this measure of mark-up to reflect competition.

But from a theoretical view-point, both these measures of PCM can have several loopholes. This has already been specified while discussing the Lerner index that the measure of profitability is not free from potential distortions in conceptualizing

competition. Similar to the HI, an increase in competition due to reallocation effect (because of more aggressive conduct by the firms) can raise market PCM, indicating incorrectly that there exists less competition. The reallocation effect can be partly eliminated by using the unweighted PCM, as in Aghion *et al.* (2005). But the disadvantage of the unweighted PCM is that, the PCM of small firms get a disproportionate effect on the industry aggregate PCM.

Profit Elasticity

Considering the drawbacks faced by the above mentioned competition indicators, we consider the profit elasticity (PE) as our final metric for competition. This novel approach was formulated by Boone (2000, 2008) and has subsequently been used in several empirical literatures. This measure rests on the concept that higher profit can be expected per unit fall in marginal cost in case of the efficient firms operating in a less competitive regime. Alternatively, in the presence of competition, the percentage fall in profit due to one percent increase in marginal cost will be of higher magnitude. Therefore, the absolute slope of the estimated coefficient for marginal cost determines the extent of competition prevailing in the market. However, owing to the lack of data on marginal cost, a ratio of variable cost and operating revenue has been used, which is in line with Boone (2008).

It can be mentioned in this context that, unlike the first three competition measures, the PE (also known as the Boone indicator) has been estimated using fixed effect panel regression on our firm-level data, over time. Again, a detailed estimation procedure can be found in chapter 3, along with the computation of the other competition measures.

6.3.3 Other firm-level determinants

The complexity of the relation between innovation and competition stems from the various factors that may influence the incentive to innovate, which includes the size and the age of firms, among others. The propensity to patent as well as the

patent productivity are affected by these firm characteristics, which we incorporate as the control variables in our model. By using the core variables obtained from the production statistics database of Statistics Netherlands, the explanatory variables are calculated.

In this chapter, we test whether the size and age of firms have a significant change in its effect on innovation output after the incorporation of competition measures in our model. As mentioned before, the logarithm of the number of employees are used as a proxy for firm size in our analysis. Empirical research suggests that the number of patents increases with firm size (Mairesse and Mohnen, 2002). Large firms have access to better financial resources and are more equipped for patenting activities, with their profound infrastructure and lesser asymmetric information. In addition, the huge patent litigation cost and the riskiness involved in the patenting process may act as a disincentive for the small firms.

Regarding the age of firms, it is measured as the difference between the entry year and the exit year of each of the firms for the successive years. We also include the age of firms as one of the control variables, owing to the fact that the life span of the incumbent firms might play an important role in determining the innovation performance in the pharmaceutical sector. The duration of their existence might throw some light on whether they accumulate the required tangible and intangible resources to engage in patenting.

Regulation indicator

To take into account the exogenous variations that affect competition, in order to control for the possible endogeneity bias, we consider the REGIMPACT indicator (as has been used in Braila *et al.*, 2010). This indicator addresses the regulatory conduct of the firms. It is attained from the OECD database and excels in its vast coverage of country and sector specifications. A brief description of this indicator is also incorporated in chapter 3.

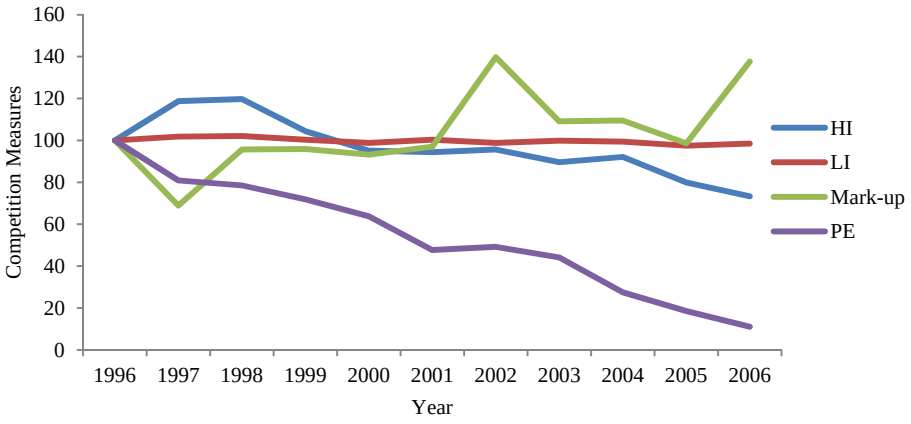
6.3.4 Data digression: A comparative approach

Competition measures

On dealing with the different alternative competition measures, the following diagram (Fig. 6.1) depicts their respective trends over the time span of 1996-2006 in the Dutch pharmaceutical industry. The four indicators are normalized to a base scale of 100 (base year 1996) for the sake of comparison. It should be noted that, for the LI (PCM without adjusting for the economies of scale) and the mark-up (PCM adjusting for the economies of scale), we have converted the firm level measure to industry level measure by calculating a simple average of individual PCM for each period of time.

In Fig. 6.1, competition does not exhibit a significant change over the span of 11 years (1996-2006) in case of the LI. But for the HI, it exhibits a decreasing trend, which implicates an increase in competition over time due to a declining level of concentration. The mark-up exhibits an overall increasing trend, exhibiting the highest level in the year 2002, and the lowest level in the year 1997. Thus, a general rise in mark-up reflects a decrease in competition with time. In line with the mark-up, the PE measure also shows a concomitant decreasing trend in competition. However the competition measured using PE exhibits a sharp decline over the specific time period. This drastic fall in profit elasticity may be due to the unobserved effects captured by the panel data.

Fig. 6.1: Trend in Competition Measures



Although both mark-up and PE indicate an overall decrease in competition over time, contrasting features are observed for HI index. However it is prudent to note that, the concentration measure indicates competition by the presence of more firms in the market due to the abatement of entry barriers, while the latter two measures quantifies competition by the change in conduct of the incumbent firms. Since a decrease in concentration might imply that more firms are entering the market, the reallocation effect can be a major reason in this context. A less aggressive conduct by firms (as denoted by mark-up and PE) might have prompted more firms to enter the market, resulting in a decrease in concentration. However the Lerner index (as in Aghion *et al.*, 2005) does not show any significant variation over our concerned time frame. Nevertheless, which indicator provides the best approximation measure cannot be deciphered yet, and hence it would be analyzed further in our empirical estimation.

Correlation Matrix

Table 6.1 represents a correlation matrix, where the Pearson's correlation coefficient for each variable is represented. Therefore, from this correlation table, we examine how the competition measures correlate to each other and to the innovation variables, along with the other determinants. It should be noted that the

values in parenthesis, under each correlation coefficient, is their corresponding p-values; with the asterisk denoting the coefficients having a p-value of 0.05 or lower.

Certain points are worth mentioning from the table. First, the LI and the mark-up show positive correlation coefficients between themselves, which is expected, as they belong to the same family of PCM. However, the competition measured using PE and mark-up again shows evidence of parity, since they bear significantly opposite sign with HI and other variables, as well as with each other. As the former is directly related to competition, while the latter is inversely related to the same, an opposite and significant correlation for these two indicators hints at a similarity between them in calibrating competition.

Table 6.1: Correlation Matrix

	HI	Log (Li)	Log (Mark-up)	PE	Citation Counts	Log (R&D Intensity)	Log(Firm Size)	Log (Age)
HI	1							
Log(LI)	0.025 [0.304]	1						
Log(Mark-up)	-0.070* [0.004]	0.145* [0.000]	1					
PE	0.821* [0.000]	0.041 [0.093]	-0.093* [0.000]	1				
Citation Counts	0.029* [0.029]	0.029 [0.227]	-0.123* [0.000]	0.030* [0.024]	1			
Log(R&D Intensity)	-0.008 [0.779]	-0.083* [0.021]	0.025 [0.482]	-0.034 [0.233]	0.131* [0.000]	1		
Log(Firm Size)	0.062* [0.000]	-0.060* [0.020]	-0.522* [0.000]	0.112* [0.000]	0.162* [0.000]	-0.152* [0.000]	1	
Log(Age)	0.010 [0.470]	-0.014 [0.569]	-0.221* [0.000]	0.002 [0.874]	0.076* [0.000]	-0.030 [0.285]	0.360* [0.000]	1

Second, although the correlation between HI (also LI) and citation counts indicates that more competition is associated with less innovation, mark-up and PE views contrary results. Concurrently, similar contrasting outcomes are obtained for log(R&D intensity), where HI and LI indicate more competition leading to greater R&D investment, while PE and mark-up infers the opposite. Third, the correlation between the log(R&D intensity) and citation counts is significant and positive, as it

is likely that higher R&D expenditure causes a greater innovation output. Fourth, the size of firms seems to have a negative and significant relation with R&D intensity while a positive and significant relation with citation counts, reflecting smaller sized firms to be more engaged in R&D activities while less in knowledge output. However, divergent albeit significant correlations are obtained for log(R&D intensity) and the various competition measures.

Finally, the firm age has insignificant correlation with most of the variables. But it remains significantly correlated with the citation counts and firm size, with a positive magnitude for both of them. This leads to the logical explanation that aged firms perform more patenting activities, and also the fact that older firms are also bigger ones. Considering the competition measures, firm age is significantly negative with mark-up, suggesting greater competition is synonymous with older firms.

In this context, it is worth mentioning that the correlation coefficient can have divergent implications with the regression estimates, as the former exhibits the strength of the mutual association between two variables, while the latter demonstrates the form of association between them such that one variable is predicted conditioned on its determinants.

6.4 Empirical methodology

In hindsight of the complexity involved in interpreting the causal effect of the competition measures, we attempt to empirically estimate the intrinsic competition-innovation relation using various econometric techniques. In an attempt to determine the best model specification to obtain robust estimation, we use three estimation strategies, namely, the zero-inflated negative binomial, the hurdle negative binomial-logit and the zero-accounting Poisson-pseudo maximum likelihood estimator, which have been detailed in the next subsection.

In general, citation counts are non-negative integers and hence the family of count data models are more applicable for this purpose. As has been discussed in chapter 5, the Poisson regression model is the core strategy that provides the groundwork for count data modeling. But in case of the Poisson model, the variance of the random variable is constrained to equal its mean. In this distribution, the random dependent variable $Y_{it} \in \{0,1,2,...\}$ has the probability distribution for observing a specific count, given x_{it}, which can be expressed as,

$$P(Y = y_{it} \mid \lambda_{it}) = \frac{\exp(-\lambda_{it})\lambda_{it}^{y}}{y_{it}!} \tag{6.1}$$

where $y = 0,1,2,...$ and $E[Y] = Var[Y] = \lambda_{it} > 0$

Hence the predicted rate is,

$$\lambda_t = E(Y_{it} \mid X_{it}) = \exp(X_{it}\beta) \tag{6.2}$$

But there exists a prominent restriction in the distribution of observed counts as it considers the equality of conditional mean and conditional distribution of the dependent variable. Hence, researchers typically employ a more general specification, like the negative binomial model. The negative binomial model accounts for the overdispersion in the data, whereby the conditional variance exceeds the conditional mean. Overdispersion may occur due to positive contagion or unobserved heterogeneity (Winkelmann and Zimmermann, 1995 & Licht and Zoz, 1996).

The negative binomial (NB) model is an extension of the Poisson model, where the unobserved heterogeneity is taken into account, by adding a random component to λ_{it}. Considering,

$$\tilde{\lambda}_{it} = \exp(X_{it}'\beta + \varepsilon_{it}) = \exp(X_{it}'\beta)u_{it} \tag{6.3}$$

where ε_{it} captures the unobserved heterogeneity of the firms. This causes a stochastic mean function with an expectation of $E[\tilde{\lambda}_{it}] = \lambda_{it}$ and the variance is $Var[\tilde{\lambda}_{it}] = \lambda_{it}^2 \sigma_m^2$.

Thus, in case of the negative binomial model, the predicted rate is the same as equation (6.2), but there exists a negative binomial distribution for the predicted probabilities. Since the variance of the predicted counts is increased by the addition of a single parameter (say γ_{it}), the predicted probabilities can be written as,

$$P(Y = y_{it} \mid x_{it}) = \frac{\Gamma(y_{it} + \alpha_{it}^{-1})}{y_{it}!\Gamma(\alpha_{it}^{-1})} \left(\frac{\alpha_{it}^{-1}}{\alpha_{it}^{-1} + \lambda_{i_t}} \right)^{\alpha^{-1}} \left(\frac{\lambda_{i_t}}{\alpha^{-1} + \lambda_{it}} \right)^{y} \quad \text{for } y = 0,1,2,... \qquad (6.4)$$

On considering the basic outline for count data analysis, the following subsections articulate on the methodological attributes of the zero-accounting models that we have used in this chapter. The reason behind using zero-accounting models is the existence of a large number of zero observations in the non-negative citation-weighted patent data. Therefore, we aim at using the models that help to deal with the "excess zeroes" in our dependent variables.

6.4.1 Zero inflated Poisson or Negative Binomial model

A description of this methodology has already been taken up in chapter 5. However, to synchronize this econometric technique with the other methods applied, we review its empirical framework in this chapter as well. Since many firms do not participate in patenting, it leads to many zero observations for citations counts. Hence, a zero inflated model is likely to be more applicable for our analysis, which permits the mechanism generating zero observations to be different from the positive observations.

The zero inflated count models reflect unobserved discrete heterogeneity in order to differentiate between those firms which will "always" have zero counts to the

ones which are “at the risk” of zero counts. The zero inflated Poisson regression model or the zero inflated negative binomial model, with a binary logit or probit, are the most prominent zero inflated count models used. In regard to its estimation technique, the inflation equation explains whether the probability that the count is logit or probit. The idea of inflation is used for defining those firms which are having “always zero” observations. In a more specific way, this class of the observations have inflate =1. Therefore, the probability of existing in this class of observations is:

$$P(always\ 0|x_{it}) = P(inflate = 1|x_{it}) = F(z_{it}\alpha) = \rho_{it} \tag{6.5}$$

In the above equation, F is the cumulative density function (cdf) for the logistic if logit is used or the cdf for the normal if probit is used for the binary model.

For calculating the probability of observing a particular count, the results that are obtained from the count equation needs to be adjusted in accordance to the probability of the observations that are in the “always zero” category. Hence, in case of the Poisson equation,

$$\begin{aligned} P(y_{it} = 0|x_{it}) &= P(always\ 0)\ +\ P(0\ by\ chance) \\ &= \rho_{it} + (1 - \rho_{it})e^{-\lambda_{it}} \end{aligned} \tag{6.6}$$

In case of the non-zero counts, the probability for observing a specific count of the Poisson distribution is,

$$P(Y = y_{it}|\lambda_{it}) = (1 - \rho_{it})\frac{exp(-\lambda_{it})\lambda_{it}^{y}}{y_{it}!} \tag{6.7}$$

Similar is the econometric explanation for the zero-inflated modeling in case of the application of negative binomial. However, we consider the zero-inflated negative binomial (ZINB) over zero-inflated Poisson (ZIP), since the former account for overdispersion.

6.4.2 Hurdle model

As an alternative count data modeling technique, we have also applied the Hurdle model, as proposed by Mullahy (1986). It may be worth noting that, although zero inflated and hurdle models are used to deal with large number of zeroes in the observed data, there lies a prominent difference in how they interpret the zeroes. While the zero inflated models consider that the zeroes occur both from structural and sampling origins; the Hurdle model assumes that the zeroes are from a structural source only. Although this difference might seem to be subtle, one model may be more appropriate than the other based on the observed sample, and might lead to prominent differences in results.

A Hurdle model reflects a two-stage decision making process, where each part being a model of a single decision. Technically, the Hurdle model specification is based on the assumption that there exists two sets of parameters that drives the data generating process. The inherent concept of the binomial probability model is that, it governs the binary outcome of whether a count variate has a zero or a positive value. As pointed out by Cameron and Trivedi (1998), it is a modified count model where the two processes generating the zeroes and the positives are not constrained to be the same. Once the Hurdle is crossed, and positive counts is observed, the data generating process is then controlled by the truncated-at-zero count model.

Another important aspect of this two-stage estimation procedure is that, the two parts are functionally independent of each other. Hence, the maximum likelihood estimation of the model can be achieved by separately maximizing the two independent terms in the likelihood model, where one corresponds to zeroes and the other corresponds to the positive values. Therefore, the first part uses the full observation sample and the second part uses only the positive count observations (Cameron and Trivedi, 2009).

In equation (6.8), we specify a two-part data generating process in the context of a panel data. A Bernoulli distribution determines the split between zero and positive values of the dependent variable, having probabilities of $f_1(0 \mid v_{it})$ and $1 - f_1(0 \mid v_{it})$ respectively. A truncated at zero variant of the count specification determines the distribution of the positives, which is given by $f_2(y_{it} \mid x_{it})$. This can be modeled as follows:

$$f(y_{it} \mid x_{it}, v_{it}) = \begin{cases} f_1(0 \mid v_{it}) & if \ y_{it} = 0 \\ (1 - f_1(0 \mid v_{it})) \dfrac{f_2(y_{it} \mid x_{it})}{1 - f_2(0 \mid x_{it})} & if \ y_{it} \geq 1 \end{cases} \tag{6.8}$$

The above equation reduces to the standard model only if $f_1(0 \mid v_{it}) = f_2(0 \mid x_{it})$ and $v_{it} = x_{it}$. Although in principle, v_{it} and x_{it} may have distinct or mutually overlapping values, in practice, they are generally the same.

As proposed by Cameron and Trivedi (2013), this model based on panel data is a pooled version of hurdle model for cross sectional data. The first step in its application procedure involves binary logit specification and subsequently the second step involves Poisson or negative binomial specification. Similar to our prior applications, we use the negative binomial specification for the second stage, due to the aforementioned reason of considering overdispersion.

However, one major drawback of this estimation technique is the implementation of the individual specific effects, as it is not possible anymore to split the likelihood into two parts that can be maximized individually. Also, the individual specific effects in the two distributions should not necessarily be independent. In addition, fixed effect variant of this model is not plausible in a short panel framework. In a longer time frame, individual specific dummy variables can be used.

6.4.3 Poisson Pseudo Maximum Likelihood model

We finally apply the Poisson Pseudo Maximum Likelihood (PPML) method as our third model specification. Although this method is generally used to estimate the gravity equation, we implement this technique to evaluate the effect of competition and other determinants on the citation-weighted patents of our concerned database dealing with the Dutch pharmaceutical industry. Although this estimation strategy has not found its niche outside trade related research yet, Silva and Tenreyro (2006) has affirmed that this technique can be used in a broad range of economic applications where the equations are log-linearized or transformed to a non-linear specification.

As has been stated before, the standard Poisson model does not account for the problem of overdispersion. To this effect, the PPML is consistent for overdispersion as well as under dispersion. Hence, this estimation technique is efficient when the conditional mean and conditional variance are proportional, and not necessarily equal to each other. In Silva and Tenreyro (2011), it is also proved that the estimation technique can perform very well even when there are excess number of zeroes in the model. Therefore, we apply this method to our dataset, using the empirical methodology as in Gourieroux *et al.* (1984). For the sake of brevity we do not incorporate the methodological details here.

But before exploring our empirical results, we consider a brief mentioning of the various empirical issues that we have taken into account, while modeling our estimations in the following subsections.

6.4.4 Controlling endogeneity

The most important problem in analyzing the effect of competition on innovation output of the firms, is the problem of reverse causality or endogeneity between them. It is very difficult to find relevant instruments that is correlated with competition, but uncorrelated with innovation. Hence we use one year lagged

values of the endogenous variable (that is, competition measures) as the possible instrument to combat the problem of endogeneity. Also, the use of lagged variables is pertinent in our data as there are some noticeable variation in the competition measure in the 11 years' time span that we have taken into account.

However, as an additional control for tackling endogeneity issues, we have also introduced a regulation indicator (REGIMPACT), obtained from the OECD database. This indicator has been improvised by Conway and Nicoletti (2006) that gauges the indirect effect of non-manufacturing industries on all industries of the economy. The main reason for adopting this indicator over the other regulation indicators is due to its vast coverage of industries, specific to different countries and years.[7]

6.4.5 Controlling unobserved heterogeneity

In analogy with chapter 4 and 5, we have also applied the maximum likelihood (ML) approach (following Wooldridge, 2005) for handling the unobserved heterogeneity problem in the empirical estimations of this chapter. This approach for handling the individual effects, conditioned on the initial values and the within-means of the time-variant independent variables, has been incorporated in the zero-inflated negative binomial and the negative binomial logit hurdle estimation analyses. Also, it is important to note that, we have applied the specification by Hesketh and Skrondal (2013), where the initial period explanatory variable is omitted from the within means, in order to prevent biased estimations. We do not restate its methodological details in this chapter and refer to subsection 5.4.1 for its elaboration.

[7] However, the regulation indicator used in our empirical estimations has not been considered in its reduced form (as in Aghion *et al.*, 2005). Hence, for future investigation on controlling for endogeneity, the regulation indicator can be incorporated as an instrumental variable (IV) in a structural model approach, where the predicted residual of the IV is applied in the final regression analysis.

