ESSAYS ON THE DRUG DISCOVERY INNOVATION SYSTEM

by

Alfred Sarkissian
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Public Policy

Committee:

_____	David M. Hart, Chair
_____	Siona Listokin
_____	Naoru Koizumi
	Ruben Jacobo-Rubio, External Reader
_____	Sita N. Slavov, Program Director
_____	Mark J. Rozell, Dean

Date: _____ Fall Semester 2017
George Mason University
Fairfax, VA

Essays on the Drug Discovery Innovation System

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Alfred Sarkissian
Master of Arts
Allameh Tabatabai University, 2002
Bachelor of Arts
Azad University, 1997

Director: David M. Hart, Professor
Department of Public Policy; Schar School of Policy and Government

Fall Semester 2017
George Mason University
Fairfax, VA

ACKNOWLEDGEMENTS

I need to express my deepest gratitude to my committee chair Prof. David M. Hart for his advice, mentorship, and patience in the course of this dissertation and the program. I also owe a sincere appreciation to other committee members Prof. Naoru Koizumi and Prof. Siona Listokin. Prof. Mark Addleson was forthcoming with his time during the field stage of the doctoral program. Special thanks to the administrative staff of the Schar School of Policy and Government.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS AND/OR SYMBOLS

European Patent Office ..... EPO
High-Throughput Screening ..... HTC
International Patent Classification ..... IPC
New Biotechnology Firms ..... NBFs
New Molecular/Chemical Entity ..... NME/NCE
Non-patent References/Literature ..... NPR/NPL
Orphan Drug Act ..... ODA
Pharmaceutical Research and Manufacturers of America ..... PhRMA
Prescription Drug User Fee Act ..... PDUFA
Small Business Innovation Research ..... SBIR
Triadic Patent Families ..... TPS
United States Patent and Trademark Office ..... USPTO

ABSTRACT

ESSAYS ON THE DRUG DISCOVERY INNOVATION SYSTEM

Alfred Sarkissian, Ph.D.

George Mason University, 2017

Dissertation Director: Dr. David M. Hart

Despite the increase in scientific, monetary, and human resources devoted to drug discovery R&D, the annual number of FDA-approved drugs has been almost unchanged for decades. This study seeks clues to the R&D productivity paradox by exploring the drivers and barriers to innovation in the pharmaceutical industry. The main focus is on the contribution of science and knowledge form external sources to innovation.

The first essay uses patent data and provides evidence in support of the "foundational view" of creativity that a deep grasp of domain knowledge is most important in developing new ideas.

Patent-based models of firm inventive output indicate a curvilinear (first negative then positive) pattern for the contribution of codified scientific knowledge to inventive output. This finding implies that firms need to assimilate and utilize scientific knowledge across multiple inventions. The use of diverse knowledge components in innovation is positively related to quality-weighted inventive output. However, a firm's reliance on knowledge

components different form own inventions, negatively influences both inventive output and quality-weighted inventive output.

Essay 2 examines breakthrough innovations (i.e. Orange Book patents) for clues about what makes inventions underpinning a drug product different from other inventions. The results indicate that inventions with fewer applications (i.e. less general or with citations across different technology classes) have a higher probability of entering the Orange Book. Broader legal scope also boosts the probability of an invention being an Orange Book patent. This finding might be caused by inventors carving out larger legal protection for more valuable inventions. In terms of knowledge recombination, higher technological knowledge diversity reduces the probability of an invention being an Orange Book patent, while technological knowledge distance (i.e. not coming from the patent's technology class) increases the probability of being an Orange Book patent. These results are in line with the "tension view" of creativity, which emphasizes the need for distant or diverse knowledge for generating new ideas.

The third essay draws on a survey of experts that explores the drivers and barriers to innovation from their perspective. The respondents possess nuanced knowledge of broad R&D spending and drug approval trends. Some barriers to innovation that they perceive are accentuated by the "molecular reductionist" drug discovery paradigm. Moreover, the essay provides evidence of several systemic failures. Lack of change in the fundamental rules of the game has created a "lock-in/path dependency failure" in which the innovation system has failed to adapt expeditiously. Deficiencies in firm capability development have lead to "transition failures". Hard (i.e. regulatory) and soft (i.e. cultural) institutional

failures, along with “regulatory capture” are also evident in the responses and other data. In line with the “foundational” view of creativity, respondents ranked “depth of knowledge” higher than “diversity of knowledge” for innovation. The latter result is in line with the insights form essay 1.

These systemic failures provide a starting point for formulating policy interventions to improve the innovative output of the drug discovery innovation system.

CHAPTER ONE: INTRODUCTION - THE DRUG DISCOVERY PARADOX

Abstract

Despite the increase in scientific, monetary, and human resources devoted to drug discovery the annual number of FDA-approved drugs has been almost unchanged for decades. This observation, coupled with the importance of drug discovery for emerging health challenges (e.g. superbugs) indicates the importance of developing better knowledge of the drug discovery innovation issues. This chapter outlines the R&D efficiency paradox of the pharmaceutical sector and briefly discussed some related topics.

Keywords: Eroom's Law, Pharmaceutical Innovation; Drug Discovery

1.1 Eroom's Law

Despite tremendous improvement in the science and technology underpinning drug discovery, improvements in R&D management, and copious R&D investments, there has been little change in the critical output of the industry which is the number of new drugs approved by the FDA (Austin, 2006).

The paradoxical phenomenon of declining R&D efficiency (i.e. the number of new FDA-approved drugs per billion US dollars of R&D spending[1]) has been characterized in different ways. Scannell et al. (2012) coin the term "*Eroom's Law*"; i.e. "*Moore's Law*" backwards; to refer to the anomaly of halving the number of new FDA-

[1] Scannell et al. (2012) use PhRMA Annual Survey of 2011 for R&D cost data

approved drugs[2] per billion US dollars of R&D spending about every 9 years since 1950 (figure 1). The trend has remained the same since the 1950s. In a similar line, some commentators have divided the industry into an "Era of Abundance" (pre-2005) and an "Era of Scarcity" (post-2005). Measuring R&D productivity in terms of fifth year product sales per billion dollars of R&D expenditure, there has been a 70% decline in R&D productivity (Reuters, 2015).

As figure 1 depicts, the decline rate is almost the same across different 10-year segments (panel b). Moreover, the trend seems to be robust to various assumptions on average delay between R&D spending and drug approval (panel c) (Scannell et al., 2012). A number of causes underlying the phenomenon are briefly discussed here.

[2] New molecular entities and new biologics

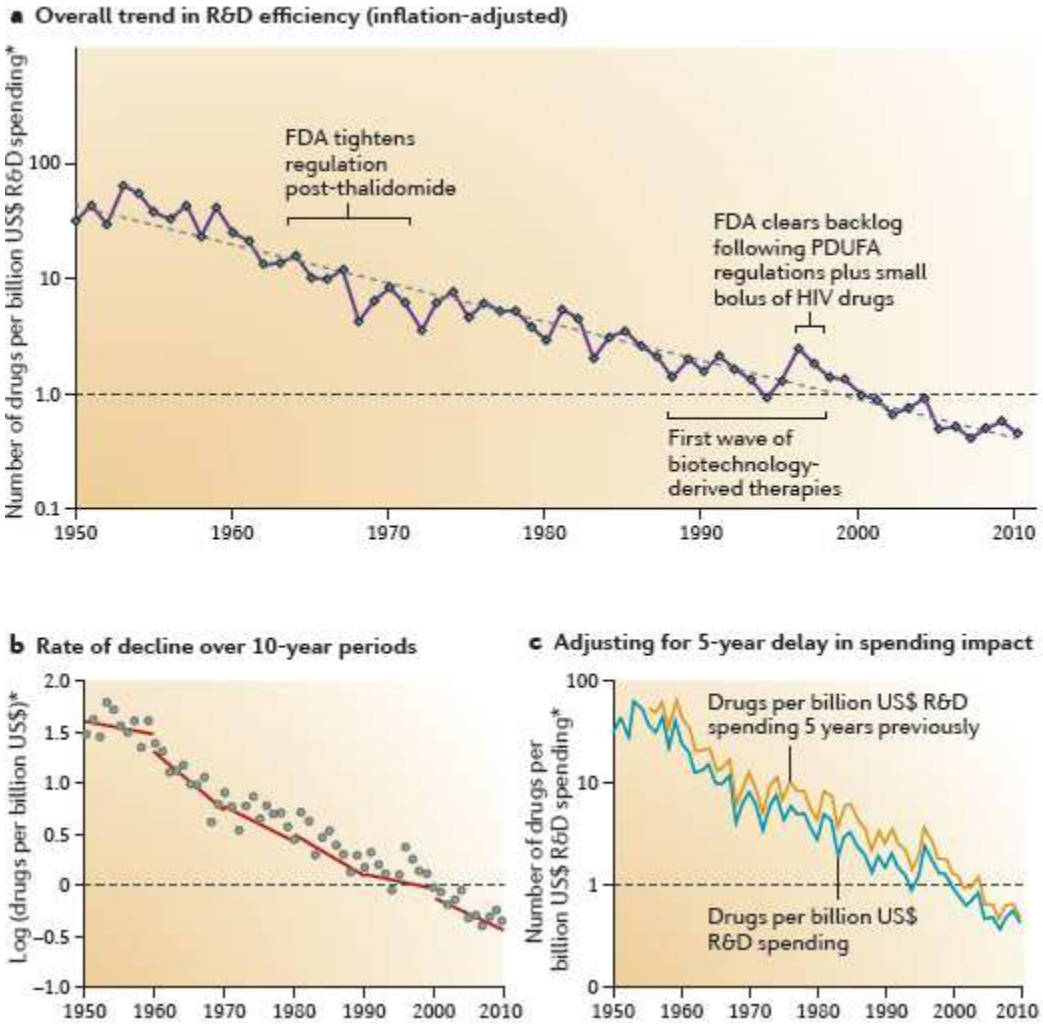

Figure 1 The Eroom's Law in Pharmaceutical R&D

Source: Scannell et al. (2012)

* Prescription Drug User Fee Act (PDUFA)[3]

a) "Better than the Beatles" Problem

The "*better than the Beatles*" problem refers to the situation where a large stock of approved drugs exists and new drugs have only a modest incremental benefit over what is already available (Scannell et al., 2012). Many diseases have been satisfactorily

[3] The Prescription Drug User Fee Act (PDUFA) was enacted in 1992 to authorize the FDA to collect fees from companies that produce certain human drug and biological products. User fees have helped expedite the drug approval process.

tackled, limiting the innovation space for big medical breakthroughs (Ding et al., 2014a). Examining R&D projects of over 28,000 compounds since 1990, Pammolli et al. (2011) report a shift form therapeutic areas with better chance of drug approval (e.g. genitourinary drugs) to areas with lower historical approval rates (e.g. immunomodulatory agents). The latter areas correspond with "*unmet therapeutic needs and unexploited biological mechanisms*" with higher risk of failure (Pammolli et al., 2011). In terms of the impact of failed innovation on innovative output, Khanna et al. (2016) examine voluntary patent expirations (i.e., patents expiring due to non-payment of renewal fees) in 97 pharmaceutical firms in the 1980-2002 period and report that the number, importance, and timing of small failures are associated with a decrease in patent count but an increase in forward citations to patents.