6.5 Empirical estimation

For our empirical analysis, we investigate on the effective impact of competition on innovation output by using three distinct empirical frameworks in the context of count data. The single parameter Poisson distribution is not very appropriate in count data modeling if the mean-variance equality fails to hold. Hence, heterogeneity due to overdispersion might be neglected in this case (Cameron and Trivedi, 1986). Due to the excess number of zeroes in our dependent variable, the zero inflated count models are supposed to be more appropriate than the negative binomial model, which is further confirmed by the Vuong test. Using the Vuong command in Stata, a significant z-test in all the cases indicates that zero-inflated models are more appropriate to our dataset. In addition, we use the 'count-fit' command (by Long and Freese, 2006) and find that the ZINB model is strongly preferred over the ZIP and negative binomial models for our given dataset. Hence, we first perform estimations using the ZINB model. In addition, as an alternative to the ZINB model, we further use the Hurdle model of counts. Finally we introduce the PPML estimation technique in the analyses of innovation and competition, as a robustness check to the former estimations.

In all the three modeling strategies, citation-weighted patents are used as the response variable. Each of the modeling techniques consist of four regression analysis, which are distinguished by the competition measure used, i.e., in the order of the Herfindahl concentration index, the Lerner index, the elasticity adjusted mark-up and the profit elasticity. Quadratic specification is applied for all the competition variables in order to ensure if there exists any non-linearity in the relation between competition and innovation output. In addition, it should be noted that we have assumed random effects for all the regression analyses using ZINB and Hurdle, following Wooldridge (2005). But random effect could not be used for PPML as it is beyond the scope of this estimation strategy. However a cluster-robust standard error specification is used for the PPML estimations, in the context of our panel data.

On arriving at the best fit model from the three estimation strategies, we apply the most robust model for continuing our analyses further. The additional analysis that is supplemented with our obtained result includes the incorporation of regulation indicator as an additional instrument to control for endogeneity, and considering a modified dataset where those firms which are closer to the technological frontier are considered. The next subsections elaborate on our empirical results.

6.5.1 Estimations using the zero-inflated negative binomial model

In the ZINB model, the first equation is a negative binomial count model which is a predictor of counts while the second is a logit model which is a predictor of zeroes. The latter regression specifies the variables that determine the probability that the count is logit. The logit is the default command used, but the probit can also be an option. In our analysis, we have kept the same set of regressors in both parts of the model. However, in table 6.2 we represent only the negative binomial part, omitting the logit analysis. This is mainly because, we intend to focus on the results of the first part of the model, as the latter part is merely the prediction of the excessive zeroes.

From the results obtained, it is found that the fitted R&D expenditure is positive and highly significant, except in the third regression where mark-up is used to measure competition. A significantly positive effect of R&D investment is obvious since innovation capacity of firms involves a concerted mechanism of channelizing investments in innovation to effective innovation output. However, a consistently lower magnitudes of the coefficients bring to notice the fact that, all R&D activities does not lead to patenting, since the whole process of drug innovation involves huge risks and innumerable failures. Also, all innovations are not accounted for by patents, since firms may evade patenting strategy owing to the cost and time involved.

Table 6.2: Zero-inflated negative binomial

Dependent Variable: Citation-weighted patents				
	1	**2**	**3**	**4**
Log(R&D fitted)	0.244***	0.236**	0.137	0.233***
	[0.091]	[0.107]	[0.097]	[0.086]
Lagged HHI	0.069***			
	[0.026]			
Lagged HHI²	-0.00007**			
	[0.00003]			
Lagged LI		2.412**		
		[1.132]		
Lagged LI²		0.414*		
		[0.249]		
Lagged Markup			2.852***	
			[0.707]	
Lagged Markup²			-0.931***	
			[0.303]	
Lagged PE				-2.252**
				[0.911]
Lagged PE²				0.338
				[0.342]
Log(Age)	-0.07100	-0.048	0.177	0.180
	[0.248]	[0.296]	[0.356]	[0.220]
Log(Employees)	0.338**	0.334***	0.028	0.287***
	[0.142]	[0.123]	[0.245]	[0.109]
Patent dummy	0.773***	0.016	2.270***	2.176***
	[0.298]	[0.328]	[0.388]	[0.298]
Intercept	-14.133***	3.538***	-2.238***	-2.174**
	[5.444]	[1.184]	[0.797]	[0.983]
Random Effect		YES		
Initial(Citation Count)	0.004***	0.007***	0.006***	0.005***
	[0.001]	[0.001]	[0.001]	[0.001]
Log Likelihood	-829.083	-443.549	-542.024	-882.445
Akaike Information Criterion(AIC)	1700.165	931.097	1122.049	1802.890
Bayesian Information Criterion (BIC)	-442.158	-289.031	-226.902	-349.607

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

For the competition indicators, non-linearity is introduced using a quadratic form. The HI and mark-up exhibit a significant inverted U-relation with citation counts. Since these two indicators are inversely related to competition, they therefore indicate a U-shaped relation between competition and knowledge output. This may signify that competition decreases with initial progress in knowledge output, reaches a minimal point, after which it tends to increase with innovation output. This serves as a contrary argument for the study by Aghion *et al.* (2005). In line with these finding, the PE also reflects a U-shaped relation between competition

and innovation output. However the squared coefficient of this direct indicator for competition remains insignificant.

But the LI depicts no sign of non-linearity, as it exhibits a persistent and significantly negative effect of competition on knowledge output. Altogether, the first-order coefficients for the competition indicators implicates that, firms are less prone to engage in patenting amidst increased competition. At the same time, evidence of an increased competition bolstering innovation output in the latter phase of the innovation process is also marked by our estimation results.

However, the total marginal effect, taking into account the complete ZINB estimations, confirms that the marginal effect obtained as the sum of the effects of the competition indicators and their squared terms yields a positive slope between the indirect competition measures (viz. HHI, Lerner index and mark-up) and the citation-weighted patents. Simultaneously, the change in PE per unit change in citation counts is found to be negative. Table 6.3 represents the total marginal effects of the competition measures based on the regressions in table 6.2.

Table 6.3: Total marginal effect from the ZINB estimations

	HHI	Log(Li)	Log(markup)	PE
Citation-weighted patents	0.012	0.443	2.120	-3.151
	[0.002]	[0.216]	[0.535]	[1.807]

The marginal effects and their corresponding standard errors reflect that the derivative effect of competition on innovation is significantly negative, taking into consideration the quadratic specification of the competition measures. Hence, the U-shaped relation between competition and innovation cannot be completely established from the economics perspective. To further check on the non-linear relation between competition and innovation, a graph for total marginal effect may be used to investigate at what level the competition becomes minimum in the U-shaped curve.

Regarding the other control variables, the effect of age on citation-weighted patents is found to be insignificant with negative coefficients in regressions 1 and 2; while positive coefficient is observed for regressions 3 and 4. Nevertheless, consistently insignificant results for firm age leads us to infer that there is no prominent effect of firm age on innovation performance in case of the Dutch pharmaceutical firms. But the effect of size of the firms is found to be consistently positive and mostly significant. This reflects that bigger sized firms have greater potential and incentives to innovate. Finally, significant evidence of a persistence in the innovation output is obtained from the coefficient for lagged patent dummy. Also, the initial dependent variable, which is incorporated for the treatment of biased initial observations, exhibits positive magnitude with high level of significance.

6.5.2 Estimations using Hurdle model

As an alternative for the ZINB estimation strategy, we use the Hurdle model with negative binomial and logit specifications. This regression method is also a two part model estimation, where the first part of the table represents the zero truncated negative binomial, that uses the sample when there exists positive citation-weighted patent counts and the second part is a logit regression that uses the full sample of the data. Both the parts are functionally independent from each other, and provide a two-step decision making process. Hence, in the regression result, a positive coefficient of an independent variable means that the variable has an increasing effect on the citation-weighted patent, conditional on a positive count. On the other hand, in the logit part of the model, a positive coefficient means the concerned regressor increases the probability of a positive observation. From the results obtained, it is found that the independent variables affect the outcome mostly in the same direction. Nevertheless, since the zero-truncated negative binomial regression is the estimation of interest, we exhibit its results in table 6.4.

Table 6.4: Hurdle Negative Binomial Logit

Dependent Variable: Citation-weighted patents				
	1	**2**	**3**	**4**
Log(R&D fitted)	0.205*	0.183*	0.157*	0.246**
	[0.105]	[0.109]	[0.092]	[0.100]
Lagged HHI	0.072**			
	[0.030]			
Lagged HHI²	-0.00007**			
	[0.00004]			
Lagged LI		2.128*		
		[1.131]		
Lagged LI²		0.345		
		[0.247]		
Lagged Markup			1.556***	
			[0.602]	
Lagged Markup²			-0.564	
			[0.377]	
Lagged PE				-1.809*
				[1.004]
Lagged PE²				-0.247
				[0.368]
Log(Age)	-0.295	-0.246	-0.250	-0.574*
	[0.352]	[0.359]	[0.311]	[0.321]
Log(Employment)	0.366**	0.373***	-0.194	0.324*
	[0.169]	[0.132]	[0.216]	[0.173]
Patent dummy	0.694**	0.041	0.610	0.907***
	[0.328]	[0.359]	[0.382]	[0.317]
Intercept	-14.663**	3.914***	-0.643	-1.411
	[6.296]	[1.194]	[0.771]	[1.073]
Random Effect		YES		
Initial(Citation Count)	0.005***	0.007***	0.005***	0.005***
	[0.001]	[0.001]	[0.0009]	[0.001]
Log Likelihood	-837.584	-450.134	-516.332	-837.531
Akaike Information Criterio (AIC)	1717.168	946.268	1074.663	1713.062
Bayesian Information Criterion (BIC)	1823.99	1051.939	1172.482	1809.71

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

From the results obtained, the coefficient for the fitted value of log R&D is found to be positive in all the regression models, with a 5% and 10% significance level. The coefficient for the age of firms is repeatedly insignificant and negative and hence, its effect on innovation output is indeterminate. But a slight exception is noticed for the last regression, where the coefficient for age is ascertained with a 10% significance level. On the other hand, a positive and significant effect of firm size is observed again in most of the cases (barring the third regression), which substantiates our previous finding that bigger firms perform more output in innovation. The more consistent effect of firm size establishes it to be an essential

determinant for a firm's decision to patent, while age may not have a preponderant role in this regard.

Plus, patent dummy is consistently positive with a significant coefficient for the first and the fourth regression analyses, thereby hinting at a persistent dynamics in terms of the innovation output in the Dutch pharmaceutical industry.

Discussing on our primary regressor, that is, the competition measures, a prominent and significant inverted U-shaped curve is observed again when using the HI and the elasticity adjusted mark-up. This suggests an U-shaped relation between competition and innovation, as observed in our ZINB estimations. But the Lerner index does not show any sign of non-linearity and points at a negative relation between competition and innovation output. Since the profit elasticity is a direct measure of competition, its negative and significant coefficient also signifies that innovation performance in the Dutch pharmaceuticals falls with the rise in competition. However, no sign of non-linearity is noticed in this specification using PE.

Further, the total marginal effect for the HHI and markup indicate a positive slope with the citation-weighted patents. Therefore, in analogy with table 6.2, it is established that the change in competition per unit change in innovation performance is negative. Similar outcomes are witnessed for LI and PE. Hence, the U-shaped relation cannot be fully justified, and calls for further analysis in this regard.

An overall inspection of the two estimation techniques (as discussed in table 6.2 and 6.3) suggests that the estimation results bear concordance, with similarity in sign and significance of the coefficients. This suggests that the two models are analogous to some extent. However the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), being asymptotically efficient model selection tests, have been used in our analysis to find whether the ZINB model or

the Hurdle model is a better fit. In addition the log-likelihood is also an essential metric regarding the choice of model selection. Although we do not arrive at a definitive inference for our model preference from this formal selection specifications, we discuss on this in greater details in subsection 6.5.4.

6.5.3 Estimations using the Poisson Pseudo Maximum Likelihood method

Table 6.5 reiterates on the contemporaneous effect of competition measures and other control variables on the citation-weighted patents, by using the PPML method. However this method did not support the Wooldridge specification to control for unobserved heterogeneity. Hence, a cluster model specification is applied for our panel data.

Table 6.5: Poisson Pseudo Maximum Likelihood

Dependent Variable: Citation-weighted patents				
	1	**2**	**3**	**4**
Log(R&D fitted)	0.652***	0.991***	0.918***	0.636***
	[0.151]	[0.142]	[0.107]	[0.143]
Lagged HHI	0.054*			
	[0.028]			
Lagged HHI²	-0.00005*			
	[0.00003]			
Lagged LI		1.909***		
		[0.674]		
Lagged LI²		0.184***		
		[0.069]		
Lagged Markup			-0.585	
			[1.839]	
Lagged Markup²			0.343	
			[0.372]	
Lagged PE				-1.623**
				[0.776]
Lagged PE²				-0.194
				[0.250]
Log(Age)	-0.362**	-0.326	0.099	-0.479**
	[0.169]	[0.218]	[0.134]	[0.195]
Log(Employees)	0.585***	0.737***	0.486**	0.434***
	[0.130]	[0.125]	[0.203]	[0.131]
Patent dummy	3.497***	4.076***	3.943***	3.566***
	[0.469]	[0.686]	[0.659]	[0.474]
Intercept	-18.288***	-4.958**	-7.100***	-7.131***
	[5.198]	[2.374]	[1.497]	[0.999]
R square	0.657	0.821	0.855	0.695
Pseudo Log Likelihood	-5025.775	-2117.030	-2600.466	-4802.911
Akaike Information Criterion(AIC)	10065.550	4248.060	5214.931	9621.821
Bayesian Information Criterion(BIC)	10101.160	4280.221	5247.537	9662.515

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

From the results obtained, the coefficient for the fitted log of R&D expenditure is found to be positive and highly significant, thereby confirming our previous results. Age is systematically negative in all the regression models. However, in regression 1 and 4, significant negative outcome for age is found. This corroborates to some extent that the young firms are the ones that engage more in innovation performance. However it maintains our previous finding that, firm size has a significantly positive effect on citation counts. Additionally, highly significant innovation persistence is observed for the coefficients of lagged patent dummy, again confirming our previous results.

Concerning the competition measures, it is found that the indicators bear parity with the previous estimations. While, the Herfindahl concentration index and markup again reflects a U-shaped link between competition and innovation, the coefficients for the former is only found to be significant. The LI is consistent with its positive and significant coefficients, thereby portraying a negative effect of competition on citation counts. Similar is the case with the PE index, with a significant negative effect being observed for the first order coefficient. The estimations in all the regression models consistently maintained a negative impact of competition on innovation, although the observed non-linearity needs to be further investigated.

On the whole, most of the coefficient estimates of the PPML estimation provide comparable results with ZINB and the Hurdle model. This undoubtedly suggests our empirical results to be robust. However to continue our analysis further, we attempt to choose the best model by evaluating the statistical tests, empirical results and the anecdotal evidences. This is elucidated in the next subsection.

6.5.4 Choice of model

According to the AIC and BIC assessment for model selection, a lower value indicates a better fit model. On the other hand, since the log-likelihood is always

negative, its higher value (closer to zero) will lead to the better model choice. Basing our choice criteria to these statistical tests, the PPML estimator can be outrightly omitted since it does not score well based on this criteria. Nevertheless, the R^2 value of the PPML estimations is considerably high (all the regressions using PPML is quoted having $R^2>0.5$), and hence we cannot completely infer if these estimation technique is inappropriate for our analysis. In addition, its comparable estimation coefficients with the ZINB and Hurdle model, ascertains that the PPML cannot be completely neglected. It should also be noted in this context that the AIC and BIC examines only the relative quality of the models, and are incapable to draw conclusions on the relevance of individual models.

Nevertheless, on the basis of these formal empirical tests, PPML will not be carried forward in our further analysis. Regarding the ZINB and the Hurdle model, our choice of best model specification is not definitive. As it can be seen from table 6.2 and 6.3, the AIC and log-likelihood favour the ZINB over the Hurdle model in case of the first two regression analysis. On the contrary, the Hurdle model performs fairly better in the mentioned statistical tests when the last two regression models are considered. However the BIC statistics considers the ZINB model to be a better fit that the Hurdle. This findings tally with several past literatures which concluded that the different statistics for the goodness of fit does not lead to the same conclusion. Therefore, for our final choice of model, it is implausible to base our conclusions on these statistical tests only. Following Tran *et al.* (2013), we also take into consideration the magnitude and significance of the estimated coefficients, economic perspectives and past empirical literatures for our choice of final model specification. In this regard, we consider the ZINB estimator for our further analysis.

6.5.5 Incorporating regulation indicator

One of the major issue is assessing the effect of competition on innovation is the fact that, innovation can also in turn affect the competitive conduct by the firms.

This causal relationship is likely to affect the estimated results. As a possible control to this endogeneity problem, we have already included the lag values of the competition measures. However the lags of the endogenous variables may not turn out to be a valid instrument. Hence in this subsection, we consider an additional control to address this endogeneity issue. Similar to Olley and Pakes (1996), we have applied a regulation indicator for the identification of an exogenous variation in competition. Therefore, in table 6.6, we present the ZINB estimation results with the incorporation of the REGIMPACT indicator.

Table 6.6: Zero-inflated negative binomial with additional instrument

Dependent Variable: Citation-weighted patents				
	1	**2**	**3**	**4**
Log(R&D fitted)	0.281***	0.282***	0.223**	0.252***
	[0.088]	[0.108]	[0.106]	[0.084]
Lagged HHI	0.173***			
	[0.042]			
Lagged HHI²	-0.0002***			
	[0.00006]			
Lagged LI		2.833**		
		[1.105]		
Lagged LI²		0.528**		
		[0.250]		
Lagged Markup			2.801***	
			[0.701]	
Lagged Markup²			-0.915***	
			[0.321]	
Lagged PE				-0.855
				[1.027]
Lagged PE²				0.350
				[0.497]
Log(Age)	-0.070	-0.010	0.346	-0.003
	[0.236]	[0.271]	[0.268]	[0.234]
Log(Employees)	0.387***	0.454***	0.551***	0.306***
	[0.143]	[0.126]	[0.157]	[0.143]
Regulation indicator	468.581***	74.883***	84.459**	116.014
	[143.423]	[26.125]	[34.232]	[82.142]
Patent dummy	0.924***	-0.008	2.541***	0.965***
	[0.295]	[0.317]	[0.351]	[0.296]
Intercept	-54.612***	-0.603	-6.889***	5.378
	[13.745]	[1.807]	[2.229]	[4.457]
Random Effect		YES		
Initial(Citation Count)	0.004***	0.006***	0.006***	0.004***
	[0.001]	[0.001]	[0.001]	[0.001]
Log Likelihood	-824.019	-437.836	-541.265	-823.771
Akaike Information Criterion(AIC)	1694.038	923.672	1124.531	1691.541
Bayesian Information Criterion (BIC)	-438.112	-287.267	-215.104	-445.695

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

The statistical tests suggest a marginal improvement in the estimations, with the addition of the REGIMPACT measure to control for endogeneity. However, a positive and mostly significant impact of the regulation indicator does suggest that patenting practices flourish in a closed environment under regulatory restrictions. However this result on the regulation indicator is merely tentative as we do not deploy any other regulatory measure to highlight its impact on the patenting phenomenon undertaken by these Dutch industries.

Considering the rest of the explanatory variables, it is seen that the estimated coefficients bear striking similarity with the ones in table 6.2, which confirms a consistency in our empirical results. However, the PE measure loses its significance with the incorporation of the REGIMPACT measure. Concurrently, the regulation measure is also found to be insignificant in this estimation using PE. Nonetheless, their coefficients bear the same sign keeping parity with the other regressions. Anyhow, the competition measures are concordant to the past results, suggesting that competition can be a necessary hindrance for innovation performance in the Netherlands' pharmaceutical sector. Altogether, the bottom line to these results indicates that the large R&D intensive pharmaceuticals, with a greater past patenting propensity are more likely to perform patenting activities, under a closed regime.

6.5.6 Distance to the technological frontier

In an attempt to verify how competition influences the innovation activity of those pharmaceuticals that are closer to the technological frontier within the industry, we have considered a subsample of our database which comprises of those firms which have a closer proximity to this frontier.

We closely follow Amable *et al.* (2008) for calculating the technological frontier from our given dataset. The labour productivity for each firm reflects their individual efficiency, which is computed as the ratio of value-added to the number

of employees. Subsequently, we consider the firm having the highest productivity in a particular year (within our reviewed time frame), as the one holding highest technological capability. On the basis of the technological frontier, so obtained, we evaluate the closeness to the frontier for individual firms, at each time period, as the ratio of their labour productivity over the technological frontier at that year.