Based on the aforementioned arguments, less innovation opportunities coupled with harder diseases to tackle may have lead to reduced quantitative output of drugs.

b) "Cautious Regulator" Problem

The "*cautious regulator*" problem emanates from the reduced risk tolerance of regulatory agencies and the concomitant increase in R&D compliance costs. R&D efficacy shows a decrease in the early 1960s after the 1962 Kefauver Harris amendment to the Federal Food, Drug, and Cosmetic Act of the 1962. The Act required demonstration of efficacy for the first time and raised the safety hurdle (Scannell et al., 2012). Jensen (1987) examines a panel data of 28 firms in the time period of 1969 through 1979 and observes that an increase in regulatory stringency decreases the expected number of new drug discoveries.

c) "Throw Money at it" Tendency

High returns, uncertain processes, and intense competition have given rise to the "*throw money at it*" tendency (Scannell et al., 2012). Some empirical research dealing with R&D spending is relevant here. At the country level, Lichtenberg and Virabhak (2002) use "*Triadic Patent Families*", i.e. sets of patents covering a single invention filed together in Europe, Japan and the US, to explore various dimensions of technical change in healthcare technologies. They observe that R&D and accumulated knowledge (i.e. patent stock) are important determinants of patenting with elasticities of 0.43 and 0.79, respectively. Jensen (1987) reports a positive correlation between a firm's R&D intensity and its probability of discovering a new drug while firm size, did not significantly affect the marginal productivity of research expenditures. Finally, Penner-Hahn and Shaver (2005) scrutinize the international R&D expansion activities, research capabilities, and patent output of 65 Japanese pharmaceutical firms in the period 1980-1991. They observe that firms benefit from international R&D only when they have existing research capabilities in the underlying technologies.

To sum up, these studies indicate that while more R&D spending may lead to higher output, there are important moderating effects such as existing research capabilities.

d) "Basic Research Brute Force" Bias

The "*basic research–brute force*" bias refers to the overconfidence in basic research and screening methods to yield safe and effective molecules for clinical trials (Scannell et al., 2012). Current drug discovery is impossible without utilizing

sophisticated modeling and computation. Computational approaches are thought to have the potential to reduce the cost of drug development dramatically by boosting the set of feasible targets (Yao et al., 2009).

This is partly because of the popularity of "molecular reductionism" as well as the commercial managers' quest for efficiency. In molecular reductionism genetics and molecular biology are the preferred ways of understanding biological systems. Automation had worked in other industries hence it was tempting to move away from "unpredictable" animal-based screening to a more efficient method (adapted from Scannel et al., 2012).

e) "Low-hanging Fruit" Problem

The "*low-hanging fruit*" problem reflects the issue of continued exploitation of drug targets that are more technically malleable. While the "*better than the Beatles*" problem means the fruit that has already been picked reduces the value of the remaining fruit, the "*low-hanging fruit*" problem contends that easy-to-pick fruit is gone. It is less important than the "*better than the Beatles*" problem because decades-worth of new drug targets is thought to be available for exploitation. Moreover, therapeutic benefit may come from interaction of multiple proteins rather than a single target (Scannell et al., 2012).

f) Other Trends Impacting Innovative Output

Potential market has been shown to impact innovative output. Acemoglu and Linn (2004) scrutinize the impact of (potential) market size on the entry of new drugs and pharmaceutical innovation. They construct a population-based potential market size

based on U.S. demographic trends and report a large effect of potential market size on the entry of nongeneric drugs and new molecular entities (NMEs)[4]. They observe that all nongenerics respond to current market size, while current and five-year leads of market size impart the strongest influence on new molecular entities and generics. They contend that pharmaceutical research responds to anticipated market size changes with a lead of 10–20 years. This strand of research may overlap with "*better than the Beatles*" problem in the sense that drugs exist for many ailments and by extension patient populations.

The Innovation strategy of the pharmaceutical sector is changing with knowledge acquisition superseding knowledge creation (Cardinal and Hatfield, 2000). There seems to be a shift away from the vertically integrated model with a strong blockbuster orientation towards more incremental and follow-on innovations, greater specialization, and focused R&D. This is meant to utilize established competencies (e.g. in marketing), boost efficiency, and create synergies (Petrova, 2014). Duplicative and nonproductive investments have emerged as the upshot (Bennani, 2011). The "*patent cliff*" may also play a role in this regard (Petrova, 2014). The "*patent cliff"* refers to the expiration of patents on highly profitable drugs of several major pharmaceutical companies. For instance, in 2011, Pfizer lost patent protection on Lipitor, its most profitable product accounting for 27 percent of its total revenues in 2006. The main challenge ensuing from the patent cliff is replacing these expiring drugs. According to one estimate, companies'

[4] The term NME or NCE (new chemical entities) refers to active ingredients being marketed for the first time in the U.S. (Drug Approvals and Databases > Drugs@FDA Glossary of Terms." Accessed April 9, 2016. http://www.fda.gov/Drugs/InformationOnDrugs/ucm079436.htm.)

new product revenues will only replace 26%of the expiring-patent revenues (Lo and Naraharisetti, 2013).

According to Scannell et al. (2012) the primary "*causes*" have led to some secondary symptoms: the narrow clinical search problem (i.e. looking for a specific outcome rather than looking for broader positive serendipitous outcomes form drug trials); the big clinical trial problem (i.e. in terms of number of patients involved); the multiple clinical trial problem (i.e. number of trials per drug has increased), and the long cycle time problem.

1.2 Scientific and Technological Advances

There have been major advancements in the plethora of scientific and technological inputs into the drug discovery process in the past 60 years. A review of these advances gives more perspective to the paradox of Eroom's Law. New inventions such as biotechnology, computational drug design, and scientific advances such as better knowledge of disease mechanisms have opened new vistas for drug discovery. Combinatorial chemistry has boosted the number of potential synthesizable molecules per chemist per year by about 800 fold (Scannell et al., 2012). Combinatorial chemistry can be characterized as "*the industrialization of chemistry*". While the underlying science of chemistry is the same, the means of carrying it out has changed. Instrumentation, robotics and extensive use of computers are incorporated into the process to analyze large quantities of data (Geysen et al., 2003).

Biotechnology is a science-based activity, drawing crucially on public research carried out in universities and government laboratories. It is a set of technologies applied

in various industries such as pharmaceutical manufacturing or agriculture. Some pharmaceutical firms draw upon a number of biotechnologies (Niosi, 2011).

DNA sequencing speed has increased a billion times since the first genome sequencing in the 1970s (Scannell et al., 2012). As a case in point, there is a world-wide endeavor underway to catalogue mutations in several cancer types that most probably will be conducive to the development of new diagnostic, prognostic and therapeutic targets. The possibility of screening numerous gene targets quickly and at minimal costs is a major driver of this technology because cancer treatment can potentially be transformed by the emergence of small molecule inhibitors and antibodies against genes (Meldrum et al., 2011).

Finally, current drug discovery practices draws on sophisticated modeling and computation techniques such as text mining for new drug leads, modeling molecular pathways and predicting drug combination efficacy (Yao et al., 2009).

1.3 R&D Management

The R&D process management knowledge has gained substantially in terms of understanding sources of project overruns, financial returns, portfolio management, and cost controls through outsourcing (Scannell et al., 2012). The current drug discovery process, a heritage of the 1960s, is cumbersome, inefficient and pricey; hence, in need of major overhaul. Major drug developers are in the process of bringing significant change in the prevailing R&D models (Kaitin, 2010).

Pharmaceutical mergers, acquisitions and strategic alliances have been the subject of some studies. According to Mittra (2007) mergers, acquisitions and strategic alliances

are a set of related strategies that provide alternatives for innovation management and productivity improvement. Firm-specific factors and environmental differences lead to differing desired mix of internal R&D and external knowledge sourcing. Large companies have the potential to shift from the "*blockbuster small-molecule R&D model*" to a biologics-based paradigm; however, the process is gradual and evolutionary, proceeding at different speeds in different companies (Mittra, 2007).

The small-molecule R&D model refers to the R&D model of developing conventional drugs that are small, well-defined and stable chemical structures, amenable to characterization by analytical methods (Declerck, 2012). Biologics on the other hand, have high molecular weight and are the result of processes in living cell cultures (see table 1 for a comparison of these two types of drugs).

As an example, So et al. (2011) explore how an amalgam of scientific and economic issues has thwarted new antibacterials development in the past few decades despite a growing global problem of resistance to existing antibiotics. Numerous bottlenecks along the pharmaceuticals value chain can be a potential intervention point. They suggest multisectoral collaboration as a solution in drug lead[5] identification and optimization. Moreover, neglected disease areas can get a boost from "*product development partnerships and South–South innovation platforms*". Decoupling product sales from the firms' return on investment is touted as a solution towards attaining the double goals of innovation and access. This is because firms with an extensive portfolio

[5] Chemical compound with likely therapeutic properties

prefer to focus on other therapeutic areas at the expense of new antibiotics that have lower expected returns on investment.

This short review points to the advances of R&D management knowledge, further making the R&D productivity decline a paradoxical phenomenon.