From this, the mean value of the technological capacity was marked at 0.042 and those firms which possessed a frontier value greater than or equal to it at any of the given years were considered as the technological leaders. Likewise, those firms which exhibited a lower value were contemplated as the technological followers. Based on this demarcation, 238 pharmaceuticals were found to have a technological edge over their 282 counterparts.

However to capture the escape competition effect fully, we were unable to base our analysis for the technological followers. The reason behind this omission is the inability to obtain successfully converged and consistent estimations with most of our competition measures when the subsample of 282 follower firms are considered.[8] This is mainly due to the fact that these firms also possess many missing values that has constituted to very less observations.

Nevertheless, consistent estimation results could be obtained when we consider the firms closer to the technological frontier. Consequently, our analysis mainly brings to light how the technological leaders react to the competition changes for their innovation performance. Table 6.7 therefore explains our empirical outcome.

[8] With forced iterations, convergence could be achieved for the HI and PE in case of the subsample of technological followers. However the results were systematically similar to our past estimations, as well as the estimations using the technological leader firms. To mention, the HHI exhibited a U-shaped relation between competition and innovation, with insignificant coefficient at the squared term. On the other hand, the PE showed a negative relation between the competition-innovation phenomena.

Table 6.7: Zero-inflated negative binomial using a subsample of firms closer to the technological frontier

Dependent Variable: Citation-weighted patents				
	1	**2**	**3**	**4**
Log(R&D fitted)	0.237***	0.313***	0.226**	0.224***
	[0.090]	[0.109]	[0.111]	[0.086]
Lagged HHI	0.127***			
	[0.044]			
Lagged HHI²	-0.0002***			
	[0.00007]			
Lagged LI		3.349***		
		[1.143]		
Lagged LI²		0.643**		
		[0.259]		
Lagged Markup			2.845***	
			[0.712]	
Lagged Markup²			-0.937***	
			[0.328]	
Lagged PE				-0.447
				[1.161]
Lagged PE²				0.317
				[0.554]
Log(Age)	-0.222	0.016	0.369	-0.182
	[0.335]	[0.276]	[0.291]	[0.348]
Log(Employees)	0.345**	0.486***	0.533***	0.301**
	[0.138]	[0.127]	[0.170]	[0.136]
Regulation indicator	363.574**	71.056***	83.089**	56.999
	[149.308]	[26.261]	[35.081]	[90.156]
Patent dummy	1.471***	0.008	2.594***	1.483***
	[0.343]	[0.324]	[0.353]	[0.353]
Intercept	-41.772***	-0.186	-6.899***	2.114
	[14.255]	[1.832]	[2.276]	[4.956]
Random Effect		YES		
Initial(Citation Count)	0.004***	0.006***	0.006***	0.004***
	[0.001]	[0.001]	[0.001]	[0.001]
Log Likelihood	-668.364	-430.284	-529.437	-669.472
Akaike Information Criterion(AIC)	1382.728	908.568	1100.873	1382.945
Bayesian Information Criterion (BIC)	-350.513	-279.614	-200.28	-355.235

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

The coefficients for most of the explanatory variables bear close resemblance to those obtained in our prior estimations. Persistently positive and significant coefficient is observed for both fitted R&D intensity and firm size, confirming our previous findings. Persistence in innovation output is also substantiated by our regression results (except in regression 2). Regarding the regulation indicator, a positive and mostly significant impact on the innovation output is found, bearing conformity to the results obtained in table 6.6. The competition measures displayed similar outcomes as well. However PE is again found to be insignificant (which was also observed in table 6.6), asserting the fact that the estimation results using

PE is not as robust when regulation indicator is applied. In this regard, the mark-up is found to provide potentially better estimated coefficients, especially after the incorporation of the REGIMPACT indicator. In any case, the estimated competition measures ratifies our previous findings on the plausible existence of an U-shaped or a negative relation between competition and citation-weighted patents. Furthermore, the log-likelihood, AIC and BIC definitely reveals an improvement in the empirical model.

By and large, it may be inferred that the technological leaders also deals with a slackening of patenting performance on encountering competition in the market. Hence, they do not show any variance in their behavior in relation to the entire corpus of pharmaceutical firms in the Netherlands.

6.6 Conclusion

This chapter conceptualizes the different competition measures from existing literatures and uses them to investigate the palpable relation between competition and innovation performance of the Netherlands' pharmaceutical industry. A descriptive exploration of the different competition measures provides cursory evidence of less aggressive conduct by the existing pharmaceuticals that has led to the influx of more firms in the market. This decrease in concentration, backs the existing pharmaceutical statistics that corroborates to a high inflow of generic firms in the Dutch pharmaceutical market.

To sum up our empirical study, a comparative investigation using three alternate econometric models, in the context of a panel data, portrays almost similar results on the effect of different regressors for predicting the innovation output of the Netherlands' pharmaceutical firms. The regression results suggest a significant U-shaped relation between competition and innovation when HI and the elasticity adjusted mark-up is used. But the LI reports a negative relation between them, and exhibits no signs of non-linearity. In conjunction with the HI and mark-up, PE also reflects a U-shaped relation between competition and innovation, albeit

insignificant second order coefficient terms. However the PE loses its significance with the incorporation of the regulation indicator.

On the whole, our empirical findings suggest a potentially negative impact of competition on citation counts, which may have the possibility to divert to a positive association at the later stage of the innovation process. Nevertheless, the existence of non-linearity cannot be ascertained and needs further exploration. Furthermore, the total marginal effect confirms that the change in competition per unit change in innovation performance is negative. Therefore, although a negative relation between competition and innovation is consistently confirmed, the presence of non-linearity between them may not exist.

This outcome remains consistent with the technological leaders as well. Therefore, our estimation results indicate that the finding of Aghion *et al.* (2005) is not foolproof. However, our finding is comparable to many existing studies that addresses the negative linkage between competition and innovation output (Gayle, 2003; Hashmi, 2013; and Beneito, 2014 among others). Nevertheless, it is worth noting that this effective relation between competition and innovation output may be sensitive to the different sectors, which have their individual institutional framework and innovation incentives. In addition, the policy variance in the different countries may also influence the probable outcome.

Concerning the auxiliary estimated regressors, the results are synonymous to that obtained in chapter 5. From the overall analysis, a highly significant and positive fitted R&D coefficient strongly confirms that greater innovation input and technological know-how of the pharmaceutical firms enable them to increase their patenting (or the value of patenting in particular). A negative relation of innovation output with age signifies that young firms are more dynamic and prone to have more innovation output than their older and well-established counterparts. However this effect of age is not significant in most of the cases and hence the proposed outcome is not fully established. On the other hand, a positive relation

between the size of firms and innovation output shows that bigger firms have more propensity to opt for patenting. Therefore, firms with higher human resource seem to engage more in knowledge output. Finally, a continuity and persistence in innovation performance is found to be maintained by the Dutch pharmaceutical sector, as affirmed by a consistently significant patent dummy.

By undertaking an exhaustive investigation to bring forth the essential relation between competition and innovation output, we may infer that our results are substantially robust to different econometric techniques which have been used to model our non-linear specification using different competition indicators. However, since many firms do not patent their innovation activities, all output in innovation may not be accounted for by the citation counts of individual patents. Therefore, as a further line of research, a non-linear econometric investigation can be carried out by substituting patenting activity with innovation investment. This will supposedly enunciate on how innovation captured by its investment is affected by the competitive conduct of the firms, and if any divergence in its impact can be accounted thereof. However, our empirical deliberation in this chapter suggests an economic disincentive for strategic patenting practices that builds a wall against competition. Since competition in the market is a fundamental aspect for its growth, innovation policies are required to buttress competitive phenomenon in the Dutch pharmaceutical industry.

Chapter 7

THE EFFECT OF INNOVATION ON FIRM PERFORMANCE: THROUGH WHICH CHANNEL DOES IT OCCUR?

Abstract

This chapter investigates on the effect of R&D expenditure and patenting on the productivity and overall output of the pharmaceutical firms in the Netherlands during 1996-2006. Special emphasis is attempted on the divergent effects of R&D activities and patenting practices on the productivity in this sector. Precisely, we examine if productivity is indirectly related to R&D investment through innovation output, or whether these two metrics to capture innovation can individually and directly affect the firm's productivity process. In view of this objective, a growth accounting approach has been used, whereby both labour productivity and total factor productivity are employed as the response variable. Our empirical results suggest that innovation measured using R&D investment as well as patent quality have a positive and significant effect on productivity. This inherently highlights the fact that not all innovations are deemed to be patented, and hence innovation captured through R&D investment can also have a prominent impact on productivity. Similar results are obtained for the innovation measures when labor productivity is used as the dependent variable. These findings provide counter argument to the CDM approach proposed by Crepon *et al.* (1998), who asserted that the R&D expenditure acts as an innovation input and hence, it cannot affect productivity directly. Concerning the effect of other determinants, firm size encourages overall productivity but attenuates labour productivity. Besides, a competitive environment is found to stimulate productivity. However the effect of

age remains indeterminate. As an ancillary to this approach, we have also identified the effect of R&D expenditure along with the factor inputs on the overall output of the firms in a production function framework. Our results suggest a significant impact of R&D investment on the pharmaceutical output in the Netherlands.

Keywords*:* R&D, Citation-weighted patents, Total factor productivity, Labour productivity

7.1 Introduction

The Netherlands is characterized by an ageing population which requires intense health care. This has consequently caused an increase in prophylactic and therapeutic drugs, and thereby spurring the demand and hence the output in the pharmaceutical industry. Inspite of being an R&D oriented industry, it is also involved in large scale operation and mass production of drugs, under rigid supervision on the quality and purity of the final output. The effective production of pharmaceutical products and its continual dissemination is fundamental for the well-being of mankind. In this context, it is prudent to finally analyze to what extent the in-house R&D expenditure and innovation performance through patenting affects the total factor productivity (TFP hereafter) and the labour productivity of the Dutch Pharmaceutical industry. In addition, the effective importance of R&D investment, as a factor input, that affects the general output in the drug sector is also an interesting avenue to explore.

Based on the statistical evidence by Nefarma (2003), the production of pharmaceutical products and materials in the Netherlands accounted to 5.6 billion euro in totality for 2001, with an annual growth rate of 15.3%. Establishing a perpetual progress at a moderate pace, it was ranked as the eight largest pharmaceutical producer in EU in the year 2006. In addition, a bulk investment in the Dutch pharmaceutical industry is dedicated to its R&D, spending almost 8% of

their turnover on R&D investments. However, as exemplified in chapter 1, the pharmaceutical production in the Netherlands is still far behind many of the major EU countries (evident from the survey by EFPIA, 2013). Taking into account this statistical facts and figures, it is worthwhile to discuss further on how the innovation mechanism in the Dutch pharmaceutical affects its output performance.

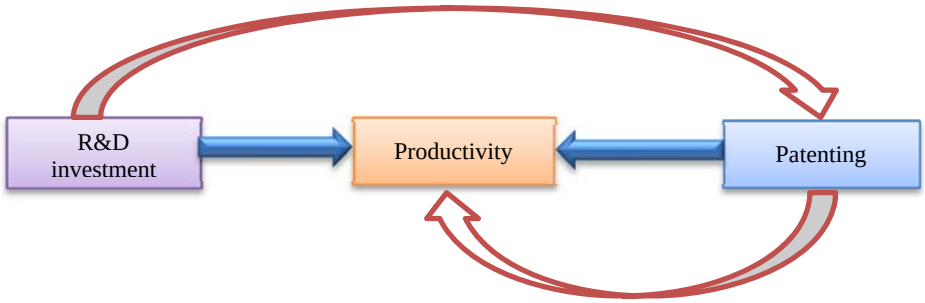
In general, the innovation activity of firms spurs productivity growth by improving the technical efficiency, which in turn, leads to the upgradation of the goods and services. This causes an increased demand and reduced production cost. Moreover, the innovative firms are likely to acquire a competitive edge over their non-innovative counterparts, with new entrants providing better and novel products can replace the inefficient ones; thus resulting in an increase in aggregate productivity. With the easing of a regulatory environment in the era of globalization, research intensive sectors resort to innovation activities to accentuate their productivity and growth, in order to gain a foothold in the competitive global market. However, despite a huge R&D investment by the innovative pharmaceuticals, a proportionate increase in output in the form of approval of new drugs is lacking, resulting in a productivity crises (Pammolli *et al.*, 2011). The influx of generic drugs and its mass production at a considerably lower cost inevitably proves to be detrimental to therapeutic innovation.

In light of the above discussion, we therefore aim at empirically analyzing the effect of innovation on the firm output and productivity in the Dutch pharmaceutical sector for the period 1996-2006, in a panel data framework. In our analysis, we attempt to capture productivity using two common practices for its measurement, that is, the total factor productivity and the labour productivity. In addition to the innovation measures, other factors that determine the level of productivity have also been considered and analyzed.

However the main purpose of this chapter is to focus on whether innovation captured through R&D investment and patent quality has divergent and distinct

effects on the level of productivity in this sector. In a way, we investigate if productivity is indirectly related to R&D investment which acts as an innovation input, through innovation output (as posited by Crepon *et al*., 1998) or whether these two innovation aspects individually and directly affects the firm's productivity process (as in Griliches and Mairesse, 1991). The inherent implication for the efficacy of the later phenomenon is based on the fact that, the measure of R&D expenditure reflects not only an input to innovation, but can also quantify the innovation process and its consequent outcome which may not be captured by patenting. This is due to the fact that not all innovations undergo patenting. This is expressed schematically in the following diagram.

Fig. 7.1: Schematic representation of the inter-linkage between R&D investment, patenting and productivity



The rest of the chapter is structured and organized in the following way. Section 7.2 provides a brief review of the past literatures dealing with innovation and productivity. Section 7.3 elaborates on the data and the descriptive statistics. Section 7.4 outlines the methodology used and section 7.5 provides the empirical results obtained. Finally section 7.6 lays out the concluding remarks on our findings.

7.2 Literature review

In response to the economic downturns or productivity gaps among the different nations, there has been a surge to research on this specific avenue for time immemorial. Recalling the pioneer work by Solow (1957), technological change is formidable for the potential growth in productivity. In the later years, proponents of the endogenous growth theory (like Grossman and Helpman, 1990 and; Smolny, 2000) have corroborated the proposition by Solow, asserting the fact that firms' innovation activity is crucial for the technological enhancement, fostering productivity growth. Klette and Griliches (1996), while formulating the innovation and productivity linkage, posited that R&D investment and innovation is the key driver of productivity and growth.

There exists a gamut of empirical literature dealing with R&D and the performance of the firms, where most of them unanimously affirms a significant and positive influence of innovation on the firm's performance (e.g. Griliches, 1986; Jaffe, 1986; Griliches and Mairesse, 1990; O'Mahony and Vechhi, 2009; and Griffith *et al*., 2006). But the estimated elasticity of productivity or output relative to R&D is found to differ significantly in the different literatures. While Griliches (1986) estimated the elasticity with respect to R&D to be approximately 0.07 for the U.S. manufacturing firms in a firm-level analysis, Jaffe (1986) observed the elasticity to be prominently higher (around 0.20) for an analogous set of firms in the U.S. In a successive study by Griliches and Mairesse (1990), the elasticity is found to be ranging from 0.25 until 0.45 for the same country data. It also posited an elasticity range for the Japanese manufacturing firms to be within 0.20 and 0.50. In a more recent paper by Griffith *et al.* (2006), prominently low level of elasticity was observed for the U.K. manufacturing firm, whereby estimated elasticity was established to be below 0.03. The difference in the elasticity can be attributed to the type of industry or the geographical location considered. In case of single sector studies, it is comparatively more pragmatic to find a larger elasticity for the R&D

intensive industries like the pharmaceuticals. Moreover the choice of the indicator for firm's performance, different econometric techniques as well as the dataset used causes the deviation in the elasticity of the firms in different literatures.

To evaluate the potential pathway through which innovation affects productivity, Crepon *et al.* (1998) propounded an empirical framework to establish the innovation and productivity relationship in a three stage structural model approach. Based on this approach, he disentangled the effect of innovation input to innovation output and the consequent influence of innovation output to the productivity of the French manufacturing firms for the period 1986-1998. The findings exhibited a robust and positive relation between innovation and productivity. Many subsequent papers used the CDM model (pioneered by Crepon *et al.*, 1998) to articulate the innovation-productivity relationship explicitly (e.g. Benavente, 2006; Griffith *et al*., 2006; Janz *et al*., 2003 and Hall *et al*., 2009 among others). Several papers exhibited variants of the CDM model, which includes Mairesse *et al.* (2005) and, Loof and Heshmati (2006) inter alia. Loof and Heshmati (2006) opted for both levels and growth rates for measuring productivity, while others have resorted to either of them. It is to be noted that, most of the literature dealing with the CDM model uses cross-sectional data for estimation purpose. Therefore, the contribution to productivity based on the timing of innovation is ignored at large under a cross-sectional framework. However exceptions were seen in the works by Peters (2006) and, Masso and Vahter (2008) where CDM estimations were carried out in a panel data context to trace out a dynamic response between them.

It can be emphasized that very few studies use observed innovation variables alone, for their estimation purposes. This is due to the lack of complete innovation data in the surveys. The inadequacy in innovation observations can jeopardize the true state of consequences if the choice of estimation techniques is not so robust. To mitigate this problem, CDM approach by Crepon *et al.* (1998) uses the predicted

value of the innovation variables. In this seminal study based on the CDM framework, the estimation methodology used for the first two stages are akin to the ones that we have undertaken in our preceding chapters, which involves generalized Tobit and count data specifications. For the last stage of this model, the effect on productivity by the predicted patents is asserted by an asymptotic least square method (ALS). However empirical studies like Criscuolo and Haskel (2003), and Duguet (2006) uses simple ordinary least square (OLS) and generalized method of moments (GMM) methods respectively, in order to obtain the productivity-innovation relationship. While the former articulated that lagged process innovation leads to TFP growth, the latter purported a TFP growth with radical innovation only.

Although a huge number of literature is dedicated to the empirical exposition of the innovation-productivity relationship, investigation on this contemporary subject in the context of the pharmaceutical industry is very sparse. Also, the overall productivity, in general, is not considerably accounted for in the academic literatures dealing with the pharmaceutical industry, whereas the knowledge productivity is adequately dealt with in the empirical studies that focus on this particular sector (like, Pammolli, 2011). Nonetheless, a study by Malmberg (2007) has brought forth the effective relation between innovation output and total factor productivity (TFP) at the micro level, in the Swedish pharmaceutical industry. Dealing with the historical period of 1952-1977, this study asserts an effective positive influence of the drug reserves and patent inventories on the TFP growth, with the former determinant having a greater impact in the 1950s while the latter gains prominence in the 1960s. Despite the fact that this analysis circumscribes an important period during which the Swedish pharmaceuticals have undergone profound institutional changes that impacted their innovation characteristics, more recent analysis is lacking for the drug sector. However a study by El-Shinnawy (2009) considers a relatively recent time span of 1993-2005 for examining the extent to which internal efficiencies of individual firms contributes to the

productivity and growth mechanisms in the Egyptian pharmaceutical sector. For this analysis, the Malmquist TFP index is estimated using a panel data framework, which procures efficiency change, technical change and TFP growth. The results suggest a positive TFP growth rate under a protective regime. However this study focuses on the technical efficiency instead of innovation capacity of the pharmaceuticals, in conjunction with the concerned socio-economic conditions and the policy reforms that prevailed in the Egyptian pharmaceuticals at that time. Hence the pertinence of this study in relation to the more developed countries, where innovation dominates the pharmaceutical arena, is debatable.

7.3 Data descriptive and TFP estimation

Since a detailed and elaborative discussion on the construction of the different variables, along with the data extraction process is provided in chapter 3, we solely focus on the descriptive statistics of our data that highlights the ad hoc features of this study. However to perceive the applicability of the different variables considered, we elucidate briefly on their structural perspective.

7.3.1: Structural outline of the variables used

Growth accounting approach

In case of the growth accounting approach, we first estimate the TFP from a Cobb-Douglas production function, which serves as our dependent variable. For the estimation of TFP, two notable factor inputs, viz. the labour input (log of the number of employees) and capital input (deflated log of capital intensity[9]) are considered. The predicted TFP is then regressed on various firm level determinants that include, firms size, age, concentration index, citation counts and R&D intensity. The last two determinants of innovation are our prime variables of

[9] As mentioned in chapter 3, the nominal values of R&D intensity and capital intensity are deflated using gross fixed capital formation price index from the EU Klems database, where 1995 is considered as the base year.

interest. Since the citation counts for each patent exemplify the value of the concerned patent, it has been preferred for our empirical analysis. Yet, we apply the patent data for our data descriptive to quantify the literal innovation output in the Dutch pharmaceuticals. As an alternative measure of productivity, we have also adapted the labour productivity as the response variable in the growth accounting framework (similar to Crepon *et al.*, 1998).