Table 1 Small Molecule Drugs versus Biologicals

Characteristic	Small molecule drugs	Biological drugs
Size	- Small (single molecule) - Low molecular weight	- Large (mixture of related molecules) - High molecular weight
Structure	Simple, well defined, independent of manufacturing process	Complex (heterogeneous), defined by the exact manufacturing process
Modification	Well defined	Many options
Manufacturing	- Produced by chemical synthesis - Predictable chemical process - Identical copy can be made	- Produced in living cell culture - Difficult to control from starting material to final API - Impossible to ensure identical copy
Characterization	Easy to characterize completely	Cannot be characterized completely the molecular composition and heterogenicity
Stability	Stable	Unstable, sensitive to external conditions
Immunogenicity	Mostly non-immunogenic	Immunogenic

Source: "Home - GaBI Online - Generics and Biosimilars Initiative." Accessed April 5, 2017. http://www.gabionline.net/Biosimilars/Research/Small-molecule-versus-biological-drugs.

1.4 R&D Spending

The pharmaceutical industry's R&D spending relative to its sales revenues is larger than other U.S. industries. The real drug R&D has burgeoned between threefold and sixfold in the course of the past quarter century (Austin, 2006). Likewise, according to the Pharmaceutical Research and Manufacturers of America (PhRMA, 2016), biopharmaceutical industry's R&D per employee investments is 12 times more than other manufacturing industries. Moreover, the R&D investment growth rate (25%) has also

been the highest across all industries between 2000 and 2012. The biopharmaceutical sector accounts for about 17% of all R&D spending in the U.S. business sector; representing the largest share of business R&D spending (PhRMA, 2016).

Figure 2 depicts the increase in R&D spending in the past two decades and figure 3 compares different industries based on their R&D spending as a percentage of sales. Figure 4 depicts the geographical breakdown of R&D expenditure. As evident, especially beginning 2000, the U.S. has consistently out-spent other important world regions in pharmaceutical R&D.

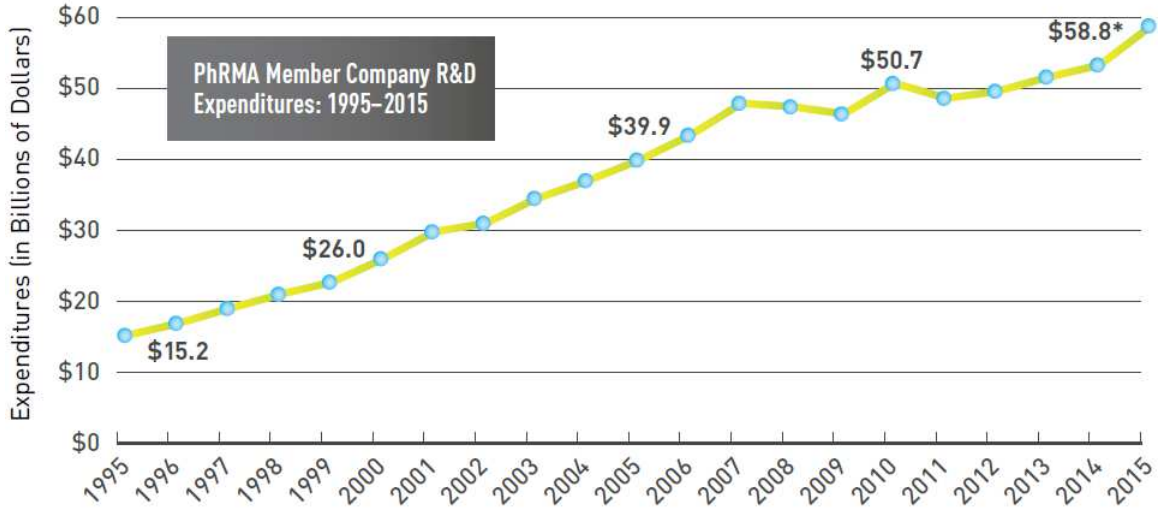

Figure 2 R&D Investment by PhRMA Member Companies
Source: PhRMA (2016)

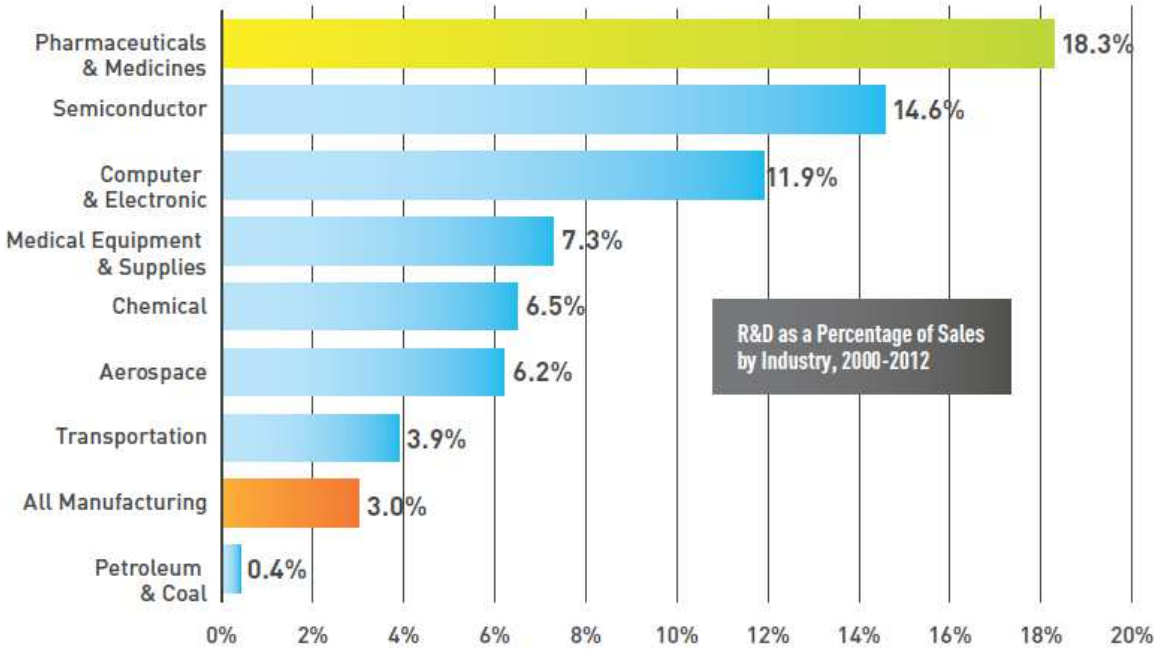

Figure 3 R&D as a Percentage of Sales
Source: PhRMA (2016)

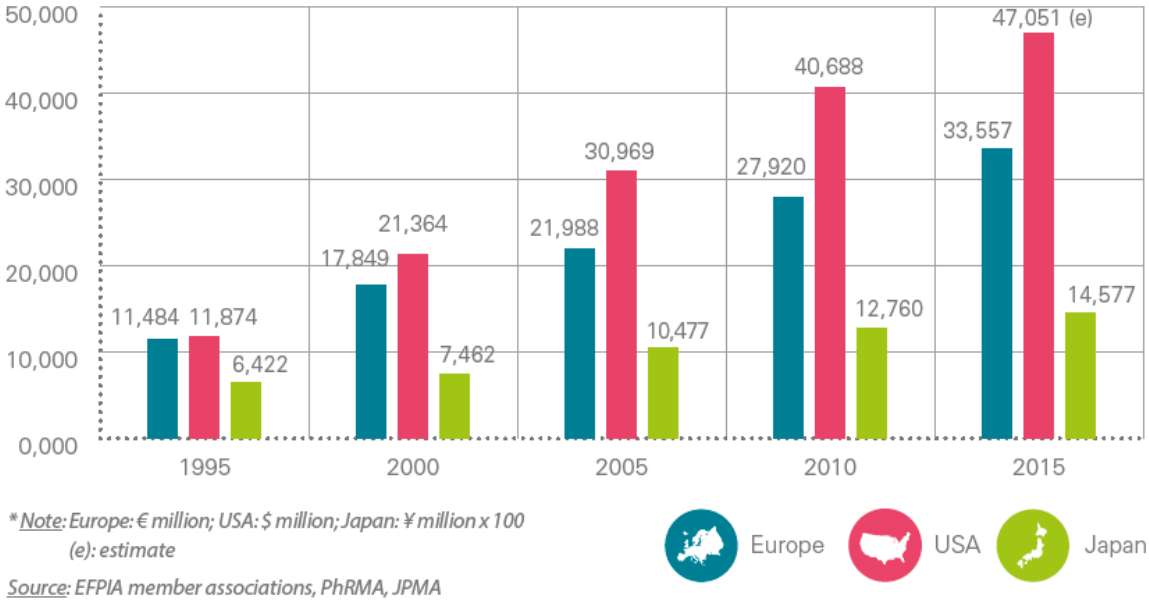

Figure 4 Geographic Breakdown of R&D Expenditure
Source: EFPIA (2017)

1.5 Conclusion

The pharmaceutical industry account presented in this chapter contradicts the traditional narrative of R&D, invention and innovation where more resources is expected to lead to more innovative output. Moreover, the crucial welfare enhancing role of the innovative output of the industry, calls for a close examination of innovation in the pharmaceutical industry for potential clues to the problems.

The remainder of this dissertation will try to explore some factors impacting innovative output of the pharmaceutical sector based on patenting data. Moreover, an

attempt will be made to ascertain the differences between patented inventions underpinning a drug product and other patents. Finally, drawing on the innovation systems literature and expert opinion, the broad systemic issues driving or hindering innovation will be explored.

CHAPTER TWO: EXPLORING THE DRIVERS OF INNOVATION IN THE PHARMACEUTICAL SECTOR - THE IMPACT OF KNOWLEDGE RECOMBINATION

Abstract

This study seeks to ascertain the drivers of inventive output (i.e. patenting by firms) in the pharmaceutical sector with special attention to the way different types of knowledge is combined. The results indicate a robust curvilinear; first negative and then positive; relationship between reliance on codified scientific knowledge (enshrined in patent references) and inventive output. Moreover, knowledge diversity positively impacts quality-weighted inventive output while knowledge heterogeneity (i.e. based on distance or differentness from own inventions) negatively influences both inventive output and quality-weighted inventive output. The Latter results are in line with the foundational view of creativity that contends a deep grasp of domain knowledge is important in developing breakthrough innovations. Nuanced subsample analysis revealed that for firms primarily active in the pharmaceutical sector higher knowledge diversity in inventive output does not improve performance. Moreover, pre-1995 reliance on codified science was associated with higher inventive output.

Keywords: Pharmaceutical Innovation; Patents; Triadic Patent Families; Science Intensity; Knowledge Diversity; Knowledge Heterogeneity (Distance)

2.1 Introduction

Drugs, as the output of the pharmaceutical industry, have a direct bearing on the people's healthcare. Moreover, healthcare and availability of therapeutics are crucial issues in the definition of welfare and democracy in the new century. However, during the past decades the pharmaceutical industry has experienced profound scientific, technological and institutional change affecting the whole spectrum of the value chain (McKelvey and Orsenigo, 2001). Nonetheless, a paradoxical parallel phenomenon is the declining R&D efficiency: the number of new FDA-approved drugs per billion inflation-adjusted U.S. dollars of R&D spending has halved about every 9 years since 1950 (Scannell et al., 2012).

Patents have long been recognized as a rich source of data pertaining to innovation and technological change. They contain detailed technical information and provide long historical records. Moreover, unlike other sources of economic data, patent information is disclosed voluntarily (Hall et al., 2001). The pharmaceutical industry is probably one of the few industries where product innovation can be adequately protected against imitation by patents. This is because small variation of a molecular structure can dramatically change its pharmacological properties (McKelvey et al., 2004).

There has been rapid growth in the reliance of patented technology on U.S. scientific papers (Narin et al. 1997). Furthermore, the patent-to-paper citation trend is more pronounced in the bioscience-related inventions (Branstetter and Ogura, 2005). Finally, with the increase in the science and technology pool that underpin drug discovery innovation (e.g. better knowledge of disease mechanisms and faster gene sequencing),

examining how science and knowledge recombination impacts innovative output is in order.

This essay draws on patent data to examine how codified scientific knowledge (i.e. enshrined in non-patent references within patents) and knowledge recombination (i.e. diversity and heterogeneity of knowledge) impact the innovative output of a panel of firms active in pharmaceutical patenting.

2.2 Literature Review

This section briefly reviews a number of relevant empirical studies on innovation in the pharmaceutical sector. First a few more general notes on the literature are discussed then the roles of science and "knowledge recombination" in innovation are introduced.

Regarding outcome variables, the literature can be divided into three categories. Some use patent counts as a proxy of innovation (e.g. Lichtenberg and Virabhak, 2002; Penner-Hahn and Shaver, 2005). A second group utilizes actual FDA drug substance approval (NME) as their dependent variable (e.g. Jensen, 1987; Graves and Langowitz, 1993; Toole 2012). Finally, a few studies have used both patent counts and approval rates (e.g. Cardinal and Hatfield, 2000).

In terms of contents, diverse issues have been examined. For instance, Graves and Langowitz (1993) examine the innovative output (NCEs) of 16 pharmaceutical firms over 19 years (1969-1987) and observe that the pharmaceutical industry suffers decreasing returns to scale in R&D. They notice a strong correlation between R&D budgets and firm

size and mention that the wave of mergers in the industry may yield less innovative productivity.

Cardinal and Hatfield (2000) use three measures for innovative productivity: patents, drug enhancement approvals (incremental innovations), and new drug approvals (i.e. NCEs or radical innovations). They observe that having an R&D laboratory boosts the productivity of drug enhancements, but has a negative impact on new drug productivity; however, having an R&D laboratory in the proximity of corporate headquarters enhances new drug productivity. They contend that proximity might be important for the integration of R&D with strategic product innovation goals.

Finally, Nesta and Saviotti (2005) study the relationship between the firm's knowledge base coherence and innovative performance. Coherence refers to complementarity between the firm's scientific and technological competencies. They find a strong link between these two properties and the firms' innovative performance measured by the annual citations weighted patents.

2.2.1 Role of Science in Pharmaceutical Innovation

The front-page of a U.S. patent contains two types of references, i.e. patent and non-patent references (NPR/NPLs[6]), which are meant to help the examiner evaluate the novelty, non-obviousness and the applicability of inventions (i.e. patentability requirements). Moreover, these references can be used in evaluating the validity of the patent claims (Van Looy et al., 2007). The distinction between patent and non-patent references has been used to study the interaction of science and technology. Science

[6] Non-patent literature (NPL or NPR) is mostly scientific references other than patents. The US patent document has a "references cited" section that divides the references into "US patent documents", "foreign patent documents" and "other publications".

intensity of a patent can be proxied as the number of non-patent reference (ideally separating scientific publications) (Callaert et al., 2006).

Patent literature can anticipate novelty. Non-patent literature, especially in technological areas closely related to the application, can foresee the inventive step (i.e. the invention should not be trivial or obvious to a skilled person in the particular field of expertise) (Sternitzke, 2007).

In an oft-quoted study on the contribution of public science[7] to industrial technology, Narin et al. (1997) track NPLs of U.S. patents and observe a rapid growth in the dependence of patented technology on U.S. papers. Referencing U.S.-authored research papers had tripled over the six-year period (1987-1988 and 1993-1994) of their data, while there was only a 30% growth in the number of patents over the period.

Patent-to-paper citations are more prevalent in the "*bio nexus*" (Branstetter and Ogura, 2005). Moreover, young and developing technological fields will cite more scientific papers. Hence, new or science-based fields are more appropriate for such studies (Van Looy et al., 2007).

Table 2 depicts selected studies dealing with science and pharmaceutical innovation using patent data[8]. In the patent literature, prior art refers to all available information on the invention and patent family refers to interrelated patents covering the same invention in different countries. The Ward and Dranove (1995) study is interesting in the sense they dissect the pharmaceutical sector research process. They examine the link between basic and applied research in the pharmaceutical sector by dividing R&D

[7] Originating from academia, government and other Federal government funded institutions
[8] Ward and Dranove (1995) do not use patent data

into three stages: government-funded basic research, publication in medical journals, and industry-funded applied R&D. Moreover, they directly link journal articles to industrial R&D.

As a hint of diversity of approaches in the literature, the unit of analysis runs the whole gamut of patent, firm, industry (or technology domain) and country. Likewise, in terms of methods, some studies have used descriptive statistics while others have drawn on more complex methods. However, regardless of methodological rigor, valuable insights have been obtained.

In terms of the link between science and inventive output[9] of the firm, a few nuances are noteworthy. There is positive relationship between using more scientific knowledge in inventions and inventive output in high-tech domains like biotechnology and pharmaceuticals but not necessarily so in other domains (Van Looy et al., 2003; Branstetter and Ogura, 2005). Moreover, there is a positive link between science citations and the generation of inventions that pass FDA approval (i.e. based on linking drugs to patents that underpin drugs) (Branstetter and Ogura , 2005). However, there might be tradeoffs between producing good science and impactful patents for firms. In the words of Gittelman and Kogut (2003) "*scientific ideas are not simply inputs into inventions*" due to different "*selection logics*". They observe that firms cluster their inventive endeavors around important internal or external scientific discoveries; however, these invention clusters are associated with less cited patents. In other words, for firms active in

[9] Strictly speaking, in the pharmaceutical sector, patent output is inventive output and FDA approved drugs are innovative output.

both publishing and patenting, important scientific papers are negatively associated with high-impact innovations (Gittelman and Kogut, 2003).

In terms of the NPL composition, the majority are journal references (50% to 55% on WoS) with "at least 42% of all USPTO non-journal references referring to scientific knowledge". On top of patent references, 40% of non-journal references also refer to technological information (Callaer et al., 2006).

Table 2 Selected Studies Examining Science, Technological Knowledge and Pharmaceutical Innovation

Study	Question	Unit of Analysis	Dependent Variable(s)	Method	Science intensity	Sample	Representative Results
Ward and Dranove (1995)	How does information flow across pharmaceutical research process (NIH funded research-publications -Industry R&D)	Disease category/ year	NIH Expenditures; Medical journal articles in MEDLINE; Industry R&D Expenditures	2 stage GLS	Journal articles	No patent data used	Government funded basic research, with a lag, generates industry-funded applied research. 1% increase in NIH Expenditures boost industry R&D by 0.57-0.76% in same therapeutic area. Cumulative effect of other NIH categories is an increase of 1.26-1.71%.
Narin et al. (1997)	What is the contribution of public science to industrial Technology ?	Patent	Descriptive	Descriptive	Scientific papers in non-patent references	430,226 non-patent references from 397,660 US patents issued in 1987-1988, and 1993-1994.	Rapid growth in the dependence of patented technology on U.S. papers
Gittelman and Kogut (2003)	How does firm's science capability (publication citations) impact the production of highly cited patents in the biotechnology field?	Firm	Cumulative forward citations to a patent (the first U.S. patent in the family)	Negative Binomial; controls for patent family by using first US patent in the family	Mean number of citations to non-patent literature across all of its patents.	Publications and patents of 116 biotechnology firms during the period 1988–1995. (scientific literature 1988–1994; patents 1992–1995)	Important scientific papers are negatively associated with high-impact innovations (publication citations having a negative impact on patent citations)
Van Looy et al. (2003)	Does science-technology interaction help in developing new technology?	Country	Per capita number of patents (country level) in a certain technological field	Variance analysis (ANOVA); no mention of patent family control	Number of references in patents to scientific articles	Granted USPTO-patents for the period 1992-1996.	Positive relationships for high tech domains like biotechnology, pharmaceuticals, organic fine chemistry and semiconductors for other domains no relationship was found.
Branstett	What drives	Firm	Number of	Fixed	" The log	Patents	• Positive

er and Ogura (2005)	the surge in patents citing academic science?		citations; number of approved new drugs; number of citations made by patents to publications; Sales	effects negative binomial; Cobb-Douglas production function; mute on patent family control	of the number of citations to academic science made by the cohort of patents applied for in year *t*	granted 1983-1999 (CHI Research) Publications 1981-1997 (Institute for Scientific Information)	link between science citations and the generation of inventions that pass FDA approval. • Increased use of the knowledge generated by university-based scientists (esp. in bio nexus).
Callaer et al. (2006)	Examining non-patent references' to assess their usefulness as an indicator of science and technology link	Patent	Number and nature of non-patent references in patents	Descriptive statistics; no mention of patent family	Number and nature of non-patent references in patents	Two 5,000 samples of non-patent references from the USPTO and the EPO databases (granted patents applied 1991-2001)	• The majority of NPR are journal references (50% to 55% on WoS) • Differ by technology field • USPTO contain less journal and more non-journal references than expected. For EPO, vice versa. • At least 42% of all USPTO non-journal references refer to scientific knowledge.
Markiewicz (2006)	How does firm-specific characteristics such as experience, knowledge stock, network position, or organizational focus, impact innovative performance?	Firm	Count of citations to non-patent prior art	Negative binomial model; no talk of patent family control	Count of citations to non-patent prior art in the patent	Data from a panel of 83 bio/pharma firms during the 1975-1995 period	• More in-house basic science research and collaboration with university scientists are associated with more exploitation of published scientific research • More citations to published scientific research and a faster pace of knowledge exploitation are associated with a superior economics performance for the firm.
Van Looy et al. (2007)	How does science intensity of	Country	Technological Productivit	"Multiple regressio	Number of references	USPTO patents granted to 8	Science and technology relationship is bi-