Focusing on the various control variables applied to determine the productivity of the firms, the firm size is determined by the log of the number of employees (e_{it}). Leung *et al.* (2008) asserted a significantly positive relation between firm size and productivity using both TFP and labour productivity for Canadian manufacturing and non-manufacturing industries. Concerning the influence of the log of firm age on productivity, Huergo and Jaumandreu (2003) found that new entrants experience very high productivity growth which tends to last for many years, albeit a convergence in productivity is achieved for the surviving firms.

We attempt for a less comprehensive analysis on the effect of the competitive conduct of firms on productivity performance, since we aim at keeping our focal point on how the different innovation criteria impact the productivity and the concomitant channel through which this takes place. Hence, as an indicator of competition, we have incorporated only the log of market share concentrated index, which is identical to the market share measure (using sales) that is applied in chapter 4 and 6. A description on its computation can be found in chapter 3.

Production function approach

Next, to stress further on the impact of innovation as a factor input in the productivity performance of this sector, we revert to the production function, and introduce R&D as an additional input together with labour and capital. The response variable is therefore the log of value-added which reflects the production output measure (this variable is deflated by the value-added price deflator from the

OECD STAN database). The production function framework is further employed in our study, as an ancillary to the growth accounting approach, to identify more precisely the role of R&D in the overall production performance in this sector.[10] However it is essential to emphasize that we have used R&D expenditure and not R&D intensity, as an essential factor input in this production function approach.

7.3.2 Obtaining the TFP variable for the growth accounting approach

For the estimation of TFP, we do not use the OLS technique as it may result in omitted variable bias, as the firm's input choices are likely to be correlated to unobserved shocks in productivity. This problem can be mitigated by using the Levinsohn-Petrin technique (Levinsohn and Petrin, 2003) as it uses a firm's intermediate inputs as a proxy, in order to correct for any productivity shocks associated with the corresponding inputs of the firms (Petrin *et al.*, 2004).[11]

Hence, we resort to the Levinsohn-Petrin approach using a Cobb-Douglas production function for the estimation of the predicted TFP. Since the methodology is already specified in chapter 3, we directly consult the estimated results obtain. Therefore, table 7.1 represents the estimated values for capital and labour inputs.

Table 7.1: Levinsohn Petrin productivity estimation

Dependent variable: log(value-added)		
	Coefficient	p-value
log(Employees)	0.440***	0.000
Log(Capital)	0.339***	0.000
Wald test(Chi^2)	8.770***	0.003

[10] We have opined to use distinctly the production function approach and the growth accounting approach, and not the entire analysis (using all the determinants) in a single production function framework, from the view point that, citation-weighted patents are the output of the innovation mechanism, and not an input of the production process. Hence we tried to analysis the production function framework from the perspective of how factor inputs affect the consequent output.

[11] The productivity technique developed by Olley and Pakes (1996) is also a consistent estimator to solve the simultaneity problem. However we could not apply this approach owing to the unavailability of investment data (since investment is used as a proxy for the unobserved shocks in productivity in this approach).

From the estimation analysis, it is found that the log of the inputs are highly significant at 1% significant level. As noticeable in table 7.1, the sum of the coefficients for log(employees) and log(capital) is found to be less than 1. But to determine whether the coefficients depict constant returns to scale (CRS), we perform the Wald test for CRS. A significant *chi²* for the Wald test denotes that we reject the null hypothesis of CRS. The TFP is subsequently predicted from the estimation.

7.3.3 Summarizing the statistical data

Capital and labour inputs for innovating and non-innovating firms

Since capital and labour inputs essentially define the production capacity of an industry, the following figures put forth the capital and labour resources for innovative and non-innovative firms. For clarifying the innovative and non-innovative pharmaceuticals, we have bifurcated them based on firms that perform R&D (or patent) as the innovative firms and the reverse for the non-innovative firms.

In case of fig. 7.2, we have categorized the non-R&D performing firms as the ones with zero R&D values, and the R&D performing are the firms which have a positive value for the concerned year. Since the R&D data is fraught with numerous missing observations, we have not considered the missing values as non-R&D performing, simply because the innovation activities of those firms are unobservable and unclear. Hence, for the moment we consider a balanced data for each year, in order to arrive at our data analysis. However, it may be noted that we have not incorporated the imputed R&D values in this section and kept it for our empirical analysis, and hence, we solely focus on the data at hand in this section. Nonetheless, we do not face the problem of missing values when we use patents as the innovation indicator (in fig. 7.3), although there exists many firms with zero patents. Considering the skewed distribution in our data sample, we have opted for

the median values of the factor inputs in each year. Simultaneously, it is also more robust than the averages in the presence of any outlier values.

Fig. 7.2: The median values of factor inputs of R&D performing & R&D non-performing firms

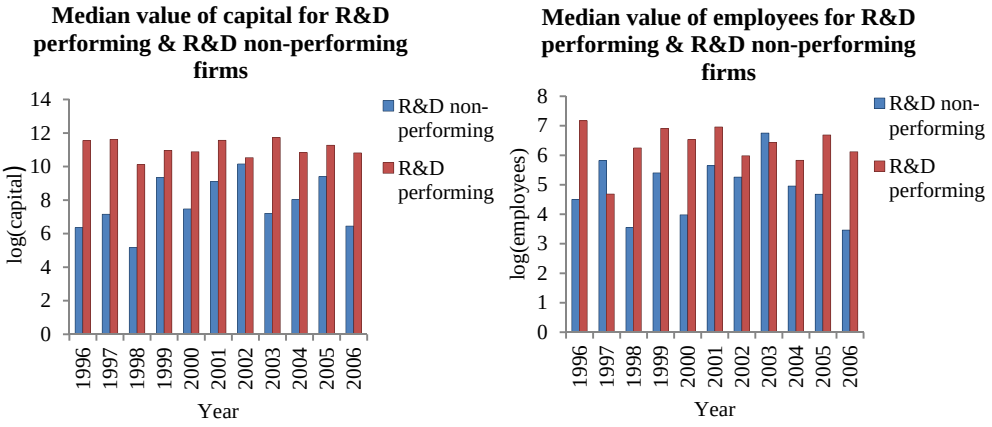
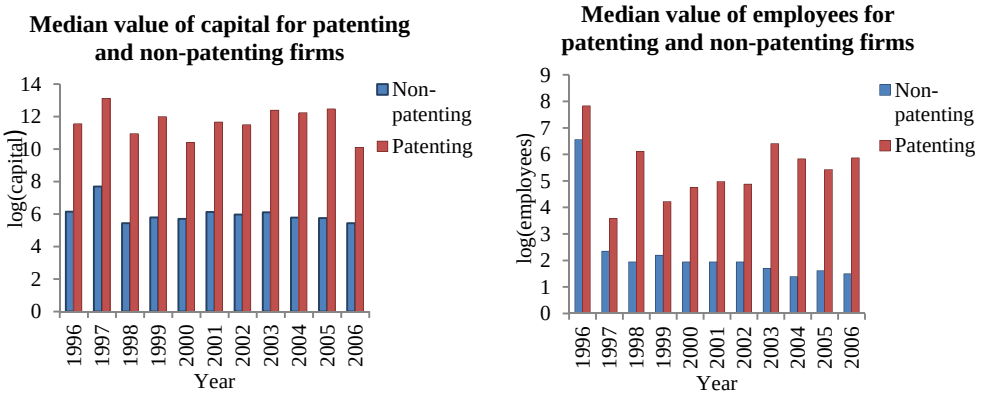


Fig. 7.3: The median values of factor inputs of patenting & non-patenting firms



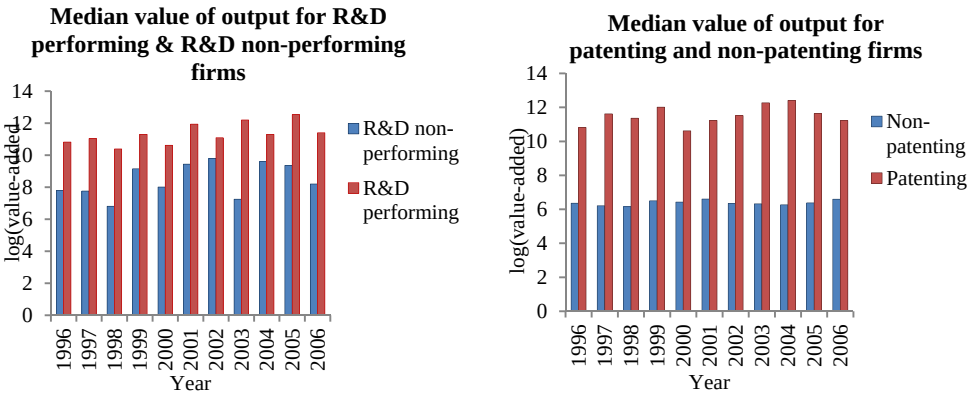
In fig. 7.2 and 7.3, capital and labour is quantified in their logarithmic values. In the displayed charts, it is observed that the innovative firms are, in general, possessing a greater reservoir of capital and labour inputs than their non-innovative counterparts. However the difference in the factor inputs is more conspicuous in case of fig. 7.3, for the patenting and non-patenting firms. This asserts the fact that,

the patenting pharmaceuticals have a prominently higher resource base at their disposal compared to the non-patenting ones. On the other hand, the differences in the input stock for the R&D performing and R&D non-performing firms are comparatively less (as evident in fig. 7.2). In spite of the capital stock being prominently higher for the R&D performing firms, the difference in the labour input is not so distinct in this case. Nevertheless, no systematic trend is observed for the concerned period.

Production output for the innovative and non-innovative firms

Subsequent to our preceding analysis on factor inputs, we further explore the overall output of the pharmaceuticals, in the context of innovative and non-innovative firms. The output is denoted by the log of value-added. Similar to fig. 7.2 and 7.3, we consider those firms that perform R&D investment or engage in patenting as the metric for innovating firms. In sync with their input capacity, their corresponding output is also explicitly higher for R&D performing or patenting firms. The divergence in the production capacity for innovative and non-innovative firms reveals the leverage of the former over the latter. However no prominent increasing or decreasing trend can be asserted for the innovating vis-à-vis non-innovating firms.

Fig. 7.4: The median values of production output for innovative & non-innovative firms



The above discussed measure of output and its corresponding factor inputs are used for our empirical estimations using the production function framework, where R&D capital is also considered as a regressor. Conjointly, the estimation of the predicted TFP is also the first stage for the growth accounting approach, where output is regressed on the capital and labour inputs to attain the TFP variable. As it has already been specified, the TFP variable acts as one of the response variables in this approach, while the labour productivity is used as an alternate. Therefore in the next subsection, we take into account the descriptive investigation for the productivity variables.

Productivity for innovative and non-innovative firms

For a better understanding of the variation in TFP and labour productivity for innovative and non-innovative firms, we plot the Epanechnikov kernel density, which portrays their probability density functions. In fig. 7.5, it is observed that the kernel density estimate for patenting firms is more peaked than the non-patenting firms. Also for patenting firms, the logarithm of TFP has a density that is right-skewed (as represented in the first graph).

However, in case of fig. 7.6, the density for log(TFP) does not show any prominent skewness between R&D performing and non-R&D performing firms. On the contrary, the peak for TFP density in case of the non-R&D performing firms is relatively higher than the R&D performing ones, although the difference is considerably less prominent and not very drastic. But, in case of labour productivity, the kernel density for R&D and non-R&D performing firms is almost in concordance with that of the patenting and non-patenting firms, suggesting greater productivity for the innovative pharmaceuticals.

Fig. 7.5: Kernel density plot of TFP and labour productivity for patenting and non-patenting firms

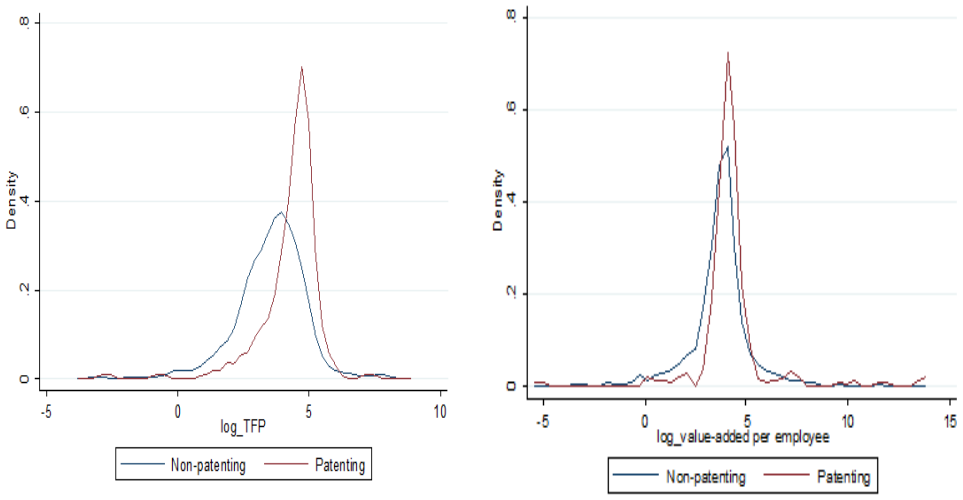
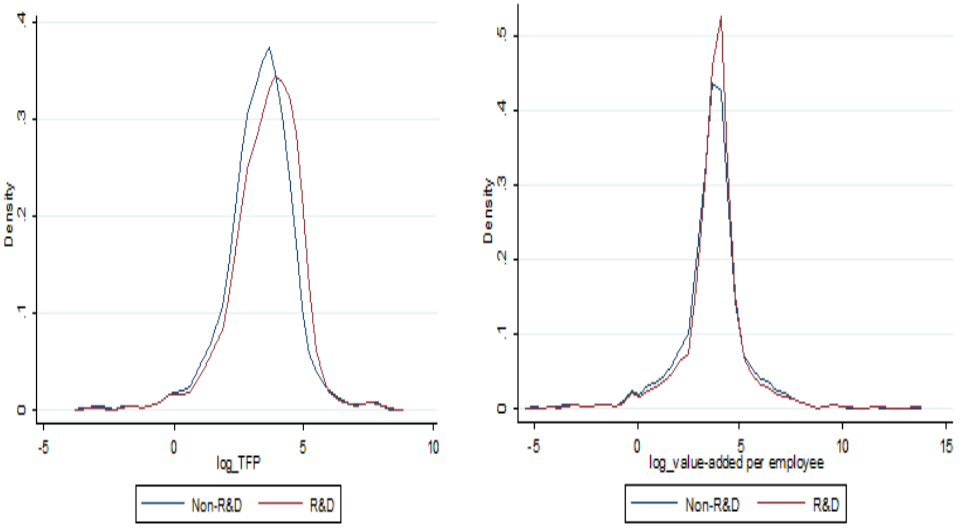


Fig. 7.6: Kernel density plot of TFP and labour productivity for R&D and non-R&D performing firms



It can summarized that, although the changes of labour productivity is not completely identical to that of TFP, their overall outcome seems to be similar. In essence, the labour productivity or TFP is found to be considerably higher and

having a stochastic dominance in case of innovating firms, as compared to the non-innovating firms. This boils down to the point that the innovative Dutch pharmaceuticals have a greater incentive in terms of their productivity, which leads to overall growth. Also, as suggested by the previous diagrammatic illustrations, the innovative pharmaceuticals have a greater endowment of capital and labour inputs, which materialize into higher output production.

We next proceed to continue our analysis, based on the empirical investigations. The following section deals with the empirical methodology that includes both the growth accounting approach as well as the production function approach, which delineates the effect of innovation on productivity and output respectively.

7.4 Empirical methodology

In order to investigate the effect of innovation on the firm's overall performance in productivity, we primarily consider the growth accounting approach. In consideration of the fecundity in patenting activities in the pharmaceutical industry, both R&D intensity and patenting are employed as the innovation criteria. An additional exploration of the effectiveness of R&D as a factor input is looked into in a production function approach.

For both the production function and growth accounting approach, a random effect Generalized Least Square (GLS) and system Generalized Method of Moments (GMM) techniques have been used as the estimation techniques. However, for the production function approach, the Levinsohn-Petrin method is included as an additional empirical modeling, which has also been adapted to estimate the predicted TFP for the growth accounting model. Before elaborating on the estimation techniques, the following subsections provide an overview of the production function and growth accounting approach.

7.4.1 The effect of innovation on productivity: The growth accounting approach

This methodology delves into the endogenous growth framework which analyses the effect of R&D intensity or patent counts on the productivity, taking into account a number of other firm characteristics as the necessary controls. We undertake this methodology by closely following the empirical outline in the works of Coe and Helpman (1995) and Atella and Quintieri (2001), among others, who promulgated the effect of R&D intensity on the estimated firm-level TFP. On the other hand, several studies look into the effect of patenting on productivity mostly by using the CDM model (as introduced by Crepon *et al.*, 1998) or a variation of it (for example, Benevante (2006), Griffith *et al*. (2006), Loof and Heshmati (2006), and many more). But the CDM approach builds on the fact that the innovation output of the firms affects productivity directly, and not the R&D expenditure.

In the light of these foregoing empirical frameworks, we develop on the growth accounting model, wherein both R&D expenditure and citation-weighted patents have been incorporated. However, we question the CDM approach for the indirect relation of R&D expenditure with productivity. The empirical structure for the CDM models build on the fact that R&D intensity is the innovation input and does not impact productivity directly, but through patenting performance which acts as an innovation output. Notwithstanding, we consider whether both R&D expenditure and patenting have a direct linkage with productivity. The pragmatic and logical reason for this phenomena to occur is mainly due to the fact that, not all innovations result in patenting. It is no doubt that the pharmaceutical firms have a predominant use of patents, which plays an important role in measuring the inventive success of the firms. But at the same time, the R&D expenditure measure reflects not only their innovation input but also circumscribes those innovation activities that are not formalized to patents.

In view of the above consideration, we adopt both R&D intensity and citation-weighted patents as our primary explanatory variables. However, an important issue that we have confronted in our empirical analysis is the lack of R&D expenditure data. It is rightly posited by Danzon and Percy (1999) that, although R&D investment is a fundamental factor to propagate pharmaceutical productivity, there exists inadequate R&D data. Although with time and more improved survey designs, the quality of R&D data at the firm level has improved significantly[12], there still remain huge gaps in accruing R&D data extensively. To circumvent this problem of large number of missing R&D observations in our dataset, we perform our estimations using two categories of fitted R&D intensity values, which are obtained by estimating a Tobit II regression (refer to A.5.2). The disparity in the two kinds of R&D fitted values lies on the fact that, one uses the fitted values as additional observations for R&D intensity; whereas the other considers the remaining missing values for R&D intensity as zero, after merging the predicted values with the R&D intensity data. It is noteworthy that, unlike the production function framework (which we will discuss shortly in subsection 7.4.2), we do not use a balanced R&D data for the productivity equation, as it causes a loss of other corresponding variables used and may not capture the true picture of their effects on productivity.

Regarding the citation-weighted patents, such problems are not faced, owing to a complete set of observations. In our estimation model, the effects of these two innovation variables are evaluated directly and individually. However, conforming to the CDM methodology[13], we also estimate on how productivity is affected by expected innovation performance, conditional on R&D intensity and other firm characteristics. The purpose for this alternative mechanism is to provide a more clear picture on the effective channel through which innovation can affect

[12] A detailed review of the innovation surveys is provided in Mairesse and Mohnen (2010).

[13] We have dealt with the first two equations of the CDM model in our previous chapters. Chapter 4 is based on the R&D equation, where R&D is a function of other firm characteristics. Subsequently the innovation equation is considered in chapters 5 and 6, where patent is related to R&D.

productivity. The corresponding specifications based on the mentioned conditionalities are enumerated as follows.

$$ln(P_{it}) = \rho_0 + \rho_1 Predicted[ln(R\&Dintensity_{it})] + \rho_2 PAT_{it} + \rho_3 X_{it} + \kappa_{it} \quad (7.1)$$

$$ln(P_{it}) = \beta_0 + \beta_1 Predicted[PAT_{it}] + \beta_2 X_{it} + \theta_{it} \quad (7.2)$$

Equation (7.1) corresponds to a productivity equation where both fitted R&D expenditure $(Predicted[ln(R\&Dintensity_{it})])$ and citation-weighted patent counts (PAT_{it}) are used as the innovation indicators. Hence this equation reflects on whether there is a direct linkage between R&D and productivity, alongside the patent performance of the firms. As an alternative methodology (as observed in equation (7.2)), we have used the last equation of the CDM model, where the predicted patent performance determines the productivity of the firms and no direct relation with R&D intensity is observed. The ZINB estimation for obtaining the fitted citation counts is annexed in A.7.1.

Regarding the dependent variables used for measuring productivity, P_{it} represents TFP $(ln(TFP_{it}))$ or labour productivity $(ln(LP_{it}))$, which are applied separately and interchangeably in our estimations. As explained before in this chapter, $ln(TFP_{it})$ is the predicted value obtained from the Levinsohn-Petrin estimation (estimation results in Table 7.1). It can be noted that, the labour productivity takes into account the amount of output per unit of labour, while the TFP encompasses the distinct effects of both capital and labour inputs. Hence, the latter is likely to capture productivity comparatively more effectively and better represent the overall performance of the Dutch pharmaceutical firms.