	patents relate to technological performance (country level)?		y (number of patents per million inhabitants)	n"; no talk of patent family control	to scientific literature* per 100 patents	European countries 1992-1996	directional and reciprocal ("technological productivity at T+2 is largely associated with past technological productivity (T), a positive and significant relationship with scientific productivity is observed and vice versa.)
Azoulay at al. (2015)	What is the impact of public R&D investments on private-sector patenting?	Disease/science/time (DST)	Patenting by private sector biopharmaceutical firms	Quasi-Maximum Likelihood Poisson; No talk of patent family	non-patent references	USPTO Chemicals and Drugs /Medical patents 1980 and 2012	A $10 million boost in NIH funding leads to a net increase of 2.3 patents

* Prior art is all the available information on the invention

2.2.2 Knowledge Recombination and Innovation

The recombinant view of innovation contends that new knowledge is created by the combination of new components or new combinations of existing components. One of the earliest references to this idea is traced to Schumpeter's research (Phene et al., 2006).

A strand of innovation literature has focused on the degree of differentness of knowledge components used in the recombination process. Different terms and proxies have been defined and used. Knowledge diversity, knowledge distance, and knowledge heterogeneity are among the common terms. For the purposes of this study, in line with the literature, knowledge diversity is defined with reference to the set of knowledge components that have been used by the firm in inventive efforts. For instance, Kaplan and Vakili (2015) draw on technological diversity in their study. Knowledge heterogeneity is

defined in terms of differentness along the set of knowledge components not in line with the firm's extant technological expertise[10]. The *term* heterogeneity is borrowed from Tsai (2017) who defines heterogeneity in terms of knowledge available for recombination on innovative outcomes at the organizational level.

In addition to the degree of differentness of knowledge components, the technological space of a firm can also be classified based on the source (or locus) of the knowledge. Technologically distant knowledge comes from outside the industry and technologically proximate knowledge is sourced from within the industry (Ahuja and Lampert, 2001).

In terms of theoretical stance, there are two competing positions. The "*tension view*" asserts that deep knowledge can lead to myopia to the extent that the recombination of distant or diverse knowledge is needed in order to see new ideas (Kaplan and Vakili, 2015). Diverse sources of knowledge are more likely to come from dissimilar contexts; hence, more likely to be novel and nonredundent to the firm (Phene et al., 2006). Innovations building on diverse knowledge sources (e.g. patents from different technology fields) are expected to lead to original outcomes (Squicciarini et al., 2013). At the firm level, experimenting with novel knowledge helps stave off the "*familiarity trap*"; i.e. preference for the refinement of familiar technologies rather than exploring new ones (Ahuja and Lampert, 2001).

On the other hand, according to the "*foundational*" view of creativity, a deep understanding of the foundations, assumptions and weaknesses of a particular knowledge

[10] The literature has tended to refer to this as knowledge distance

domain helps detect anomalies hence conducive to breakthrough innovations. Wide recombination can be disadvantageous to innovation because only a penetrating attempt can produce breakthroughs (Kaplan and Vakili, 2015). Rather than being independent of the past, even the most radical innovations are grounded in the past. A frame of reference to the past provides coherence and resonance with the intended audience. Moreover, without knowing the discipline, going beyond it might be a tall order (Weisberg, 1999).

To sum up, the degree of differentness of knowledge components used in innovation has been studied in various contexts and research designs. In addition, the degree of familiarity (i.e. depth vs. breadth) with knowledge components and the locus of knowledge components (form inside or outside firm/country) can be other dimensions. Finally, the impact of diversity and heterogeneity of knowledge on innovative output may be contingent upon various organizational factors.

2.3 Research Question and Hypotheses

Major advancements in scientific and technological inputs into the drug discovery process coupled with the R&D productivity paradox, calls for examining the link between scientific and technological knowledge recombination and innovative output of firms. Hence, the research question of this study: "*How does codified science use and the knowledge recombination profile of firms impact innovative output?"*

The unit of analysis is "firm-year-patent[11]", in line with the literature focusing on firm patenting. For instance, the units of analysis in Fabrizio (2009) is firm-year and firm-patent class-year. The focus on patents gives us the chance to look at the earlier

[11] More specifically "patent family"

stage of the innovation pipeline. New drug and drug enhancement innovations are more representative of directed research. Patents, on the other hand, are more closely associated with earlier stages of the research endeavor (Cardinal and Hatfield, 2000). Thus, using patents helps edge closer to the discovery phase of drugs. Moreover, there is not much variation in annual FDA-approved new drugs ("*broadly flat*" since the 1950s according to Scannell et al., 2012) and firms filing the drug applications. For instance, according to Munos (2009) half of the approved NMEs since 1950 are produced by 21 companies; ironically, half of these companies do not exist today.

The first hypothesis examines the link between codified science and inventive output. The literature reviewed here indicates that firstly, biopharmaceutical inventions rely on academic science more than other fields (Branstetter and Ogura, 2005; Van Looy et al., 2003). Hence, it is tempting to assume that more reliance on scientific knowledge in the inventive process, will lead to higher quantitative output of inventions. However, utilization of science in the inventive process calls for certain capabilities. For instance, Markiewicz (2006) observes that in-house basic science research and collaboration with university scientists (i.e. "absorptive capacity-related activities") are related to more exploitation of published scientific research.

Small number of backward citations per patent may indicate an invention in a relatively new technology area with few sources of information available. Alternatively, it can reflect failure to cite relevant prior art (Lanjouw and Schankerman 2001). It is a reasonable assumption that the former case would be a rare occurrence as breakthroughs

do not happen frequently (hence the term). If the latter holds, the patent has less chance of passing patent office examination.

Combining the absorptive capacity notion with the recombinant view of innovation, bits of codified science are considered components of potential new combinations; however, there needs to be a certain critical mass of components to allow for experimentation and capacity building. Unless a certain critical mass (of components and capabilities) is achieved there will be no increased inventive output. Hence,

Hypothesis 1: Higher reliance on codified science in inventions, *ceteris paribus*, will improve firms' inventive output only after fulfilling a certain threshold.

The second hypothesis is based on the properties of the knowledge used in the inventive process. Knowledge diversity refers to different knowledge components that the firm is using in inventive efforts. Drawing on the recombinant view and the foundational view of creativity, diverse knowledge components can be assembled into many innovations, if the firm has already used these knowledge components and has knowledge of their properties. Hence:

Hypothesis 2a: Technological knowledge diversity, *ceteris paribus*, will be associated with higher inventive output.

Knowledge heterogeneity refers to differentness along the set of knowledge components not in line with the firm's extant technological expertise; hence, less familiar to the firm. As such, most probably, they will be recombined into fewer innovations. Based on the tension view these components of knowledge can help the firm break the familiarity trap and lead to original outcomes (Kaplan and Vakili, 2015; Ahuja and

Lampert, 2001; Squicciarini et al., 2013). This chimes with the common view in the literature that local search is expected to be associated with exploitation rather than exploration (Kaplan and Vakili, 2015). In other words, given the effort required in recombining these knowledge components, it can be expected that the output would be of higher quality at the expense of quantitative output. Hence:

Hypothesis 2b: Knowledge heterogeneity, *ceteris paribus*, will be associated with higher *quality* inventive output.

Patent data will be used to test these hypotheses. The remainder of this section will deal with operationalizing the aforementioned concepts. Inventive output (dependent variables) will be measured by the firm's annual patent family output. Annual patent family output weighed by forward citations (i.e. of the first US patent of the family in the first 7 years) will proxy quality weighted innovative output.

Non-patent references will be used to proxy codified science which refers to the scientific knowledge expressed in journal papers, books, texts, and other technical media. A successful recombination is proxied by the grant of a patent. Knowledge diversity and heterogeneity are measured by the patent references of the focal patent. Hence, a more accurate term would be "*technological*" knowledge diversity/heterogeneity (in the spirit of Kaplan and Vakili, 2015).

The patent originality index, commonly used as a measure of technological knowledge diversity, captures the breadth of the technology fields on which a patent rests and is constructed as follows (Squicciarini et al., 2013).

Equation 1 The Originality Index

$$Originality_p = 1 - \sum_j^{n_p} {S_{pj}}^2$$

S_{pj} is the percentage of citations made by patent *p* to class *j* out of the n_p IPC four-digit (or seven-digit) patent classes in the patents cited in the patent *p*.

The patent radicalness index is used to proxy for the technological knowledge heterogeneity (or distance). Radicalness (à la Shane) measures the degree to which a patent builds on (through backward citations) a diversified array of technologies; i.e. patents from fields other than own assigned field (Johnstone et al., 2015). It is specified as follows:

Equation 2 The Radicalness Index

$$\text{Radicalness}_P = \sum_j^{n_p} CTj/n_p;\ IPC_{pj} \neq IPC_p$$

CTj is the count of IPC four digit codes of patent *j* cited in patent *p* that is not assigned to patent *p*, out of n IPC classes in the backward citations counted at the most disaggregated level available (up to the 5th hierarchical level). The higher the ratio, the more diversified the range of technologies underlying the patent (Squicciarini et al., 2013).

2.4 Data

To test the hypotheses, firm level patenting data and firm level financial data (for control variables) were required[12]. The OECD Patent Quality Indicators database and its Triadic Patent Families database (February 2016 release) were used for main patent data and indicators. The LENS[13] patent search engine and the COMETS[14] (Connecting Outcome Measures in Entrepreneurship, Technology, and Science) patent assignee data files were used for comparison and verification of assignees (i.e. patent owners at filing).