The X_{it} in the above equations is the vector of control variables that influence the productivity of the Dutch pharmaceuticals and includes, firm size (e_{it}), age (a_{it})

and concentration index (s_{it}). These explanatory variables are discussed in section 7.3 of this chapter. The constant terms ρ_0 and β_0 represents the firm specific heterogeneity and; κ_{it} and θ_{it} are the corresponding measurement errors.

For the purpose of econometric modeling we have used the random effect GLS and the system GMM methods to estimate our productivity equations, which accounts for the simultaneity bias problem. These techniques are discussed in details in the subsection 7.4.3.

7.4.2 The effect of R&D expenditure on output: The production function approach

In order to inspect the efficaciousness of R&D as a necessary factor input, in company with labour and capital inputs, we attempt to consider this issue in greater details, under a production function framework. In other words, we examine the effect of R&D investment on the output of the firms. Similar to subsection 7.3.2, the output is measured by the log of value-added. However, unlike the TFP estimation in the mentioned subsection (which has been used for the growth accounting approach), the additional input in this case is the R&D as a capital input. Since the pharmaceutical industry is heavily involved in innovation activities, it is pertinent to include the R&D as an additional factor input in the production process. This is in line with a number of empirical works, ranging from Schankerman (1981) to Branstetter and Chen (2006), who investigated the impact of innovation on the output of the firms by using a production function.

It is worth-mentioning that this approach is a separate identification for R&D innovation, whereby it enters into the production function as a necessary input to production. Also, as mentioned before in the footnote, we re-emphasize on the fact that the entire modeling framework of this chapter is not corroborated to a single production function framework, from the perspective that the citation-weighted

counts are the output of innovation and hence may not suffice to the concept in which we base the production function approach for this study.

Consequently, the baseline specification for the production function is expressed in the following equation, which demonstrates the effect of various capital inputs (in the form of R&D capital, net tangible capital and human capital) on the output of the Dutch pharmaceuticals.

$$ln(Y_{it}) = \eta_0 + \eta_1 ln(R\&D_{it}) + \eta_2 ln(L_{it}) + \eta_3 ln(C_{it}) + \lambda_{it} \quad (7.3)$$

where, the dependent variable Y_{it} is the value-added of the firms; $R\&D_{it}$, L_{it} and C_{it} are the explanatory variables representing R&D capital, labour and physical capital respectively; η_0 is the firm specific heterogeneity and; λ_{it} is the corresponding error term.

This production function equation (equation (7.3)) can be categorized as an augmented Cobb-Douglas production function with an additional incorporation of innovation input as an explanatory variable. In this context it can be noted that although a translog production function is less restrictive, the restrictions on the Cobb-Douglas function does not cause any significant difference empirically and hence we prefer the latter.

Although the R&D values, has an advantage of being measured in similar units (in million euro) to output and capital for a better statistical overview of the data, the missing value problem is strikingly evident. Hence, in consonance with the growth accounting approach, the fitted R&D expenditure is computed, based on the two different specifications for the prediction of R&D. Additionally, a balanced R&D data set is employed, which contains only the reported R&D measures and their corresponding other variables. But this strategy has resulted in a loss of huge number of observations.

However in case of the production function estimation at the micro level, the fundamental problems of simultaneity bias exists. The Levinsohn-Petrin approach, as used before in subsection 7.3.2, deals with the unobserved productivity shocks in a two-step estimation technique (similar to Olley and Pakes, 1996). However, Wooldridge (2009) points out that the unknown function of the semi-parametric Levinsohn-Petrin approach is not well approximated in case of higher order polynomials. Therefore, he preferred the GMM framework and suggested a variant of the Levinsohn-Petrin approach. Since the Levinsohn-Petrin two-stage estimation method does not identify parameters of the variable inputs in the first stage, we propose to use the GMM technique as an alternative estimation strategy for our augmented production function. To address the simultaneity bias problem, we incorporate the system GMM technique (Blundell and Bond, 2000), which accounts for the unobserved heterogeneity in a dynamic panel context. In addition, we also perform a complementary estimation strategy using a random effect GLS method, considering the panel structure of our dataset. In this estimation procedure, λ_{it} is considered to be a random variable and it is based on the assumption that λ_{it} and the other explanatory variables are uncorrelated. Based on this discussion on the econometric modeling issue, we present a comprehensive elaboration of these techniques in the next subsection.

7.4.3 Econometric models used

Random Effect GLS

By virtue of our constructed panel dataset, it is possible to address endogeneity or causality issues, which is difficult to be achieved in a cross-sectional framework. Therefore, as the primary estimation technique in regard to a panel data context, we use the GLS method, allowing for random effects. We have adopted the random effect technique instead of the fixed effect, as the Hausman specification test (Hausman, 1978) leads to the non-rejection of the null hypothesis and hence classifies the individual effects in our model to be random. Therefore, this signifies

that the inference of the estimation model is drawn randomly in respect to population of all effects.

Although both OLS and GLS methods can be technically unbiased in nature, the latter is superior to the former in terms of being asymptotically efficient. To describe the random effect GLS we consider the basic model in the matrix-form as,

$$Y = \beta X + W \tag{7.4}$$

In regard to the above specification, $E(WW') = \Omega$, where Ω represents the variance-covariance matrix. The GLS provides efficient parameter estimates, provided that the structure of Ω is known. Therefore,

$\hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y$ and,

$$Var(\hat{\beta}_{GLS}) = \sigma_{\mu}^{2}(X'\Omega^{-1}X)^{-1}$$

Ω^{-1} is calculated by using the following formula,

$$\Omega^{a} = (\sigma_{\mu}^{2})^{a} I + (T\sigma_{\phi}^{2} + \sigma_{\mu}^{2})^{a} R \tag{7.5}$$

where, $\mu + \phi = W$, a is any arbitrary scalar, T is the time matrix, and I denotes idempotency whereas R denotes orthogonality.

However, the GLS is an efficient and consistent estimation method for panel data if the explanatory variables are exogenous in nature. Hence, in the presence of endogeneity, the random effect GLS may lead to inconsistency in the estimated coefficients. This leads us to use an alternative estimation technique, called system GMM.

System GMM

The system GMM prominently mitigates the weak correlation problem and provides with more reliable estimates (Blundell and Bond, 2000), as it effectively deals with the potential endogeneity of its regressors and the bias due to the lagged dependent variable in a dynamic panel data framework. Thus, this model structure enables us to utilize our panel data in a dynamic context, where unobserved heterogeneity is considered among individuals at both past and current levels of output or productivity.

However the standard GMM estimation which considers the first-difference has been found to provide unsatisfactory and inconsistent estimations by Mairesse and Hall (1996). Blundell and Bond (1998) asserted that this bias in the estimation results of Mairesse and Hall (1996) is due to weak instruments, which could be lessened when stationarity restrictions are exploited on the initial conditions. Blundell and Bond (2000) confirmed that the extended GMM estimator provides more consistent parameter estimates with the addition of new instruments and also introduced autoregressive components to reduce productivity shocks. Therefore, we use the extended GMM in levels, developed by Arellano and Bover (1995) and Blundell and Bond (1998), which is more asymptotically efficient and suitable to study the growth accounting process.

To empirically model the system GMM, we consider the first-difference GMM with AR(1) as the basic model, assuming unobserved individual specific effects in a dynamic panel dataset.

$$y_{it} = \alpha y_{it-1} + \mu_{it} + \phi_{it} \tag{7.6}$$

where $\mu_{it} + \phi_{it} = w_{it}$, (μ_{it} and ϕ_{it} are the component structures of the standard error.)

The assumption, following the moments restriction can be written as,

$$E(y_{i,t-k}\Delta\phi_{it}) = 0 \text{ for } t=3,..,T;\ s \geq 2 \text{ and } s \neq t \tag{7.7}$$

The additional assumption for the system GMM estimator, which is more suitable for a autoregressive model, in the context of a persistent panel data, is,

$$E(\mu_{it}\Delta y_{i2}) = 0 \text{ for } i=1,...,N. \tag{7.8}$$

The fundamental requirement of this assumption is the stationery restriction on the initial condition y_{i1}.

Combining equation (7.7) and (7.8), this assumption provides further *T-2* moments condition,

$$E(w_{it}\Delta y_{it-1}) = 0 \text{ for } i=1,...,N \text{ and } t=3,...,T. \tag{7.9}$$

Therefore, considering equation (7.7) and (7.8), the instrument matrix of the system GMM can be written as,

$$M_i^+ = \begin{bmatrix} M_i & 0 & 0 & ... & 0 \\ 0 & \Delta y_{i2} & 0 & ... & 0 \\ 0 & 0 & \Delta y_{i3} & ... & 0 \\ . & . & . & ... & . \\ 0 & 0 & 0 & ... & \Delta y_{iT-1} \\ \end{bmatrix} \tag{7.10}$$

where M_i is a $(T-2)\times m$ matrix which finally expresses the complete second order moment condition as,

$$E(M_i^+ w_i^+) = 0, \text{ where, } w_i^+ = (\Delta\phi_{i3},....,\Delta\phi_{iT}, w_{i3},...,w_{iT}) \tag{7.11}$$

Therefore, in the system GMM estimation technique, suitable lagged levels of instruments is employed to the standard set of equations in first-differences, by

using additional equations in levels, instrumented by suitable lagged first differences.

It can be noted that, contrary to the empirical modeling in the previous chapters, we do not use the Wooldridge specification (Wooldridge, 2005) for random effects in this chapter. It is because Wooldridge (2005) himself points out that, for linear panel data model the problem of unobserved effects is reduced effectively by using GMM estimation developed by Blundell and Bond (1998), which greatly improves the GMM method by imposing restriction on the distribution of the initial conditions.

Levinsohn-Petrin

For estimating the production function, we use the semi-parametric Levinsohn-Petrin method as an additional estimator. This approach (Levinsohn and Petrin, 2003) uses intermediate inputs as a proxy to eliminate simultaneity problems. It specifically focuses on minimizing the measurement error problem, especially in calculating capital. If the capital and labour variables are positively correlated and the capital data is marked with error, then the probable estimated coefficient for capital may tend towards zero and make the effect of labour on output more pronounced. In this estimation methodology, we incorporate a intermediate material input, other than capital and labour in the production function.

$$ln(y_{it}) = \eta_0 + \eta_1 ln(l_{it}) + \eta_2 ln(c_{it}) + \eta_3 \ln(i_{it}) + \mu_{it} + \phi_{it} \tag{7.12}$$

where, $\ln(y_{it})$ is the value-added[14], $ln(l_{it}), ln(c_{it}), ln(i_{it})$ are the labour, capital and intermediate inputs respectively; μ_{it} and ϕ_{it} denotes the error terms. Hence, the input demand function is given by,

[14] Levinsohn and Petrin (2003) asserted that, in the estimation of value-added production function, the separability condition should hold for the inputs. This condition is generally observed in firm-level data. However, following Petrin *et al.* (2004), we consider the STATA command that takes into account value-added as the dependent variable in the production function.

$$i_{it} = i_{it}(\mu_{it}, c_{it}), \tag{7.13}$$

which is monotonic in $\mu_{it} \forall c_{it}$ to be a valid input proxy. Thus, under conditions of monotonicity, the inversed input demand function can be written as,

$$\mu_{it} = \mu_{it}(i_{it}, c_{it}) \tag{7.14}$$

Hence, the unobservable productivity term is expressed as a function of capital and intermediate input.

Therefore, considering value-added as the output measure (obtained as a net of gross output), equation (7.12) can be re-written as,

$$ln(y_{it})^{*} = \eta_1 ln(l_{it}) + \xi_{it}(i_{it}, c_{it}) + \phi_{it} \tag{7.15}$$

where, $\xi_{it}(i_{it}, c_{it}) = \eta_0 + \eta_2 ln(c_{it}) + \eta_3 \ln(i_{it}) + \mu_{it}(i_{it}, c_{it})$

As has been explained in Petrin *et al.* (2004), a polynomial approximation of the third-order in i_{it} and c_{it}, instead of $\mu_{it} = \mu_{it}(i_{it}, c_{it})$, provides with consistent estimation with the use of value-added production function. A detailing of the polynomial approximation is provided in Petrin *et al.* (2004). The Stata command for this estimation strategy takes into account this empirical issue.

Recalling equation (7.3), we have incorporated an additional R&D input in the production function. Hence, it is to be seen as to what extent the estimator provides accurate results when R&D capital is incorporated as additional regressors.

7.5 Empirical estimation

In this section, we enumerate the empirical findings of our estimations. We first elaborate on the growth accounting approach and subsequently discuss on the production function approach.

7.5.1 The growth accounting approach

In this section, we focus on how the various firm-level determinants affect the level of productivity in the Netherlands' pharmaceutical sector. We first consider the total factor productivity as the dependent variable, following which we use the labour productivity as an alternative quantification. We re-state on the fact that, the former is predicted from the estimated Levinsohn-Petrin model, while the latter is a ratio of value-added and the number of employees. A detailed explanation on the productivity variables and the corresponding regressors are provided a priori.

Regarding the empirical modeling, a random effect GLS is used as the primary estimation method. As an alternative candidate for this task, a system GMM estimation is further used. Nevertheless, the empirical methodology essentially suggests that the system GMM technique is a better estimator than the random effect GLS, as it adequately handles endogeneity of its regressors and biased estimations that are likely to occur. In this approach, we also use the lagged dependent variables as regressors in the random effect GMM estimations, such that a more robust result can be expected.

The main focus of our empirical analysis is to examine how patent performance and R&D intensity of the firms affect the level of productivity. In a way, we try to explore through which channel innovation influences productivity, for the essential growth in this particular Dutch sector. The estimations for the productivity equation in table 7.2 and 7.3 have the same sequence of regression analysis. For first regression analysis, we have used only citation-weighted patents, as an innovation measure, to indicate if innovation output alone has a prominent impact on productivity. Subsequently, we have added the fitted R&D intensity values in regression 2 and 3 for identifying to what extent the innovation captured through the R&D measure and citation-weighted patents affect productivity individually. The criteria for the two distinct fitted R&D has already been discussed in subsection 7.4.1. Finally we use the predicted value of the citation-weighted

patents, following the CDM approach, where an indirect relation for innovation input and productivity is considered. The ZINB estimation from which the predicted value for citation counts is computed can be found in appendix A.1.[15]

Using TFP as the dependent variable

We first discuss the estimation results obtained using TFP as the regressand, as represented in table 7.2. Focusing primarily on the innovation variables, the coefficient for the citation-counts are found to be systematically insignificant with a very low level of magnitude in case of random effect GLS. However, in case of system GMM, a comparatively higher magnitude for citation counts are observed, with high significance level. A noticeable similarity in regression 1 and 4 is observed, where the coefficients for the citation counts and the fitted citation counts in addition with other regressors, bears uncanny resemblance. This is due to the fact that a very nominal number of citation count variables are fitted.

Regarding the effect of R&D intensity, both regression 3 and 4 confirms its significant effect on TFP, exhibiting a greater magnitude than the coefficients for patent counts. These observations are analogous to both estimation techniques used. But the magnitudes of the coefficients are relatively higher for system GMM. Nevertheless, the R&D intensity and patents differ in their measurement approach. While the former is quantifiable as R&D investment on a per worker basis, the latter is a count variable. Hence the magnitude of their coefficients cannot be compared. Howbeit, the R&D intensity consistently portrays a significant impact on TFP, which is found to be absent for citation counts, when random effect GLS is used.

Regarding the other control variables used, firm size, as depicted by the log of the number of employees, bears a significantly positive relation with TFP. This

[15] In future research, an interaction term for R&D and patenting may be included in the regression models, to emphasize and evaluate the value-capturing and profit-making criteria of patents. This is particularly pertinent for the pharmaceutical industry.

signifies that bigger sized firms are more likely to perform better in terms of overall productivity. However the magnitude of the coefficients is below 0.2 which indicates a low elasticity. On the other hand, age is found to have no significant effect on the TFP of the firms, as reflected by their insignificant coefficients and indeterminate signs.

Since more concentration indicates less competition among the firms and vice-versa, the results suggests that an increase in competition greatly influences the TFP in the industry. In the random effect GLS estimations, the absolute value of the coefficients for the log of market share is greater than -0.5 is evidently far higher than the other independent variables used. Moreover, in case of system GMM, its absolute value is greater than unity which implies that TFP increases more than proportionately with competition. This is in conformation to the study by Nickell (1996) who asserted that higher market share has a detrimental effect on TFP. Similar finding was observed in the earlier industry level study by Haskel (1991) and firm-level study by Nickell *et al.* (1992) which revealed similar phenomena. In all these studies, market concentration has been proxied by market share with identical computation strategy.

Table 7.2: Estimation of the productivity function using TFP as the dependent variable

Dependent variable: Log(TFP)	Estimation Method: Random Effect GLS				Estimation Method: System GMM			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
Log(employees)	0.089***	0.072***	0.086***	0.089***	0.123***	0.083***	0.114***	0.123***
	[0.017]	[0.016]	[0.017]	[0.017]	[0.023]	[0.025]	[0.023]	[0.023]
Log(age)	0.011	0.005	0.001	0.012	-0.031	-0.005	-0.029	-0.03
	[0.028]	[0.028]	[0.028]	[0.028]	[0.033]	[0.036]	[0.032]	[0.033]
Log(market share)	-0.624***	-0.523**	-0.647***	-0.626***	-1.305***	-1.392***	-1.308***	-1.315***
	[0.216]	[0.258]	[0.215]	[0.216]	[0.200]	[0.240]	[0.197]	[0.200]
Citation Counts	0.001	0.0002	0.0009		0.036***	0.024***	0.032***	
	[0.001]	[0.0008]	[0.001]		[0.004]	[0.004]	[0.005]	
(Citation Counts)Fitted				0.001				0.036***
				[0.001]				[0.004]
Log(R&D capital)Fitted 1		0.068***				0.111***		
		[0.023]				[0.042]		
Log(R&D capital)Fitted 2			0.086***				0.117**	
			[0.025]				[0.056]	
Log(TFP)[t-1]	0.373***	0.544***	0.367***	0.373***	0.170***	0.291***	0.175***	0.170***
	[0.029]	[0.038]	[0.029]	[0.029]	[0.052]	[0.055]	[0.051]	[0.052]
Intercept	5.641***	4.469***	5.777***	5.657***	10.278***	10.391***	10.232***	10.337***
	[1.286]	[1.542]	[1.281]	[1.285]	[1.225]	[1.469]	[1.206]	[1.224]
Number of Observations	1122	631	1122	1123	1122	631	1122	1123
Wald Chi²	360.780***	530.000***	375.460***	361.000***	330.260***	254.100***	345.180***	332.340***
AR(2)					1.51	1.3	1.46	1.51

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

Our regression provides positive and significant coefficients for lagged TFP values and the constant terms. Taking into account the asymptotic properties of the random effect GLS and the system GMM estimators, the χ^2 value reports that our coefficients are highly significant when considered jointly.[16]

Using labour productivity as the dependent variable

Next, we proceed to analyze the estimation results obtained when labour productivity is used as the response variable. The results of the estimations are provided in table 7.3. Concerning the innovation measures, citation counts and fitted citation counts provide identical results. In general, regression 1 and 4 exhibits comparable results, similar to Table 7.2. This again stresses on the fact that not much alteration was achieved when predicted citation counts, conditioned on R&D intensity and others firm characteristics was used. In regard to the regression results 2 and 3, citation counts are no longer significant when random effect GLS estimator is used. In addition, consistently very low magnitudes are observed for coefficients of citation counts in this first estimation technique. However, contrary to the random effect GLS, the system GMM estimations provide highly significant estimated coefficients for citation counts in their original as well as fitted values. Despite the disparity in the nature of the coefficients for citation counts in the two estimation methods used, the values for fitted R&D coefficients are found to be analogous with highly significant positive values. Furthermore, the coefficients for R&D intensity are consistently high in magnitude which proclaim that the R&D measure has a prominent impact on labour productivity.

[16] The Sargan test for overidentification has provided inconsistent results in our estimations. We have omitted the Sargan test from the estimation table because statistical literatures (e.g. Pitt, 2011) suggest that, with the xtabond2 command in STATA, the Sargan test does not bear any statistical validity from the perspective of statistical theory. According to Pitt(2011), the built-in command for system GMM by David Roodman, with this additional test can provide erroneous results causing an over-rejection of the null hypothesis due to the over-identifying restrictions that this test holds. Even Roodman (2009) exclaimed that the estimation results should not rely on the Sargan test, since it is a weak indicator, and susceptible to errors. Also, in case of one-step system GMM, the Sargan test becomes inconsistent with the non-sphericity of the error terms.

However, a contrasting finding is observed between firm size and labour productivity. A systematically negative and significant coefficient for firm size is obtained, using both the regression techniques. Although this outcome is not a general expectation and directly refutes most of the literatures on the same (like Leung *et al*., 2008), it is possible to draw a logical explanation on our result. The outcome entails a decrease in output per unit of labour with an increase in the number of employees in the Dutch pharmaceutical industry. Stated differently, a noticeable loss of production efficiency is observed when firm size increases.