The Compustat database (by the Wharton Research Data Services: WRDS) was the source of financial data. It is more reliable and comprehensive than alternative data sources (i.e. PhRMA and NSF data) for the pharmaceutical industry (Golec and Vernon, 2008).

To extract pharmaceutical patents, patents assigned to the international patent class (IPC) A61K, representing pharmaceutical patents, were selected. This patent class excludes patents on cosmetics (Squicciarini et al., 2013; Schmoch, 2008). This yielded 117,442 observations.

The concept of "*Triadic Patent Families*" was used to identify unique inventions. A patent family is a set of related patents linked by one or more patents called priority filings that have been filed in several countries (Squicciarini et al., 2013). The priority patent of a patent family is the first patent application filed for the invention of each family (Criscuolo, 2006). Triadic families have been filed in the U.S., Europe and Japan. Drawing on the OECD Triadic Patent Families (TPF) concept, 80,683 patents were

[12] A more detailed data construction account is available in appendix A.
[13] "The Lens." Accessed April 11, 2017. https://www.lens.org/lens/.
[14] "COMETS." Accessed April 11, 2017. http://www1.kauffman.org/COMETS/.

identified. After assignee verification, 56,804 patents remained in the sample. To construct the panel and link to Compustat, assignees with five or greater years of patenting activity (1170 assignees) were selected.

Figure 5 depicts the usable data points after merging financial data (total of 631 firms). The complete list of firms is available in Appendix A. The final data set, comprising of 29,554 patents (i.e. in firm-year-Triadic-Patent-Family unit of analysis), is unbalanced and firm data points are not necessarily chronologically complete (e.g. 1998, 1999, 2010, 2011, 2012).

Figure 6 depicts the Global Industry Classification (GIS)[15] of the firms, indicating the diversity of firms contributing to the pharmaceutical invention pool. These firms may not be a member of pharmaceutical industry associations such as PhRMA. This indicates one advantage of the bottom-up sample building (i.e. starting from a patent pool rather than a list of companies) used in this study.

[15] The GIS is an industry taxonomy developed in 1999 by MSCI and S&P for the global financial community use. It consists of 11 sectors, 24 industry groups, 68 industries and 157 sub-industries (Wikipedia).

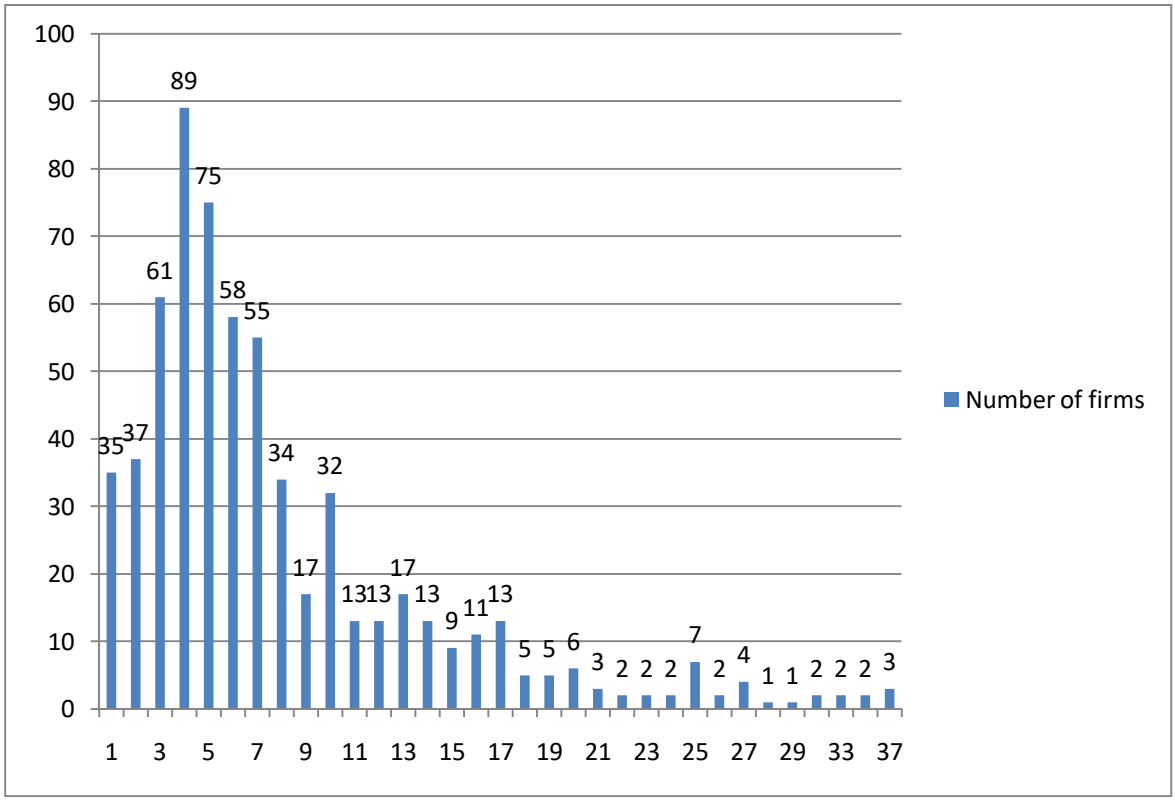

Figure 5 Years of Data for Total Panel Firms (total firms=631)

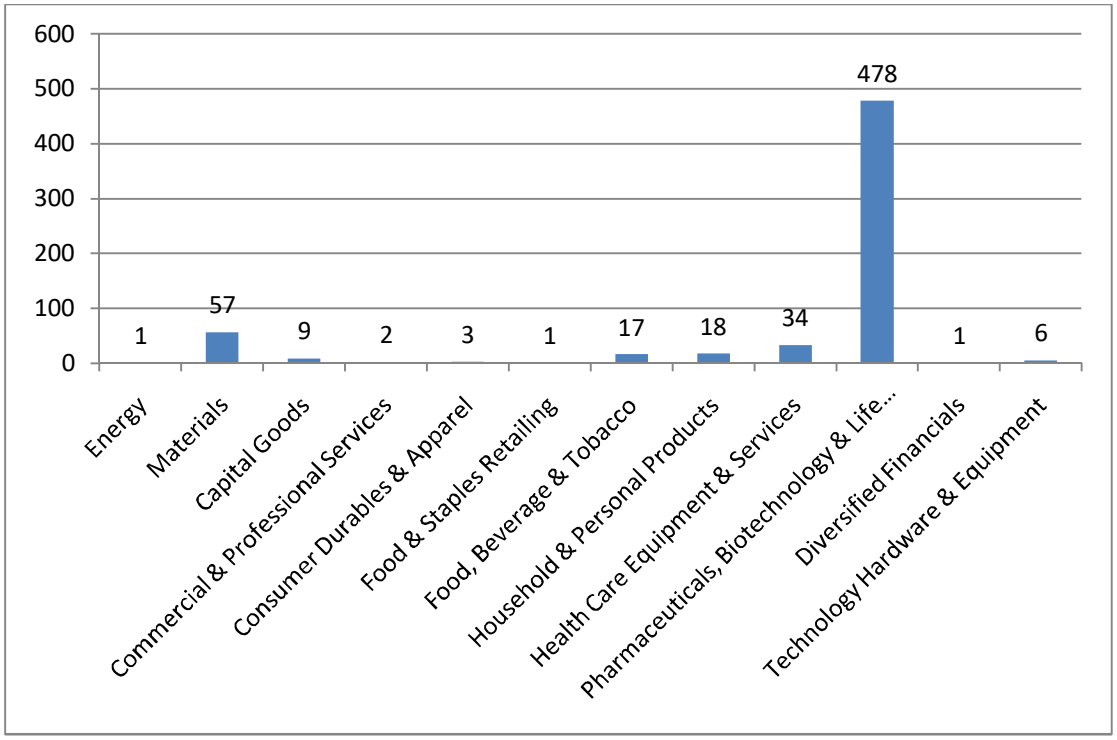

Figure 6 Firms by Global Industry Classification (GIC) Groups

2.5 Methods

Equation 3 depicts the basic model used to test the hypotheses. An attempt has been made to construct a model representing the main drivers of innovation in the sector; i.e. R&D spending, firm size, firm R&D profile (i.e. science intensity. radicalness and originality of patents), and total realized market size proxied by annual sales of firms active in patenting in the same year. Table 3 defines the variables, including the weighted dependent variable.

The alternative dependent variable is constructed to test hypothesis 2b. It follows Trajtenberg's (1990) method of weighting patents by forward citations (i.e. citations received). NPL intensity is calculated by equation 4 which is the ratio of total number of

NPL per year for each firm over total prior art of the firm in the same year, normalized by the number of patent families the company has in the same year.

Equation 3 Basic Regression Model

$$Patents_{it} = \beta_0 + \beta_1 R\&Dintesity_{it} + \beta_2 NPLintensity_{it} + \beta_3 NPLintensity_{it}^2 + \beta_4 logAnnualSales_t + \beta_5 logEmploy_{it} + \beta_6 Original_{it} + \beta_7 Radical_{it} + Y_t + F_i + \epsilon$$

Equation 4 NPL Intensity

$$\text{NPL Intensity}_{\text{it}} = (\sum NPL_{it} \div \sum Prio\ Art_{it}) \div \sum Patent\ families_{it}$$

Table 3 Variables and Definitions

Variable	Definition
	Dependent variable
$Patents_{it}$	The count of the number of Triadic Patent Families of firm *i* in year *t*
$WeightedPatents_{it}$	The citations weighted first patents of Triadic Patent Families of firm *i* in year *t*. Calculated (Trajtenberg, 1990): $WeightedPatents_{it} = \sum_{i=1}^{n_t}(1 + C_i)$
	Explanatory Variables
$NPLintensity_{it}$	NPL intensity for the patents of firm *i* in year *t*.
$NPLintensity_{it}^2$	The quadratic term of NPL intensity
$Original_{it}$	Average originality index of the patents of firm *i* in year *t*
$Radical_{it}$	Average radicalness index of the patents of firm *i* in year *t*
	Control Variables
$R\&Dintesity_{it}$	Is the contemporaneous ratio of R&D expenditures on sales calculated per year per firm
$logAnnualSales_t$	Logarithm of annual sales of all panel firms active in year *t*
$logEmploy_{it}$	Logarithm of number of employees of firm *i* in year *t*
Y_t	Year fixed effect
F_i	Firm fixed effect

A brief note on control variables is useful. R&D intensity and employment of firms represent firm-level control variables. According to Austin (2006) there is a relatively close positive relationship between pharmaceutical firms' current R&D expenditure and current sales. The relative stability of the relationship implies that firms

find it most profitable to invest a steady portion of any additional sales revenues in their own drug research. A large number of studies have dealt with the impact of different R&D lags on patenting performance and concluded that R&D lags are "*very poorly identified due to the high within-firm correlation in R&D*". Also, previous research has indicated that inclusion of firm [fixed] effects may turn lags insignificant. As an upshot of these studies, it is almost "*de rigueur*" to use contemporaneous R&D to predict patenting performance (Somaya et al., 2007). This study controls for the contemporaneous *R&D intensity* (R&D/Sales) instead of straight R&D expenditures as a measure taking account of also the sales of the firm. This is a variation of the abovementioned argument.

The annual employment size represents the size of the firm. A vast literature has dealt with the impact of firm size on innovation, for the purposes of this study, suffice it to say that, some studies have explored a direct link between firm size and innovation Most of these studies are motivated by the seminal work of Joseph Schumpeter. He espoused two different takes on the relationship between firm size and innovation during his career. He first exhorted that the entrepreneur outside the firm (i.e. small entities) is important in innovation. This is position is known as "Schumpeter Mark I"; "widening" or "creative destruction". Later, he emphasized the importance of large firms in promoting innovation via endogenous innovative processes. This is referred to as "Schumpeter Mark II"; "deepening" or "creative accumulation" (Acs and Audretsch, 1988; Breschi et al., 2000). Other studies have explored the moderating effect of e.g. market structure and technological regimes, or have proposed nonlinear inverted-U

shaped relationship between firm-level R&D effort and number of competitors (Peneder and Wörter, 2013).

"Total annual sales" of panel firms is meant to be a proxy variable for the market size in that year. Admittedly, it is a noisy proxy since some firms are not strictly speaking exclusively in the pharmaceutical industry. However, it may somehow be indicative of resources available from cognate industries (e.g. chemicals) that also patent in pharmaceutical fields. At the estimation level, including year dummies may effectively cancel it out[16]. Emphasis on market goes as far back to Schmookler's seminal work, "*Invention and Economic Growth*", economists have emphasized the importance of profit incentives and target market size in innovation. To emphasize this point, Shmookler titles two chapters "*The amount of invention is governed by the extent of the market*" (Acemoglu and Linn, 2004).

Firm fixed effects are dealt with in the panel fixed effect estimation model. Year dummies are included to control all industry-level changes affecting all firms (Fabrizio, 2009) such as economic cycles or new technology emergence (e.g. biotechnology).

In addition to the abovementioned variable, subsample analyses will be conducted to test pre-1995 science intensity based on a dummy variable representing patenting activity before 1995. Moreover, a dummy variable for biotechnology firms will test the possible difference in biotechnology firms' inventive output.

In terms of estimation method, Poisson and negative binomial regression have been applied to patent count data in previous studies (Hall and Ziedonis, 2001; Gittelman

[16] Courtesy of a colleague

and Kogut, 2003). The Poisson distribution assumption is that the dependent variable has equal mean and variance. When the data is overdispersed, the assumption does not hold and the negative binomial analysis is commonly used (Penner-Hahn and Shaver, 2005). The Stata overdispersion test based on a straight negative binomial regression [17] on both dependent variables indicates that overdispertion exists (e.g. for patent family count: Likelihood-ratio test of alpha=0: chibar2(01) = 6543.65 Prob>=chibar2 = 0.000). This is consistent with Gittelman and Kogut (2003) who mention that patent citations exhibit a great deal of overdispersion therefore they use a negative binomial model. They also use cluster regression models for robustness checks hence the same strategy is adopted here.

In terms of functional form, the expected number of patents applied for during a year is usually conceived as "an exponential function of the firm's R&D spending and other characteristics" (Hall and Ziedonis, 2001); hence, the logarithmic transformation in most models in the literature. However, Somaya et al. (2007) transform all "*size-dependent variables*" by logs. Consequently, employment is log transformed. The "annual sales" is log transformed to reduce its skewness and bring it closer to other variables.

2.6 Results and Discussion

Analysis was conducted by Stata statistical software package version 12.0. For practical and theoretical reasons, size of the coefficients will not be interpreted. Firstly, main variables are indices (e.g. originality and radicalness) and ratios (science intensity). Secondly, the dataset is comprised of proxy variables for a complex phenomenon and

[17] "*nbreg*" no fixed effects or clustering)

some are consciously manipulable (e.g. number of non-patent references). Hence, stating a magnitude of relationship could project an image of false accuracy or causality. To illustrate the latter point, according to Callaert et al. (2006) "*in absolute terms, USPTO patents hold approximately 3 times more references on average than EPO patents*" that could reflect the broader differences in "*rationale of citing prior art*". In other words, same invention might be patentable with varying number of backward citations even in one patent office due to examiner preferences and idiosyncrasies[18]. Finally, there are studies only looking at the direction of relationships with no attempts at coefficient size interpretations such as Gittelman and Kogut (2003).

Table 4 depicts summary statistics of the variables based on Stata panel summary command[19] decomposing the standard deviation into between and within components[20]. For example, looking at the first entry, the average number of patent families for all year-patent rows of data (N =4940) is about 9.3. The average annual number of patent per each 631 firms varied between 1 and 47.54. The within refers to the deviation from each firm's average, i.e. some firms deviated from their average -39.9 to 94 patents. T-bar refers to the average number of years the variable was observed for each firm (Adapted from Schüpbach, 2014).

Table 5 contains the pairwise correlation of variables. The negative correlation between employment and NPL intensity may indicate that larger firms on average cite less codified scientific material (e.g. journal papers).

[18] There is a literature on examiner inserted backward citations
[19] xtsum
[20] https://www.stata.com/manuals13/xtxtsum.pdf

Table 6 depicts the results. Negative binomial panel with firm fixed effects and negative binomial with firm clustered errors (for robustness checks) were run on the two types of dependent variables. Models 1 and 4 are the main models because they include year dummies and are based on the two types of dependent variables. The first hypothesis regarding the impact of codified science is supported across all models, indicating that there is a negative significant association and then a positive significant association of incorporating codified science into inventions. The initial negative impact is in line with arguments that large numbers of backward citations are associated with more incremental innovations (Squicciarini et al., 2013). Drawing on the recombination analogy, the latter phase positive impact could be due to the assimilation of a larger pool of scientific knowledge (i.e. building blocks of innovation) in the inventive efforts; hence, being able to make a larger number of inventions.[21]

Hypothesis 2a based on the knowledge diversity argument is not supported in the main model (1) based on count dependent. The variable (originality) is positively and significantly associated with better quality-weighted inventive output (model 4). An explanation behind the observation could be that assimilation and recombination of diverse knowledge components is challenging but when achieved leads to better quality inventive output.

Hypothesis 2b based on the knowledge heterogeneity argument is not supported as formulated. In fact, the reverse relationship holds with heterogeneous knowledge

[21] As mentioned earlier, coefficient size interpretation is not sough here. Coefficients indicate the impact of independent variables on the outcome variables' "log of expected counts" or "log counts"; e.g. in model 1 one unit increase in NPL intensity reduces log of expected counts of patent families by 4.607 (statistically significant at the 1% level) before the critical mass is achieved.

recombination significantly associated with reduced inventive output in both outcome variables. One reason for this could be that drawing on heterogeneous knowledge (i.e. distant or not close with the firm's extant technological expertise), requires more time and resources which in turn reduces the inventive output of the firm.

A brief note on the control variables is in order. If we focus on model 1 and 4, employment is negatively associated with inventive output but not with quality weighted output. It might be that smaller firms have less inventive output but the quality of their output does not differ from large firms. Given the R&D intensive nature of innovation in the sector, this observation seems plausible. The annual sales variable is not significant. The explanation could be because of its poor proxy quality or being cancelled out by year fixed effects. One explanation on the usefulness of such a variable would be a cross subsidization argument. Given the fact that not all firms come from the pharmaceutical sector (about 24% from outside), some financial resources may be transferred to the pharmaceutical inventive efforts from other sectoral activities. Likewise, the constructed R&D intensity measure does not show any significance. Measurement and variable construction issues aside, a real explanation could be that in the R&D-intensive pharmaceutical sector, it does not differentiate inventive output of firms.

Table 4 Summary Statistics

Variable		Mean	Std. Dev.	Min	Max	Observations
$Patents_{it}$	O	4.592308	9.29794	1	137	N =4940
	B		4.537825	1	47.54054	N=631
	W		6.010186	-39.9482	94.05177	T-bar=7.82884
$Patents_{it}$ Weighted by 7-year citations	O	39.32733	85.92369	1	1460	N=4940
	B		38.0221	1	325.973	N=631
	W		64.52015	-283.646	1248.354	T-bar=7.82884
$Annual\ Sales_t$	O	$54,700,000.00	$47,700,000.00	$10,237.56	$155,000,000.00	N=4940
	B		$36,000,000.00	$10,237.56	$155,000,000.00	N=631
	W		$38,100,000.00	-$46,400,000.00	$173,000,000.00	T-bar=7.82884
$R\&D\ Intesity_{it}$	O	16.67462	441.4469	-90.3846	25684.4	N=4367
	B		361.2019	-16.2359	8666.188	N=593
	W		327.6426	-7442.21	17034.89	T-bar=7.36425
Employment	O	18.61606	37.55868	0.002	361.796	N=4043
	B		29.33049	0.007333	343.9965	n=567
	W		9.478934	-81.3815	117.2185	T-bar=7.13051
$NPL\ Intensity_{it}$	O	0.251009	0.250072	0	1	N=4934
	B		0.175546	0	0.946429	n=631
	W		0.192682	-0.40051	1.144114	T-bar=7.81933
$Original_{it}$	O	0.838273	0.113728	0	0.981557	N=4885
	B		0.074331	0.394531	0.962883	n=631
	W		0.095405	0.034581	1.232804	T-bar=7.74168
$Radical_{it}$	O	0.270516	0.187558	0	1	N=4885
	B		0.123775	0	0.732303	n=631
	W		0.153606	-0.26529	1.121602	T-bar=7.74168