Further, the variable age still does not provide any definitive inference on its effect on labour productivity. However it is found to be significant in regressions 1 and 4 using both the estimators, albeit opposite signs for the corresponding magnitudes. A consistently negative and highly significant effect of market share on productivity is observed again in the system GMM estimations, establishing the fact that a positive relation between competition and productivity in the Dutch pharmaceuticals exist. This finding backs the general consensus that competition instigates productivity (Bridgman, 2010).

Lastly, in consonance with our previous findings, highly significant and positive estimates are obtained for lagged labour productivity and the intercept terms. Additionally, the Wald χ^2 suggest our models to be robust.

Table 7.3: Estimation of the productivity function using labour productivity as the dependent variable

Dependent variable: Log(Labour productivity)	Estimation Method: Random Effect GLS				Estimation Method: System GMM			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
Log(employees)	-0.171***	-0.075***	-0.155***	-0.171***	-0.090***	-0.076**	-0.077***	-0.090***
	[0.022]	[0.026]	[0.020]	[0.022]	[0.027]	[0.034]	[0.020]	[0.027]
Log(age)	0.085**	0.03	0.019	0.085**	-0.110*	-0.1	-0.009	-0.111*
	[0.040]	[0.050]	[0.038]	[0.039]	[0.058]	[0.065]	[0.045]	[0.058]
Log(market share)	-0.377	-0.296	-0.616***	-0.38	-1.473***	-1.365***	-1.346***	-1.482***
	[0.263]	[0.382]	[0.249]	[0.263]	[0.278]	[0.369]	[0.219]	[0.278]
Citation Counts	0.004*	0.0004	0.001		0.140***	0.084***	0.076***	
	[0.002]	[0.002]	[0.002]		[0.010]	[0.009]	[0.007]	
(Citation Counts)Fitted				0.004*				0.140***
				[0.002]				[0.010]
Log(R&D capital)Fitted 1		0.368***				0.353***		
		[0.035]				[0.066]		
Log(R&D capital)Fitted 2			0.371***				0.393***	
			[0.029]				[0.046]	
Log(Labour productivity)[t-1]	0.200***	0.184***	0.195***	0.200***	0.062*	0.098***	0.092***	0.061*
	[0.024]	[0.033]	[0.023]	[0.024]	[0.036]	[0.036]	[0.028]	[0.036]
Intercept	5.697***	4.753**	6.968***	5.715***	12.530***	11.385***	11.267***	12.589***
	[1.561]	[2.257]	[1.475]	[1.559]	[1.641]	[2.160]	[1.288]	[1.640]
Number of Observations	1312	697	1312	1313	1312	697	1312	1313
Wald Chi²	116.860***	177.530***	293.980***	117.000***	228.540***	190.950***	319.770***	228.600***
AR(2)					2.3	1.83	1.96	2.3

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

7.5.2 The production function approach

As has been described in the subsection 7.4.2, we present in Table 7.4 the empirical findings using the three distinct empirical modeling techniques, namely, the Random effect GLS, the system GMM and the Levinsohn-Petrin estimator. Each of the models consists of four regressions estimates, with the variations based on the regressors used. Regression 1 in each model represents the basic production function equation with labour and capital as the only factor inputs to determine the level of output. From regression 2 onwards, R&D investment is used as an additional input. However, in regression 2, we only use a dataset where R&D expenditure variable is balanced. This is evident from the number of observations for regression 2, where it is conspicuously lower than the other regression models for obvious reasons. But in regression 3 and 4, we use the R&D fitted values from a Tobit II estimation, wherein a modified fitted R&D is used in the latter case, in which the missing values after fitting the predicted outcomes for R&D expenditure is substituted by zero. Thus, in regression 4, the number of observations increases in respect to regression 3. In all the estimations, the variables are expressed in their logarithmic forms. Considering the various forms of R&D data that we have used, we proceed with the discussion on the empirical analysis.

Highly significant coefficients for labour and capital inputs are obtained consistently in all our regression results. Besides a few exceptions, the magnitude of the coefficients for capital is found to be greater than labour, exhibiting that the level of output may depend more on their physical capital, than the number of employees. Regarding R&D, which acts as a supplementary factor input, the balanced R&D capital is not significant in case of random effect GLS and the semi-parametric Levinsohn-Petrin estimator. This may be attributed to the low observations, which does not portray the true R&D investment in the Dutch pharmaceutical sector. However the coefficient for balanced R&D in the system GMM method is found to be significant. The fitted R&D measures are found to have a significant effect on output, as exemplified by its coefficients. Although the

magnitudes of the R&D coefficients are prominently lower than the other inputs, their significant outcomes do indicate that their effect on the overall output of the pharmaceutical firms cannot be overlooked. In case of the system GMM estimations, the coefficient for the lagged dependent variables, as one of the regressors, depicts systematically positive and significant values. Similarly, the intercepts, as obtained for the first two regression models, provides significant coefficients.

Altogether, the estimation results in the production function approach annotate the effective relation between the production inputs and production output, reflecting the potential driving sources that lead to productivity and overall economic growth in this particular sector. In addition, the empirical exposition of R&D as a factor input clearly reflects its significant impact on the output level of the Netherlands' pharmaceuticals.

Table 7.4: Estimation of the Production Function

Dependent variable: Value-added	Estimation Method: Random Effect GLS				Estimation Method: System GMM				Estimation Method: Levinsohn-Petrin			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
Log(employees)	0.221***	0.207***	0.298***	0.250***	0.447***	0.124***	0.287***	0.435***	0.440***	0.256***	0.394***	0.448***
	[0.026]	[0.030]	[0.031]	[0.027]	[0.051]	[0.040]	[0.042]	[0.052]	[0.052]	[0.061]	[0.064]	[0.051]
Log(capital)	0.464***	0.628***	0.531***	0.451***	0.324***	0.385***	0.394***	0.326***	0.339***	0.603***	0.480***	0.336***
	[0.018]	[0.025]	[0.022]	[0.018]	[0.036]	[0.034]	[0.034]	[0.040]	[0.090]	[0.104]	[0.125]	[0.071]
Log(R&D capital)!=0		0.004				0.081**				0.019		
		[0.023]				[0.035]				[0.026]		
Log(R&D capital)Fitted 1			0.102***				0.067*				0.069*	
			[0.023]				[0.033]				[0.040]	
Log(R&D capital)Fitted 2				0.113***				0.126**				0.087*
				[0.024]				[0.056]				[0.045]
Log(Value added)[t-1]					0.196***	0.356***	0.299***	0.191***				
					[0.042]	[0.040]	[0.041]	[0.045]				
Intercept	3.381***	2.748***	3.849***	3.297***	2.189***	1.499***	1.474***	2.171***				
	[0.123]	[0.218]	[0.170]	[0.122]	[0.177]	[0.202]	[0.161]	[0.184]				
Number of Observations	1552	441	805	1552	1298	376	699	1298	1519	440	799	1519
Wald Chi²	1348.570***	1337.390***	1620.480***	1406.550***	3201.520***	4882.000***	5113.980***	2910.420***				
AR(2)					1.46	2.63	1.09	1.42				
Wald test of CRS(Chi²)									8.770***	1.66	0.19	2.15

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

7.6 Conclusion

This chapter presents a comprehensive investigation on the essential pathways that lead to effective output and productivity growth, giving special impetus on the outcomes due to the innovation processes undertaken in the Netherlands' pharmaceuticals.

Based on our statistical data descriptive, the innovative pharmaceuticals are found to outperform their non-innovative counterparts in terms of not only their reserve in factor inputs, but also their consequent overall output. Likewise, it is also established that the innovative firms have a dominance in terms of their total factor productivity or labour productivity. This outcome is identical for both innovation measured using R&D intensity or citation-weighted patents.

However, to effectively delineate the divergence in the effect of the two alternative measures of innovation on productivity, we resort to our econometric estimations. Our empirical results suggest a significantly prominent impact of innovation that is picked up by the R&D intensity measure along with the citation-weighted patents. Although the divergence of the measurement approach of the two different innovation measures does not allow a comparison of their coefficients, the consistency in the significance level of R&D intensity is apparently more pervasive in both the estimation techniques compared to the citation-weighted patents. This substantiates that the innovation performance of the firms, in the form of patent quality does not essentially capture the effect of innovation on productivity fully. On the other hand, R&D intensity has a noticeable relevance for productivity in the Dutch pharmaceutical sector. Therefore, the proposition by Danzon and Percy (1999) that, R&D investment plays an important role in the productivity of the pharmaceutical firms holds true. Simultaneously innovation output is also evidenced to be enviable for the productivity in this sector, although it may be comparatively less impactful than the innovation captured using R&D expenditure.

Hence, from a conceptual notation, R&D intensity can directly affect the productivity of the firms. This supports the fact that, not all innovations lead to patenting as many pharmaceuticals patent their innovations mostly for strategic and economic reasons. Plus, many firms circumvent the process of patenting due to its tedious nature, which involves huge time and cost. This arguments against the CDM model (Crepon *et al*. 1998), which builds upon a recursive 3 stage estimation technique, whereby R&D investment bears an indirect effect on productivity via patenting. However, our findings are construed within the periphery of the pharmaceutical sector in the Netherlands and may vary in case of other industries where innovation output may be calibrated using other measures, like innovative sales.

In the context of a production function framework, R&D expenditure is also revealed to have a positive influence on the level of output in the production process, as an essential factor input. This claims that the R&D investment is indispensable in the production process of the pharmaceutical industry, along with physical capital and labour inputs. In essence, our study confirms a significant direct effect of R&D investment on pharmaceutical output and productivity.

Regarding the firm-size and productivity linkage, bigger firms are found to positively influence total factor productivity. However, the output per unit labour is seen to decrease with the increase in firm size. In other words, the productive efficiency of the workers diminishes with higher employment at the firm level. This finding might provide an interesting avenue for further research on this issue, by demarcating the skilled and unskilled workers and investigating on their role in the pharmaceutical industry.

Furthermore, our analysis largely provides evidence of higher competition leading to greater productivity. A positive effect of competition on productivity, by using market share as the concentration index, is broadly confirmed in the studies by Haskel (1991) and Nickell (1996), among others. In consideration of the logical

plausibility, it is of a general consensus that higher competition among the firms provides greater incentive to increase their productivity levels (as in Bridgman, 2010 and Aghion *et al.*, 2007).

However, keeping parity with most of our earlier findings, this study further asserts that the age of the pharmaceuticals bear little or no significance in the productivity performance of the Dutch pharmaceuticals. This stems from the fact that new firms can be the result of possible spin-offs, which already possess adequate tangible and intangible reserves. In addition, the influx of venture capital markets provides prominent financial support for budding firms to compete with their already established counterparts.

On a whole, this study brings forth prominent findings that may provide relevant and essential policy levers, which will be discussed in our final chapter. Nevertheless, it can be prominently inferred from this study that innovation plays a significant role in the overall productivity of the Dutch pharmaceutical industry, although the impact of patenting activity is not strikingly radical. In addition, bigger firms with higher competitive conduct are found to maintain higher drug productivity. However the effect of competition has not been dealt with in greater details in this chapter. Hence further research can be carried out to indicate the effect of alternative measures of competition on the productivity in the Dutch pharmaceutical industry.

A.7 APPENDIX

A.7.1 The ZINB model from which the predicted citation counts are obtained

Table A.1: ML-regression results using citation counts

Dependent Variable	Forward Citation Counts
	ZINB
Log(R&D per employee)Fitted	0.186*
	[0.100]
Log(Employees)	0.423***
	[0.102]
Age	-0.065
	[0.431]
Entry	-3.330***
	[0.780]
Exit	-0.547*
	[0.291]
Dummy(Patent)	0.655**
	[0.310]
Intercept	2.767***
	[0.829]
Time Dummy	YES
Initial(Patent)	0.005***
	[0.001]
Random Effects	YES
Log likelihood	-839.489
N Observations	1196
Nonzero observations	169
Zero observations	1027

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

Chapter 8

CONCLUDING REMARKS

8.1 Synopsis

This PhD dissertation, in its entirety, circumscribes the different aspects of the pharmaceutical industry; whereby their innovation, competition and productivity performance is interwoven to arrive at a logical inference. The pharmaceutical industry is not only fundamental for the overall economic growth of a country, but it is indispensable from a social perspective. Bearing the task of providing single-handedly medicinal cure for saving human lives and improving health conditions, this sector undoubtedly holds its preeminence over other industries. However, the discovery and final dissemination of medicines is a complex and creative procedure, where the inventive performance plays a crucial role. Based on this purview, our research analyses successfully canvas the crucial attributes in the innovation process of this sector, and the accompanied characteristics that define their economic performance.

Our research work in this thesis is essentially concentrated on the evidences from the pharmaceutical sector in the Netherlands. However, as a prelude to our main topic of discussion, chapter 1 and 2 provides a comprehensive overview of the pharmaceutical industries operating in the world economy, with a special emphasize on the Netherlands. In a way, these studies, based on a review of past literatures and statistical evidences bear a comparative approach from the viewpoint of the Netherlands' drug sector. Chapter 1 essentially deals with the status of the Netherlands' pharmaceutical industry, in regard to its neighboring European counterparts. It is asserted that although the drug sector in the Netherlands performs considerably well, it stands at a modest position with

reference to the other more established European countries. Nonetheless, the incessant uprising in this Dutch sector is evident from the statistical data.

Concurrently, chapter 2 embarks upon a more wider view, where the countries outside Europe come to prominence. This chapter theoretically investigates on the pharmaceutical industry in the global platform, where the USA is found to hold its sway in terms of commercial success as well as inventive expertise. Next in line, remains Japan and the European continent. In addition, focusing on the developing countries, we have explored the pharmaceutical regime in India, since it stands as the biggest pharmaceutical exporter in the Netherlands, amongst all the developing countries. This chapter also portrays the huge diversification in the pharmaceutical operations, prevailing in the different economic scenarios. Although the developing countries like India mostly resort to bulk drug production, the innovation incentives in these countries still lags far behind. However the latest policy enforcements entitles most developed countries to perform product and process innovation and therefore, take active participation in the novel invention which is authenticated by the international intellectual property rights. Although this acts as a steppingstone for the betterment of this sector in the developing arena, the accompanying wrath cannot be completely overlooked. This new policies for the introduction of pharmaceutical patents have been vehemently criticized by several economic researchers (like, Siebeck, 1990 and Maskus, 2000), as the developing economies lack the essential financial infrastructure to carry out the complicated and ambitious inventive process. Nonetheless, it cannot be gainsaid that this policies have triggered novel innovation in the developing pharmaceuticals, which holds optimism for the future. Statistical evidence suggests a substantive improvement in the patenting performance in the Indian pharmaceutical sector. At the same time, the infux of many multinational pharmaceuticals that has led to mergers have also put forward the Indian pharmaceutical industry at a better position (Srivastava, 2001; and Balakrishnan *et al.,* 2000). On the other hand, the developed countries engage in conspicuous medicinal research and therefore, the predominant area of

our study, that is based on the innovation phenomena of the Netherlands pharmaceuticals, holds pertinence.

The theoretical review in the first two chapters has been inherently built on the three main features that we focus upon in this thesis, viz., innovation, competition and firm performance. Conforming to this structural framework, our study concentrates on the micro-econometric analysis for the Netherland's pharmaceutical sector. For this, an extensive micro-level data has been accrued and compiled. The data sources, its extraction process and the final computation of the different variables have been enumerated in chapter 3. As pointed by Griliches and Mairesse (1991), micro level data can provide a more accurate result than the macro level ones, with more appropriate data structure as it evades any aggregation problem. However, the detrimental issue for firm level data is the lack of adequate number of observations. In addition, this type of data is also subject to higher rate of error, which gets averaged in case of macro level data. Nonetheless, using proper estimation techniques, we have been able to circumvent the concomitant problems associated with our firm level data. With the compilation of our dataset, we have been able to formulate a panel data for the Netherlands' pharmaceutical sector, spanning over the period of 1996-2006. This dataset has been subsequently used for our exhaustive empirical analysis undertaken in the succeeding chapters.

Chapter 4 inspects on how the various firm-level determinants affect the propensity of the Dutch pharmaceuticals to engage in R&D activities. Investment in R&D is an essential criterion for the research intensive pharmaceutical industry. Owing to a lack of observations for the R&D data, a Heckman's generalized Tobit II estimation technique has been adopted. Our primary empirical findings assert that smaller sized pharmaceuticals, having a sufficient reserve of tangible capital and enjoying a less competitive environment are more inclined to engage in R&D expenditure.

It is de-facto expected that there exists a substantial synergy between the R&D investment and the innovation performance through patenting, which is proved to be true in chapter 5. It is empirically evidenced that R&D intensity is a major determinant for generating patents. However, these two innovation processes do not necessarily call for complementary effects on the other determinants. They can provide with considerably different attributes, which is anticipated, in a way, as they are two different facets of innovation. From the perceptions of our empirical findings, although the smaller budding pharmaceuticals are more susceptible to engage in R&D (as observed in chapter 4), patenting practices are usually undertaken by the larger firms who have the adequate human resource and financial capabilities to expend through the tumultuous and risky process of patenting. In this regard, sequential innovation cannot be neglected, where large firms with higher established power engage in strategic patenting process to evade competition in the market. Besides, chapter 5 also portrays that prominent barriers to entry and exit can lead to greater patenting performance by the Dutch pharmaceuticals. For this analysis in chapter 5, a ZINB estimation technique has been applied, due to the positive and non-negative patent data with many zero values. Our results hold robust when alternative response variable, in the form of citation-weighted patent counts, or its bifurcation into EPO and USPTO patents (and citations), has been used.

This study on the effective relation of the various determinants on innovation performance is furthered in chapter 6, with a more in-depth view on the competition aspect. The different competition measures put forth in this study leads to varied conceptual grounds in their anecdotal evidences. Over the period of 1996-2006, the concentration in the Dutch pharmaceuticals is seen to have decreased, which bears congruence with the entry of many generic pharmaceutical companies. However the mark-up with the adjustment for scale elasticities and the profit elasticity reveals a decreasing competition, with an increased level of profitability. This may hint at a reallocation effect, where lesser competition has led to higher

influx in the market. Moreover, a detailed empirical analysis is attempted in this chapter, using different econometric modeling tools based on count data (which includes ZINB, Hurdle and PPML estimation techniques), and conditioning on several additional controls. Besides, special attention is given for the endogeneity issue, by incorporating an exogenous regulation indicator and using lagged competition measures in the estimations. Our findings suggest a negative or a U-shaped relation between competition and innovation performance, and no sign of a positive relation between them could be ascertained. Hence competition in the Dutch pharmaceutical market either negatively affect the patenting performance for the entire phase, or for the initial phase until it reaches a threshold point.

The findings in chapter 6 consolidate our prior establishment in chapter 5, in which barriers to entry and exit proclaimed competition to be a disincentive for patent performance. Hence, firms who aim at patenting their innovations would create a potential barrier to ward off competition from the market. On the other hand, it is seen in chapter 7 that lesser concentration and hence, greater number of firms in the market is essential for the overall productivity that leads to effective growth in this sector. Hence competition is imperative for the overall growth of this sector, and hence, its absence can cause prominent market inefficiencies. This leads us to question whether patenting is really desirable. Concurrently, chapter 7 also affirms that the innovation captured through R&D investment has a prominent impact on productivity along with patenting performance (a detailed estimation is carried out in this regard using random effect GLS and system GMM estimation techniques). Therefore, the empirical findings suggest that, not all pharmaceutical innovations are patented, and the innovation captured through R&D expenditure can propagate pharmaceutical productivity effectively. This intrinsically leads us to question whether patenting is really desirable. Although it cannot be denied that intellectual property protection is a prerequisite for introducing novel therapeutic drugs in the market, it is of utmost importance to effectively channelize this strategy for a more efficient market structure and flexible pricing policy. Hence, policymakers should

give more impetus to the social drawbacks of patents and consider their possible alternatives (as suggested by Hubbard and Love, 2004), especially for the growing economies where access to drugs largely depend on the pharmaceutical pricing.

8.2 Policy implications

In view of the structural summarization of the dissertation, we aim at providing certain policy perspectives. Our micro-econometric evidence from chapter 4 brings to light the upsurge of venture capital markets that play a prominent role of encouraging the smaller and newly established firms to perform R&D investments. This also instigates the possibility of entrepreneurship and spin-offs in this Dutch sector. Hence, this argument favors policies like size-dependent R&D taxation or patent regulations.

A concerted indication from chapter 4, 5 and 6 suggests that the intellectual property protection adopted by the innovative pharmaceuticals to safeguard their innovations, buffers competition. This might hinder the social benefits at large, by creating a tendency towards a non-competitive regime and consequently raising the drug price level. This aptly justifies the propagation of managed competition in the Dutch Health Care Reform Act of 2006, and the promotion of smaller sized pharmaceuticals as these firms are instrumental to control the drug prices. However, it is fundamental for the policy makers to monitor patenting and licensing practices, and provide free access to basic inventions.

At the same time, to mitigate the adversity due to patenting, policy makers should give priority to alternative strategies that can pay off the huge R&D expenditure procured for the discovery of novel therapeutic drugs. A study by DiMasi and Grabowski (2004) advocated two alternative policies to patenting. This includes a compensation through financial award after the completion of a successful innovation or the participation of the government for the procurement of R&D expenditure by contracting grants. Although these plausible substitutes for

patenting look theoretically prudent, they are difficult to implement against the long-standing establishment of the system of patenting. However the government should frame their policies tactfully and cautiously such that the short term inefficiencies offered by the patenting system can be minimalized. This is especially pertinent for the developing countries, where the hindrance due to patenting can be more drastic. However Philipson and Mechoulan (2003) conjectured that the disincentives due to patenting remains only for the short-run and is consequently obliterated by the apparent benefits that pharmaceutical patents offer.

Chapter 7 reaffirms that, although patenting is an inherent characteristics for pharmaceuticals to safeguard their innovations, its effect on productivity is noticeable but not immensely substantial. Also, a competitive environment proves to be highly beneficial for the productivity in the Dutch pharmaceuticals. Therefore reducing regulatory practices might have a favorable effect on the productivity growth in this sector.

Nonetheless, chapter 7 further asserts that, knowledge input and innovation activity is crucial for this research-intensive industry, as it provides the foundation for scientific advancement to promote productivity. Hence encouraging higher R&D investments, through venture capital markets and governmental aid can further propagate productivity and growth in this sector.

8.3 Further research

Our study mainly provides a comprehensive and detailed inspection of the Netherlands' pharmaceutical sector, based on econometric and theoretical justifications. However there are numerous avenues through which this study can be further explored.

Although a global pharmaceutical overview is provided in the initial introductory chapters, they are solely based on the review of foregoing surveys. Hence empirical

investigations are only ventured for the Dutch pharmaceutical industry. Since our study is based on a single sector and a single country, there is no scope for empirical comparison with other countries or other industries. Hence, as a later step to this research work, this analysis can be furthered and diversified among other industries to have a better knowledge on how the various determinants and their causal effects are functional in other industrial regime. Moreover, a systematic digression with the pharmaceutical industry of other countries would provide an essential stepping stone for structural and policy implications in a broader perspective.

In the era of globalization, another important aspect that we have not dealt with in this paper is, the competitive edge at which the Netherlands' pharmaceutical industry operates in respect to other countries at its vicinity or otherwise. In addition to considering their differences in institutional framework and policy regimes that varied from the Netherlands, another important criterion that should essentially be looked upon is the import and export practices that the Netherlands' drug sector undertakes.

Placing a more microscopic view on our discussed chapters, Chapter 6 can be further extended by investigating on how the various competition measures affect the R&D intensity, instead of patenting performance, in a non-linear framework. It would be intriguing to scrutinize if there exists an alternative outcome, rather than what we perceived by using innovation output as the response variable. Also, chapter 7 can be developed into greater details to accommodate the effect of the various competition indicators on the productivity of the Dutch pharmaceutical industry.

Focusing on the empirical investigations, a pre sample mean-quasi differencing fixed effect technique (following, Blundell *et al.*, 1999) may be used as an alternative to maximum likelihood approach (following, Wooldridge, 2005) to control for unobserved heterogeneity at the firm level. However to incorporate the

former methodology, information on the pre-sample innovation data is required that replaces the unobserved firm fixed effects.

In addition to accruing data for previous years, an extension of our data sample beyond 2006 may provide an interesting avenue to examine and identify the exogenous shocks that might occur with the emergence of the Dutch Healthcare Reform Act in 2006. This is particularly pertinent for competition analysis, since the radical changes in the healthcare system with the introduction of a regulated market, might lead to new inferences for the Dutch pharmaceutical sector.

Lastly, the innovation mechanism of the pharmaceutical industry is an interesting avenue of research and can be further investigated, based on the innovation management and strategies. The innovation practice in this sector is unique and intriguing, that calls for further investigation on the pharmaceutical research in practice and its corresponding state-of-the-art. Emphasizing on the sequential patenting largely practiced by innovative pharmaceuticals, it results in prominent hindrances to competition by preventing the influx of generic firms. However the high risk ventures of drug development and definite shelf-life of patents entail the pharmaceuticals to engage in optimal strategic choices. Hence further studies may be devoted to the portfolio management and policy perspectives, in view of the market characteristics, that lead to short-term profits and long-term development for the pharmaceutical industry.

BIBLIOGRAPHY

Abelson, P. (1993), "Improvements in Health Care", *Science*, **260(5104)**, 11.

Abrol, D., P. Prajapati and N. Singh (2011), "Globalisation of the Indian Pharmaceutical Industry: Implications for Innovation", *International Journal of Institutions and Economies*, **3(2)**, 327-365.

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

Acs, Z. J. and D. B. Audretsch (1989), "Patents as a Measure of Innovative Activity", *KYKLOS*, **42(2)**, 171-180.

Acs, Z. J. and D. B. Audretsch (1988), "Innovation in Large and Small Firms", *The American Economic Review*, **78(4)**, 678-690.

Acs, Z. J. and D. B. Audretsch (1987), "Innovation, Market Structure and Firm Size", *Review of Economics and Statistics*, **69(4)**, 567-574.

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

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

Aghion, P., M. Braun and J. Fedderke (2006), "Competition and productivity growth in South Africa", CID Working Paper No. 132.

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.

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

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

Akerlof, G. (1970), "The Market for "Lemons": Quality Uncertainty, and the Market Mechanism", *The Quarterly Journal of Economics*, **84(3)**, 488-500.

Allison, M., A. Lemly, K. Moore and R. Trunkey (2009), "Valuable patents", *The Georgetown Law Journal*, **93**, 435-479.

Amable, B., L. Demmou and I. Ledezma (2008), "Competition, Innovation and Distance to Frontier", CES Working Paper.

Amoroso, S., P. Kort, B. Melenberg, J. Plasmans and M. Vancauteren (2010), "Firm level productivity under imperfect competition in output and labour markets" CESifo Working Paper Series No. 3082.

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.

Arellano, M. and O. Bover (1995), "Another look at the instrumental variable estimation of error-components models", *Journal of Econometrics*, **68(1)**, 29-51.

Artés, J. (2009), "Long-run versus short-run decisions, R&D and market structure in Spanish firms", *Research Policy*, **38(1)**, 120-132.

Arundel, A. and I. Kabla (1998), "What percentage of innovations are patented?: Empirical estimates for European firms'', *Research Policy*, **27(2)**, 127-141.

Askenazy, P., C. Cahn and D. Irac (2008), "Competition, R&D and the cost of innovation", PSE Working Papers 2008-32, PSE (Ecole normale supérieure), Paris.

Atella, V. and B. Quintieri (2001), "Do R&D expenditures Really Matter for TFP?", *Applied Economics*, **33(11)**, 1385-89.

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

Badinger, H. (2007), "Has the EU's Single Market Programme Fostered Competition? Testing for a Decrease in Mark-up Ratios in EU Industries", *Oxford Bulletin of Economics and Statistics*, **69(4)**, 497-519.

Balasubramanian, N. and J. Lee (2008), "Firm age and innovation", *Industrial and Corporate Change*, **17(5)**, 1019-1047.

Balakrishnan, P., K. Pushpangadan and M. S. Babu (2000), "Trade liberalisation and productivity growth in manufacturing: Evidence from firm-level panel data", *Economic and Political Weekly*, **35(41)**, 3679-3682.

Bartus, T. (2005), "Estimation of marginal effects using margeff", *The Stata Journal*, **5(3)**, 309-329.

Basu, S. and J. G. Fernald (2002), "Aggregate productivity and aggregate technology", *European Economic Review*, **46(6)**, 963-991.

Battese, G. E., D. S. P. Rao (2002), "Technology potential, efficiency and a stochastic metafrontier function", *International Journal of Business and Economics*, **1(2)**, 87-93.

Bhattacharjea, A. and F. Sindhwani (2014), "Competition Issues in the Indian Pharmaceutical Sector", Centre for Development Economics, Delhi School of Economics.

Belderbos, R., M. Carree and B. Lokshin (2004), "Cooperative R&D and firm performance", *Research Policy*, **33(10)**, 1477-1492.

Benavente, H. 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.

Beneito, P., M. E. Rochina-Barrachina and A. Sanchis (2014), "Patents, Competition, and Firms' Innovation Incentives", *Industry and Innovation*, **21(4)**, 285-309.

Bevan, G. and W. P. M. Van de Ven (2010), "Choice of providers and mutual healthcare purchasers: can the English National Health Service learn from the Dutch reforms?", *Health Economics, Policy and Law*, **5(3)**, 343-363.

Bikker, J. A. and M. Van Leuvensteijn (2008), "Competition and efficiency in the Dutch life insurance industry", *Applied Economics*, **40(16)**, 2063-2084.

Bishop, S. and M. Walker (2002), "The Economics of EC Competition Law: Concepts, Applications and Measurement", 2nd ed., London: Sweet and Maxwell Publishers.

Blundell, R., R. Griffith and F. Windmeijer (2002), "Individual effects and dynamics in count data models", *Journal of Econometrics*, **108(1)**, 113-131.

Blundell, R. and S. R. Bond (2000): "GMM Estimation with Persistent Panel Data: An Application to Production Functions", *Economics Reviews*, **19(3)**, 321-340.

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(3)**, 529-554.

Blundell, R. and S. R. Bond (1998), "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics*, **87(1)**, 115-143.

Blundell R., R. Griffith and J. V. Reenen (1995), "Dynamic Count Data Models of Technological Innovation", *The Economic Journal*, **105(429)**, 333-344.

Boone, J. (2000), "Competition", CEPR Discussion Paper 2636, CEPR.

Boone, J., J. Van Ours and H. Van der Wiel (2007), "How (not) to measure competition", CPB Discussion Paper 91, CPB.

Boone, J. (2008), "A New Way to Measure Competition", *The Economic Journal*, Royal Economic Society, **118(531)**, 1245-1261.

Boonen, L. H. H. M., S. A. Van der Geest, F. T. Schut and M. Varkevissern (2010), "Pharmaceutical policy in the Netherlands: From price regulation towards managed competition", in: Avi Dor ed., Pharmaceutical Markets and Insurance Worldwide (Advances in Health Economics and Health Services Research, vol. 22), Emerald Group Publishing Limited, 53-76.

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.

Braila, C., G. Rayp and S. Sanyal (2010), "Competition and regulation, Belgium, 1997-2004", Federal Planning Bureau Working Paper 3-10.

Branstetter, L. G. and J. R. Chen (2006), "The impact of technology transfer and R&D on productivity growth in Taiwanese industry: Microeconometric analysis using plant and firm-level data", *Journal of the Japanese and International Economies*, **20(2)**, 177-192.

Bridgman, B. (2010), "Competition and Productivity", Bureau of Economic Analysis.

Bureau of Census (2007), "Statistics of U.S. Businesses: NAICS 3254".

Cameron, A. C. and P. K. Trivedi (2013), "Count Panel Data", *Oxford Handbook of Panel Data Econometrics*, B. H. Baltagi ed., Oxford University Press, (forthcoming).

Cameron, A. C. and P. K. Trivedi (2009), "Microeconometrics Using Stata", A Stata Press Publication, StataCorp LP, College Station, Texas.

Cameron, A. C. and P. K. Trivedi (1998), "Regression Analysis of Count Data", New York: Cambridge University Press.

Cameron, A. C. and P. K. Trivedi (1986), "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests", *Journal of Applied Econometrics,* **1(1)**, 29-53.

Caves, R. E. and M. Porter (1978), "Market Structure, Oligopoly, and Stability of Market Shares", *The Journal of Industrial Economics*, **26(4)**, 289-313.

CBI Market Survey (2010), "The pharmaceutical products market in the Netherlands", CBI Market Information Database.

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

Cefis, E. and L. Orsenigo (2001), "The Persistence of Innovative Activities: A Cross- Countries and Cross- Sectors Comparative Analysis," *Research Policy*, **30(7)**, 1139-1158.

Chamberlain, G. (1984), "Panel Data", in: Handbook of Econometrics, II, edited by Zvi Griliches and Michael Intriligator, Amsterdam: North Holland.

Chen, Y. (2000), "Strategic Bidding by Potential Competitors: Will Monopoly Persist?", *Journal of Industrial Economics*, **48(2)**, 161-175.

Cincera, M. (1997), "Patents, R&D and Technological Spillovers at the Firm level: Some Evidence from Econometric Count Models for Panel Data", *Journal of Applied Econometrics*, **12(3)**, 265-280.

Cockburn, I. M. (2007), "Is the Pharmaceutical Industry in a Productivity Crisis?", *Innovation Policy and the Economy*, National Bureau of Economic Research, **7**, 1-32.

Cockburn, I. M. (2004), "The changing structure of the pharmaceutical industry", *Health Affairs*, **23(1)**, 10-22.

Coe, D. T. and E. Helpman (1995), "International R&D spillovers", *European Economic Review*, **39(5)**, 859-887.

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

Cohen, W. M., R. R. Nelson and J. Walsh (2000), "Protecting their Intellectual Assets: Appropriability Conditions and Why U.S. manufacturing Firms Patent (or Not)", NBER Working Paper 7552.

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 W. M. and S. 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 S. Klepper (1996b), "A Reprise of Size and R&D", *The Economic Journal*, **106(437)**, 925-951.

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.

Comanor, W. S. (1967), "Market structure, product differentiation, and industrial research", *The Quarterly Journal of Economics*, **81(4)**, 639-657.

Compendium of Patent Statistics (2008), OECD Report, Organization for Economic Co-operation and Development.

Comtrade (2012), UN Comtrade, http://comtrade.un.org/

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

Conway, P. and G. Nicoletti (2006), "Product market regulation in the non-manufacturing industries of OECD countries: measurement and highlights", OECD, Working Paper No. 530.

Corrales-Leal, W. (2007), "Basic concepts and proposals on the use of policy spaces in trade-supported strategies for sustainable development", ICTSD Series on Trade-Supported Strategies for Sustainable Development, International Centre for Trade and Sustainable Development, Geneva.

Correa, J. A. (2011), "Innovation and Competition: An unstable relationship", *Journal of Applied Econometrics*, **27(1)**, 160-166.

Crépon, B., E. Duguet and J. Mairesse (1998), "Research, Innovation, and Productivity: An Econometric Analysis at the Firm Level", *Economics of Innovation and New Technology*, **7(2)**, 115-158.

Crépon, B. and E. Duguet (1996), "Innovation: measurement, returns and competition", INSEE studies in economics and statistics, **1**, 82-95.

Criscuolo, C. and J. E. Haskel (2003), "Innovation and Productivity growth in the U.K.: evidence from CIS2 and CIS3", CERIBA Working Paper, London, U.K.

Cutler, D. M. and M. McClellan (2001), "Is technological change in medicine worth it?", *Health Affairs*, **20(5)**, 11-29.

Czarnitzki, D., K. Kraft and S. Thorwarth (2009), "The knowledge production of 'R' and 'D'", *Economic Letters*, **105(1)**, 141-143.

Danzon, P. M. and A. Percy (1999), "The Effects of Price Regulation on Productivity in Pharmaceuticals", National Bureau of Economic Research, 371-418.

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.

Diewert, E. (1993), "Fisher Ideal Output, Input and Productivity Indexes Revisited", in: Essays in Index Number Theory, E. Diewert and A. Nakamura eds., Elsevier Science Publishers.

DiMasi, J. A. and C. Paquette (2004), "The Economics of Follow-on Drug Research and Development: Trend in Entry Rates and the Timing of Development", *Pharmaceoeconomics*, **22(2)**, 1-14.

DiMasi, J. A. and H. Grabowski (2004), "Patents and R&D Incentives: Comments on the Hubbard and Love Trade Framework for Financing Pharmaceutical R&D", Tufts Center for the Study of Drug Development, Tufts University.

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.

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

Duguet, E. (2006), "Innovation height, spillovers and TFP growth at the firm level: Evidence from French Manufacturing", *Economics of Innovation and New Technology*, **15(4/5)**, 415-442.

Duguet, E. and S. Monjon (2004), "Is Innovation Persistence at the Firm Level? An Econometric Examination Comparing the propensity Score and regression methods", Cahiers de la Maison des Sciences Economiques, Universite Pantheon- Sorbonne, France, Working Paper No. 04075.

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

El-Shinnawy, A. (2009), "Trends of Total Factor Productivity in Egypt;s Pharmaceutical Industry: Evidence from the Nonparametric Malmquist Index Approach", ERF Annual Conference 2009, Cairo, Egypt.

EFPIA report (2008), "The Pharmaceutical Industry in Figures", The European Federation of Pharmaceutical Industris and Associations, www.efpia.eu.

EFPIA report (2012), "The Pharmaceutical Industry in Figures", The European Federation of Pharmaceutical Industries and Associations, www.efpia.eu

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.

Enzing, C., A. Van der Giessen and S. Kern (2004), "OECD Case Study On Innovation: The Dutch Pharmaceutical and Food Biotechnology Innovation Systems", TNO Strategy, Technology and Policy, 1-116.

Espicom (2009), The Pharmaceutical and Medical Device Knowledge Centre.

European Commission, DG Competition (2009), "Pharmaceutical Sector Inquiry", Final Report.

Gambardella, A., L. Orsenigo and F. Pammolli (2000), "Global Competitiveness in Pharmaceuticals. A European Perspective", Enterprise Directorate-General, European Commission.

Gambardella, A. (1995), "Science and Innovation: The US Pharmaceutical Industry during the 1980s", Cambridge Univ. Press, Cambridge.

Gayle, P. (2003), "Market Concentration and Innovation: New Empirical Evidence on the Schumpeterian Hypothesis", University of Colorado Boulder, Center for Economic Analysis, Working Paper 01-14.

Gayle, P. (2002), "Market Structure and Product Innovation", Manuscript, Kansas State University.

Geroski, P. A., V. J. Reenen and C. F. Walters (1997), "How Persistently Do Firms Innovate?", *Research Policy*, **26(1)**, 33-48.

Geroski, P. (1989), "The Effect of Entry on Profit Margins in the Short and Long Run", *Anales d'Economie et de Statistique*, **15/16**, 333-353.

Gourieroux, C., A. Monfort and A. Trognon (1984), "Pseudo-Maximum likelihood: Applications to Poisson Models", *Econometrica*, **52(3)**, 701-720.

Ghose, A. and C. Chakraborty (2012), "Total Factor Productivity Growth in the Pharmaceutical Industry: A Look Using Modern Time Series Approach with Indian Data", *The Journal of Industrial Statistics*, **1(2)**, 250-268.

Gilbert, R. (2006), "Looking for Mr. Schumpeter: Where Are We in the Competition-Innovation Debate?, in: Jaffe, A. B., J. Lerner, and S. Stern eds., Innovation Policy and the Economy, **6**, 159-215.

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

Graham, S. J. H., B. H. Hall, D. Harhoff and D. C. Mowery (2002), "Post-Issue Patent "Quality Control": A Comparative study of U.S. Patent Re-examinations and European Patent Oppositions", NBER Working Paper, 8807, National Bureau of Economic Research, 1-44.

Griffith R., E. Huergo, J. Mairesse and B. Peters (2006), "Innovation and Productivity across four European Countries", *Oxford Review of Economic Policy*, **22(4)**, 483-498.

Griliches, Z. and J. Mairesse (1991), "R&D and Productivity Growth: Comparing Japanese and U.S. manufacturing Firms", University of Chicago Press, 317-348.

Griliches, Z. (1986), "Productivity, R&D and Basic Research at the Firm level in the 1970s", *American Economic Review*, **76(1)**, 141-54.

Griliches, Z. (1979), "Issues in Assessing the Contribution of Research and Development to Productivity Growth", *Bell Journal of Economics*, **10(1)**, 92-116.

Grossman, G. M. and E. Helpman (1990), "Product Development and International Trade", NBER Working Paper No. W2450.

Guevara, H. H., A. Tübke and A. Brandsma (2008), "The 2008 EU Industrial R&D Investment Scoreboard", European Comission Report, 1-99.

Hall, B. H. and J. Lerner (2010), "The Financing of R&D and Innovation", in: Hall, B. H. and N. Rosenberg eds., Handbook of the Economics of Innovation, Elsevier, **1**, 609-639.

Hall, B. H., F. Lotti and J. Mairesse (2009), "Innovation and productivity in SMEs: empirical evidence for Italy", *Small Business Economics,* **33(1)**, 13-33.

Hall, B. H., A. Jaffe and M. Trajtenberg (2005), "Market Value and Patent Citations", *RAND Journal of Economics*, **36(1)**, 16- 38.

Hall, B. H., Z. Griliches and J. A. Hausman (1986), "Patents and R and D: Is There a Lag?", *International Economic Review,* **27(2)**, 265-283.

Hall, D. (2000), "Zero-Inflated Poisson and Binomial Regression with Random Effects: A Case Study", *Biometrics*, **56(4)**, 1030-1039.

Hall, R. E. (1988), "The relation between price and marginal cost in U.S. industry", *Journal of Political Economy*, **96(5)**, 921-947.

Hamberg, D. (1964), "Size of firm, oligopoly, and research: The evidence", *Canadian Journal of Economics and Political Science*, **30(1)**, 62-75.

Harhoff, D., F. Narin, F. M. Scherer and K. Vopel (1999), "Citation Frequency and the Value of Patented Innovation", *Review of Economics and Statistics*, **81(3)**, 511-515.

Hashmi, A. R. (2013), "Competition and Innovation: The inverted U-relationship revisited", *The Review of Economics and Statistics*, **95(5)**, 1653-1668.

Haskel, J. (1991), "Imperfect Competition, Work Practices and Productivity Growth", *Oxford Bulletin of Economics and Statistics*, **53(3)**, 265-279.

Hausman, J. A., B. Hall and Z. Griliches (1984), "Econometric Models for Count Data with an Application to the Patents-R&D Relationship", *Econometrica*, **52(4)**, 909-938.

Hausman, J. A. (1978), "Specification Tests in Econometrics", *Econometrica*, **46(6)**, 1251-1271.

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(4)**, 1705-1745.

Hesketh, S. R. and A. Skrondal (2013), "Avoiding biased versions of Wooldridge's simple solution to the initial conditions problem", *Economic Letters*, **120(2)**, 346-349.

Holmstrom, B. (1989), "Agency costs and innovation", *Journal of Economic Behaviour and Organisation*, **12(3)**, 305-327.

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.

Hubbard, T. and J. Love (2004), "A New Trade Framework for Global Healthcare R&D", *PloS Biology*, **2(2)**, 147-150.

Huergo, E. and J. Jaumandreu (2004), "How Does Probability of Innovation Change with Firm Age?", *Small Business Economics*, **22(3/4)**, 193-207.

Huergo, E. and J. Jaumandreu (2003), "Firms' age, process innovation and productivity growth", *International Journal of Industrial Organization*, **22(4)**, 541-559.

IMS Health Report (2010), Press Release.

Jaffe, A. B. (1986), "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and market Value", *American Economic Review*, **76(5)**, 984-1001.

Jaffe, A. B., M. Trajtenberg and M. S. Fogarty (2000), "Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors", *American Economic Review*, **90(2)**, 215-218.

Jaimovich, N. and M. Floetotto (2008), "Firm Dynamics, Markup Variations, and the Business Cycle", *Journal of Monetory Economics*, **55(7)**, 1238-1252.

Janz, N., H. Lööf and B. Peters (2003), "Firm level innovation and productivity- Is there a common story across countries?", ZEW Discussion Paper No. 03-26, Manheim, Germany.

Janz, N., H. Lööf and B. Peters (2004), "Innovation and productivity in German and Swedish manufacturing firms- Is there a common story?", *Problems and Perspectives in Management*, **2**, 184-204.

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

Kamble, P., S. Ghorpade, R. Kshirsagar and B. Kuchekar (2012), "Progress of the Indian pharmaceutical industry: a shifting perspective", *Journal of Intellectual Property Law and Practice*, **7(1)**, 48-51.

Kamien M. and N. Schwartz (1982), "Market Structure and Innovation", Cambridge University Press.

Kaur, P. (2012), "Mergers in India: Exploiting Financial Synergies", Academic Foundation, New Delhi.

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

Klepper, S. (1996), "Entry, Exit, Growth and Innovation over the Product Life Cycle", *American Economic Review*, **86**, 562-583.

Klette, T. J. and Z. Griliches (1996), "The inconsistency of common scale estimators when output prices are unobserved and endogenous", NBER Working Paper No. 4026.

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.

Klomp, L. and G. V. Leeuwen (2001), "Linking Innovation and Firm Performance: A New Approach", *International Journal of the Economics of Business*, **8(3)**, 343-364.

Klomp, L. and G. V. Leeuwen (1999), "The Importance of Innovation for Firm Performance", Statistics Netherlands.

Konings, J., P. Van Cayseele and F. Warzynski (2001), "The dynamics of industrial mark-ups in two small open economies: Does national competition policy matters?", *International Journal of Industrial Organization*, **19(5)**, 841-859.

Lanjouw, J. O and M. Schankerman (2001), "Characteristics of Patent Litigation: A Window on Competition", *The Rand Journal of Economics*, **32(1)**, 129-151.

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

Leung, D., C. Meh and Y. Terajima (2008), "Firm Size and Productivity", Bank of Canada Working Paper.

Levin, R. C., A. K. Klevorick, R. R. Nelson and S. G. Winter (1987), "Appropriating the Returns from Industrial Research and Development", *Brookings Papers on Economic Activity*, **18(3)**, 783-831.

Levin, R. C., W. M. Cohen and D. C. Mowery (1985), "R&D Appropriability, Opportunity and Market Structure: New Evidence on Some Schumpeterian Hypotheses", *American Economic Review*, **75(2)**, 20-24.

Levinsohn, J. and A. Petrin (2003), "Estimating Production Functions Using Inputs to Control for Unobservables", *Review of Economic Studies*, **70(2)**, 317-341.

Licht, G. and K. Zoz (2000), "Patents and R&D: An Econometric Investigation using Applications for German, European and US Patents by German Companies", in: Encaoua, D., B. Hall, F. Laisney and J. Mairesse eds., The Economics and Econometrics of Innovation, 307-338.

Licht, G. and K. Zoz (1996), "Patents and R&D: an econometric investigation using applications for German, European and US patents by German companies", ZEW Discussion Papers, Center for European Economic Research (ZEW).

Lichtenberg, F. R. (2001), "Are the Benefits of Newer Drugs Worth Their Cost? Evidence from the 1996 MEPS", in: Triplett, J.E., Measuring the Prices of Medical Treatments, 241-251.

Lichtenberg, F. R. (2004), "Public policy and innovation in the U.S. pharmaceutical industry', in: Holtz-Eakin, D. and H. Rosen eds., Entrepreneurship and Public Policy, MIT Press, Cambridge, MA.

Loecker, J. D. and F. Warzynski (2012), "Markups and Firm-Level Export Status", *American Economic Review*, American Economic Association, **102(6)**, 2437-71.

Long, J. S. and J. Freese (2006), "Regression Models for Categorical Dependent Variables Using Stata", 2nd edition. College Station, TX: Stata Press.

Lööf, H. and A. Heshmati (2006), "On the Relationship Between Innovation and Performance: A Sensitivity Analysis", *Economics of Innovation and New Technology*, **15(4/5)**, 317-345.

Lööf, H. and A. Heshmati (2002), "Knowledge capital and performance heterogeneity: A firm-level innovation study", *International Journal of Production Economics*, **76(1)**, 61-85.

Mahlich, J. (2007), "The Japanese Pharmaceutical Industry in Transition: Has Higher Research Orientation Resulted in Higher Market Value?", *Asian Business and Management*, **6**, 75-94.

Mairesse, J. and P. Mohnen (2010), "Using Innovation Surveys for Econometric Analysis", NBER Working Paper No. W15857.

Mairesse, J. and P. Mohnen (2005), "The Importance of R&D for Innovation: A Reassessment Using French Survey Data", *Journal of Technology Transfer*, **30(2/2)**, 183-197.

Mairesse, J., P. Mohnen and E. Kremp (2005), "The Importance of R&D and Innovations for Productivity: A Re-examination in light of the 2000 French Innovation Survey", in: Mairesse, J. and M. Trajtenberg eds., Special issue of *Annales d'Economie et de Statistique* 79-80 in memory of Z. Griliches, 487-528.

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.

Mairesse, J. and B. H. Hall (1996), "Estimating the Productivity of Research and Development in French and US Manufacturing Firms: An Exploration of Simultaneity Issues with GMM Methods", in: Wagner, K. and B. Van Ark eds., International Productivity Differences and their Explanations, Elsevier Science, 285-315.

Malerba, F and L. Orsenigo (1996), "Schumpeterian patterns of innovation are technology- specific", *Research Policy*, **25(3)**, 451-478.

Malmberg, C. (2007), "The effects of institutional change on innovation and productivity growth in the Swedish pharmaceutical industry", CIRCLE Working Paper No. 2007/02, Lund University.

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.

Maskus, K. E., S. M. Dougherty and A. Mertha (2005), "Intellectual Property Rights and Economic Development in China", in Fink, C. and K. E. Maskus eds., Intellectual Property and Development: Lessons from Recent Economic Research, New York: World Bank and Oxford University Press.

Maskus, K. E. (2000), "Intellectual Property Rights in the Global Economy", Institute of International Economics, Washington DC.

Masso, J. and P. Vahter (2008), "Technological innovation and productivity in late-transition Estonia: econometric evidence from innovation surveys", *European Journal of Development Research*, **20(2)**, 240-261.

Mazumdar, M. (2013), "Performance of pharmaceutical companies in India: A critical analysis of industrial structure, firm specific resources, and emerging strategies", Physica Verlag Heidelberg: Springer Publication.

Melitz, M. and G. Ottaviano (2008), "Market size, trade and productivity", *Review of Economic Studies*, **75(1)**, 295-316.

Mestre-Ferrandiz, J., J. Sussex and A. Towse (2012), "The R&D cost of a new medicine", Office of Health Economics.

Michel, J and B. Bettels (2001), "Patent citation analysis: A closer look at the basic input data from patent search reports", *Scientometrics*, **51(1)**, 185-201.

Min, Y. and A. Agresti (2005), "Random Effect Models for Repeated Measures of Zero- Inflated Count Data", *Statistical Modeling*, **5(1)**, 1-19.

Mohnen, P., J. Mairesse and M. Dagenais (2006), "Innovativity: A Comparison across Seven European Countries", *Economics of Innovation and New Technology*, **15(4/5)**, 391-413.

Montalvo, J. G (1997), "GMM Estimation of Count-Panel-Data Models with Fixed Effects and Predetermined Instruments", *Journal of Business and Economic Statistics*, **15(1)**, 82-89.

Mossialos, E. and A. Oliver (2005), "An overview of pharmaceutical policy in four countries: France, Germany, the Netherlands and the United Kingdom", *The International Journal of Health Planning and Management,* **20(4)**, 291-306.

Mullahy, J. (1986), "Specification and testing of some modified count data models", *Journal of Econometrics*, **33(3)**, 341-365.

Murphy, K. M and R. H. Topel (2006), "The value of health and longevity", *Journal of Political Economy*, **114(5)**, 871-904.

Nefarma (2003), Annual report 2002, Den Haag.

Nefarma (2011), Pharma Facts and Figures 2011, Den Haag.

Nevo, A. (2001), "Measuring market power in the ready-to-eat cereal industry", *Econometrica*, **69(2)**, 307-342.

Netherlands Pharmaceutical Country Profile (2011), Ministry of Health, Welfare and Sports; and World Health Organization.

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

Nesta, L. and P. Saviotti (2005), "Coherence of the Knowledge Base and the Firm's Innovative Performance: Evidence from the U.S. Pharmaceutical Industry", *Journal of Industrial Economics*, **53(1)**, 123-142.

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

Nickell, S. J., B. S. Wadhwani and M. Wall (1992), "Productivity Growth in U.K. Companies, 1975-1986", *European Economic Review*, **36(5)**, 1055-1085.

Nilsen, O. A. and F. Schiantarelli (2003), "Zeroes and Lumps in Investment: Empirical Evidence on Irreversibilities and Nonconvexities", *The Review of Economics and Statistics*, **85(4)**, 1021-1037.

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

O'Mahony, M. and M. Vecchi (2009), "R&D, Knowledge Spillovers and Company Productivity Performance", *Research Policy*, **38(1)**, 35-44.

Olley, S. and A. Pakes (1996), "Market share, market value and innovation in a panel of British manufacturing firms", *Econometrica*, **64(6),** 1263-1297.

Ophem, V., H. E. Brouwer, A. Kleinknecht and P. Mohnen (2002), "The Mutual Relation between Patents and R&D", in: Kleinknecht, A. and P. Mohnen eds., Innovation and Firm Performance, Hampshire and New York: Palgrave, 56-70.

Pakes, A. and M. Simpson (1991), "The Analysis of Patent Renewal Data", Brookings Papers on Economic Activity, Microeconomic Annual, 331-401.

Pakes, A. (1985), "On Patents, R&D, and the Stock Market Rate of Return", *Journal of Political Economy*, **93(2)**, 390-409.

Pakes, A. and Z. Griliches (1980), "Patents and R&D at the firm level: A first report", *Economics Letters*, **5(4)** , 377-381.

Pammolli, F., L. Magazzini and M. Riccaboni (2011), "The productivity crisis in pharmaceutical R&D", *Nature Reviews: Drug Discovery*, **10**, 428-438.

Pammolli, F and M. Riccaboni (2007), "Innovation and industrial leadership: lessons from pharmaceuticals", Center for Transatlantic Relations, The John Hopkins University, Washington D.C.

Parisi, M. L., F. Schiantarelli and A. Sembenelli (2006), "Productivity, Innovation and R&D: Micro Evidence for Italy", *European Economic Review*, **50(8)**, 2037-2061.

Pavitt, K., M. Robson and J. Townsend (1987), "The Size Distribution of Innovating Firms in the UK: 1945- 1983", *Journal of Industrial Economics*, **35(3)**, 297-316.

Peters, B. (2009), "Persistence of innovation: stylized facts and panel data evidence", *Journal of Technology Transfer*, **34(2)**, 226-243.

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

Peters, B. (2006), "Productivity effects of innovation activities", Chapter 4 of Peters, B., Innovation and Firm Performance: An Empirical Investigation for German Firms, University of Würzburg: PhD thesis.

Petrin, A., B. P. Poi and J. Levinsohn (2004), "Production function estimation in Stata using inputs to control for unobservables", *The Stata Journal*, **4(2)**, 113-123.

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

Philipson, T. J. and S. Mechoulan (2003), "Intellectual Property and External Consumption Effects: Generalizations from Pharmaceutical Markets", NBER Working Paper Series 9598.

PhRMA (2010), "Pharmaceutical Industry Profile 2010", Pharmaceutical Research and Manufacturers of America, Washington DC.

Pitt, M. M. (2011), "Overidentification Tests and Causality: A Second Response to Roodman and Morduch", Draft, Brown University.

Polder, M., G. V. Leeuwen, P. Mohnen and W. Raymond (2009), "Productivity Effects of Innovation Modes", Statistics Netherlands, Discussion Paper 09033.

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.

Quillen, C. D., O. H. Webster, and R. Eichmann (2002), "Continuing Patent Applications and Performance of the U.S. Patent and Trademark Office-Extended', *The Federal Circuit Bar Journal*, **12(1)**, 35-55.

Raymond, W., P. Mohnen, F. Palm and S. Schim der Loeff (2010), "Persistence of Innovation in Dutch Manufacturing: Is it Spurious?", *The Review of Economics and Statistics*, **92(3)**, 495-504.

Raymond, W., J. Mairesse, P. Mohnen and F. Palm (2013), "Dynamic models of R&D, Innovation and Productivity: Panel Data Evidence for Dutch and French Manufacturing", CIRANO Working Paper.

Raymond, W., 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.

Reichert, J. M. and C. Milne (2002), "Public and private sector contributions to the discovery and development of impact drugs", Boston, MA: Tufts Center for the Study of Drug Development.

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.

Roodman, D. (2009), "How to do xtabond2: An introduction to difference and system GMM in Stata", *The Stata Journal*, **9(1)**, 86-136.

Roper, S. and N. H. Dundas (2008), "Innovation Persistence: Survey and Case-Study Evidence", *Research Policy*, **37(1)**, 149-162.

Roy, L. and A. Wickelgren (2001), "Competition Policy Issues for Regulators: A U.S. Perspective on Pharmaceutical Industry Cases Before the Federal Trade Commission" in: Kettler, H. ed., Consolidation and Competition in the Pharmaceutical Industry, London: Office of Health Economics, 106-117.

Sakakibara, M. and L. Branstetter (2001), "Do stronger Patents Induce More Innovation? Evidence from the 1988 Japanese Patent Reforms", *The Rand Journal of Economic*, **32(1)**, 77-100.

Sakakibara, M. and M. E. Porter (2001), "Competing at Home to Win Abroad: Evidence from Japanese Industry", *The Review of Economics and Statistics*, **83(2)**, 310-322.

Santos Silva, J. M. C and S. Tenreyro (2006), "The Log of Gravity", *The Review of Economics and Statistics*, **88(4)**, 641-658.

Santos Silva, J. M. C and S. Tenreyro (2011), "Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator", *Economic letters*, **112(2)**, 220-222.

Saranga, H. and R. Banker (2010), "Productivity and technical changes in the Indian pharmaceutical industry", *Journal of Operational Research Society*, **61(12)**, 1777-1788.

Schankerman, M and A. Pakes (1986), "Estimates of the Value of Patent Rights in European Countries during the Post-1950 Period", *Economic Journal*, **96(384)**, 1052-1076.

Schankerman, M. (1981), "The Effect of Double-Counting and Expensing on the Measured Returns to R&D", *Review of Economics and Statistics*, **63(3)**, 454-458.

Scherer, F. M. (2007), "Pharmaceutical Innovation", Faculty Research Working Paper Series, Harvard University.

Scherer, F. M. (1980), "Industrial Market Structure and Economic Performance", Rand McNally, Chicago.

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

Schumpeter, J. (1942), "Capitalism, Socialism and Democracy", New York: Harper and Roe Publishers.

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

Shapiro, C. (2002), "Competition Policy and Innovation", OECD STI Working Papers 2002/11.

Sharma, C. (2011), "R&D and productivity in the Indian pharmaceutical firms", MPRA Working Paper No. 31681.

Siebeck, W. E. (1990), "Strengthening Protection of Intellectual Property in Developing Countries: A Survey of the Literature", World Bank Discussion Paper, Washington DC.

Smolny, W. (2000), "Endogenous Innovations and Knowledge Spillovers: A Theoretical and Empirical Analysis", ZEW Economic Studies, **12**, Heidelberg.

Solow, R. M. (1957), "Technical progress and the aggregate production function", *The Review of Economics and Statistics,* **39(3)**, 312-320.

Srivastava, V. S. (2001), "The impact of India's Economic Reforms on Industrial Productivity, Efficiency and Competitiveness- A panel study of Indian companies 1980-1997", NCAER study.

Stiglitz, J. (1989), "Imperfect information in the product market", in: R. Schmalensee, R. Willig (Eds.), Handbook of Industrial Organization 1, Elsevier Science Publishers, Amsterdam, 769-847.

Sutton, J. (1991), "Sunk Costs and Market Structure", Cambridge, MA.

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

Tingvall, P. G. and A. Poldahl (2006), "Is there really an inverted U-shaped relation between competition and R&D?", *Economics of Innovation and New Technology*, **15(2)**, 101-118.

Tirole, J. (1988), "The theory of industrial organization", MIT Press.

Trajtenberg, M., R. Henderson and A. Jaffe (1997), "University versus Corporate Patents: A Window on the Basicness of Invention", *Economics of Innovation and New Technology*, **5(1)**, 19-50.

Trajtenberg, M. (1990), "A Penny for Your Quotes: Patent Citations and the value of Innovations", *The Rand Journal of Economics*, **21(1)**, 172-187.

Tran, N., N. Wilson and D. Hite (2013), "Choosing the Best Model in the Presence of Zero Trade: A Fish Product Analysis", in: Frontiers of Economics and Globalisation-Non-Tariff Measures with Market Imperfections, Trade and Welfare Implications, Emerald Group Publishing Limited.

Van de Ven, W. P. M. M. and F. T. Schut (2008), "Universal mandatory health insurance in the Netherlands: A model for the United States?", *Health Affairs*, **27(3)**, 771-781.

Vancauteren, M., P. Kort, B. Melenberg and J. Plasmans, "Innovation and Productivity of Dutch Firms: A Panel Data Analysis", (forthcoming).

Vancauteren, M. (2012), "EU harmonization of regulations and mark-ups in the Dutch food industry", *European Review of Agricultural Economics*, **40(1)**, 163-189.

Vancauteren, M. and B. H. de Frahan (2011), "Trade Policy, Competition and Productivity: The Impact of EU Harmonization in the Dutch Food Processing Industry", *De Economist*, **159(4)**, 483-509.

Vuong, Q. H. (1989), "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses", *Econometrica*, **57(2)**, 307-333.

Winkelmann, R. and K. F. Zimmermann (1995), "Recent Developments in Count Data Modelling: Theory and Application", *Journal of Economic Survey*, **9(1)**, 1-24.

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(1)**, 39-54.

universiteit hasselt
KNOWLEDGE IN ACTION

www.uhasselt.be
Hasselt University
Martelarenlaan 42 | BE-3500 Hasselt
T + 32(0)11 26 81 11