* O=Overall; B=Between; W=Within; Dollar amounts are in millions; Employment is in thousands

Table 5 Pairwise Correlations

Variable	Correlation (Significance)					
	$R\&D\ Intesity_i$	$NPL\ Inten$	$Annual\ Sale$	Employment	$Original_{ii}$	$Radical$
$R\&D\ Intesity_{it}$	1					
$NPL\ Intensity_{it}$	0.0044	1				
	(0.7731)					
$Annual\ Sales_t$	0.0072	0.0590	1			
	(0.6342)	(0.0000)				
Employment	-0.0190	-0.2062	-0.0413	1		
	(0.2413)	(0.0000)	(0.0087)			
$Original_{it}$	0.0198	0.0746	0.2712	-0.0840	1	
	(0.1937)	(0.0000)	(0.0000)	(0.0000)		
$Radical_{it}$	-0.0135	0.0765	-0.0660	0.0092	0.23	1
	(0.3737)	(0.0000)	(0.0000)	(0.5620)	(0.0000)	

Table 6 Regression Results – Full Sample

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	**Number of Patent Families**			**Weighted Number of Patent Families (first US patent 7-year citations)**		
Variables	**Neg. Binomial Panel (Fixed Effects)**	**Neg. Binomial Panel (Fixed Effects)**	**Neg. Binomial (Clustered Errors)**	**Neg. Binomial Panel (Fixed Effects)**	**Neg. Binomial Panel (Fixed Effects)**	**Neg. Binomial (Clustered Errors)**
$R\&D\ Intesity_{it}$	-0.0000224	-0.0000113	-0.0000333***	0.00000395	0.000011	-0.0000114**
	(0.0000337)	(0.000034)	(0.00000681)	(0.0000316)	(0.0000319)	(0.00000498)
$NPL\ Intensity_{it}$	-4.607***	-4.609***	-7.810***	-2.628***	-2.966***	-8.179***
	(0.219)	(0.220)	(0.488)	(0.196)	(0.207)	(0.597)
$NPL\ Intensity_{it}{}^2$	3.571***	3.549***	6.880***	1.730***	2.227***	6.825***
	(0.272)	(0.273)	(0.546)	(0.244)	(0.256)	(0.653)
Natural log_$Annual\ Sales_t$	0.129	0.156***	0.0729	0.0393	0.0453***	-0.0431
	(0.115)	(0.00866)	(0.106)	(0.126)	(0.00833)	(0.0795)
Natural log _Employment	-0.141***	-0.175***	0.148***	-0.0101	-0.0596***	0.0264
	(0.0191)	(0.0170)	(0.0153)	(0.00972)	(0.00921)	(0.0186)
$Original_{it}$	0.144	0.583***	1.670***	0.334**	0.438***	2.165***
	(0.149)	(0.146)	(0.236)	(0.148)	(0.145)	(0.280)
$Radical_{it}$	-0.244***	-0.321***	-1.315***	-0.211***	-0.202**	-0.910***
	(0.0855)	(0.0850)	(0.175)	(0.0790)	(0.0821)	(0.185)
Year Dummies	Yes	No	Yes	Yes	No	Yes
Constant	0.0879	-0.136	-0.673	-1.268	-0.402***	1.117
	(1.371)	(0.160)	(1.051)	(1.533)	(0.153)	(0.874)
lnalpha			-0.982***			0.00612
			(0.0921)			(0.0335)
Observations	3,740	3,740	3,787	3,740	3,740	3,787
Wald chi2(43)/(7)	1654.49	1201.03	1712.05	1310.39	458.58	2713.44
Prob > chi2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Log likelihood	-5937.3553	-6111.1785	-8475.2234	-12840.66	-13241.817	-16261.001
Number of Firms (gvkey)	493	493	(540)	493	493	(540)

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1; gvkey is Compustat unique firm code

To explore some of the aforementioned nuances and results, a subsample analysis was conducted. Table 7 depicts the details of this effort. An important subsample would be the firms primarily active in the pharmaceutical sector (models 1 and 2). This was achieved by separating firms active in the "Pharmaceuticals, Biotechnology & Life Sciences" industry group based on the GIC industry classification. Four firms were missing this designation in the dataset and were excluded from the regression.

Another nuance would be the impact of longer history of inventive activity or more experienced (models 5 and 6). This was achieved by separating firms with 20 years of patenting data; however, the eventual analysis has firms with minimum of 13 years of observations due to other missing data. This analysis could be important for two reasons. Firstly, a longer period of activity could potentially point to hidden regularities in innovative activities. Secondly, these firms could be different because of history, survival or other attributes not fully accounted for in the models. In all four models the curvilinear relationship between science intensity and inventive output still holds. However, the knowledge diversity proxy is not significant in all models, even the model based on quality weighted output. This may indicate for pharmaceutical firms, the ability to recombine diverse knowledge does not give them an edge in inventive output. In other words, this ability is pervasive and a qualifier for inventive activity.

Hypothesis 2a based on the knowledge diversity argument is not supported in the main model (1) based on count dependent. The variable (originality) is associated with better quality-weighted inventive output (model 4). An explanation behind the observation could be that assimilation and recombination of diverse knowledge

components is challenging but when achieved leads to better quality inventive output. The negative impact of knowledge heterogeneity on inventive output is observable in three models. It is only not significant in the model based on quality weighted output for all pharmaceutical firms. Hence, recombining heterogeneous knowledge can be assumed to negatively impact inventive output possibly because of the effort involved.

Another interesting analysis would be to see how science intensity might have changed in the wake of the 1990s scientific and technological advances. To this end a dummy variable representing pre-1995 inventive activities was created and interacted with the science intensity variable (models 3 and 4). First wave of biotechnology drugs were granted FDA approval in the 1990s (see figure 1). Results indicate that before 1995 higher science intensity was associated with more and better quality inventive output. One explanation could be that science intensity increased over time and without assimilating and utilizing a critical mass it does not have any advantage in inventive activities.

The last subsample analysis pertains to biotechnology firms. A dummy variable was created for firms active in biotechnology[22]. Model 7 run on quality weighted inventive output[23] indicates that biotechnology firms have a better performance within the broader pharmaceutical sector. This is consistent with arguments in the literature touting biotechnology firms as the innovative engine of the pharmaceutical sector (e.g.

[22] i.e. assignment to the GIC Industries Code of 352010

[23] The model would not run on the count dependent variable because some firms have only one output per year (i.e. "*constant within group*" dropping of biotechnology dummy).

Perova, 2014 trifecta model[24]). However, they have less total output. This point is corroborated by summing the total citations weighted dependent variables for biotechnology (i.e. 49774) and non-biotechnology (i.e. 117565) firms, and a significant *t-test* of weighted dependent variable based on a biotechnology dummy. The first activity year for biotechnology firms is 1982, versus 1977 for non-biotechnology firms.

One caveat of the abovementioned results is that prior art[25] on patent documents contain citations inserted by the patent examiner (Fabrizio, 2009). The data used in this study does not distinguish between examiner and applicant cited prior art. However, the underlying argument of characterizing inventions based on cited prior art can still be valid. For instance, the argument would be that firms with inventions underpinned by larger NPLs (i.e. either determined by examiner or inventors) have better inventive output performance than those firms not meeting the unknown threshold. Even if we manage to separate inventor and examiner citations, the inventor might be able to base the invention on larger number of fragmented codified science references (e.g. journal articles) or a few broader science references covering more ground. Hence, the issue of qualitative contents of cited prior art arised which is rarely addressed in studies. This might be another issue why coefficient interpretation might be redundant.

[24] Simple model depicting division of labor between: biotech firms (doing applied research and biomolecule design), public institutions (creating new fundamental knowledge and mapping the scientific landscape) and large pharma firms (doing clinical trials, manufacturing and commercialization) (see Appendix A).

[25] As a reminder, NPL intensity; originality and radicalness indices are all based on references cited on the patent document.

Table 7 Subsample Regression Results – Negative Binomial Panel Fixed Effects

	(1) Only Pharmaceutical Sector	(2)	(3) Pre-1995 Science Intensity (all sample)	(4)	(5) Pharma Sector Firms with 20 or more years of patenting*	(6)	(7) Biotech Dummy (Pharma Sector)
Dependent Variable	**Count**	**Weighted**	**Count**	**Weighted**	**Count**	**Weighted**	**Weighted**
$R\&D\ Intesity_{it}$	-0.0000217	0.00000306	-0.0000169	0.00000617	0.0721	0.0614	0.00000185
	(0.0000333)	(0.0000304)	(0.0000347)	(0.0000327)	(0.0526)	(0.0627)	(0.0000301)
$NPL\ Intensity_{it}$	-5.351***	-3.350***	-2.258***	-1.549***	-9.967***	-8.025***	-3.380***
	(0.258)	(0.232)	(0.0986)	(0.0882)	(0.741)	(0.784)	(0.233)
$NPL\ Intensity_{it}^{2}$	4.239***	2.471***			9.960***	7.711***	2.481***
	(0.308)	(0.277)			(1.023)	(1.032)	(0.276)
Natural log_$Annual\ Sales_t$	0.140	0.0228	0.113***	0.00544	0.00161	-0.0742	0.0186
	(0.128)	(0.140)	(0.00978)	(0.0102)	(0.192)	(0.206)	(0.140)
Natural log _Employment	-0.170***	0.00242	-0.184***	-0.0548***	-0.145***	0.0251	0.0153
	(0.0203)	(0.0117)	(0.0170)	(0.00921)	(0.0446)	(0.0319)	(0.0136)
$Original_{it}$	0.0241	0.236	0.153	0.148	-0.00760	0.459	0.206
	(0.182)	(0.183)	(0.144)	(0.143)	(0.358)	(0.424)	(0.184)
$Radical_{it}$	-0.178*	-0.133	-0.327***	-0.186**	-0.421**	-0.471*	-0.138
	(0.104)	(0.0979)	(0.0853)	(0.0823)	(0.214)	(0.252)	(0.0979)
Pre-1995 *$NPL\ Intensity_{it}$			0.793***	0.854***			
			(0.176)	(0.161)			
Pre-1995 Dummy			-0.496***	-0.453***			
biotechnology							0.139*
							(0.0727)
Year Dummies	Yes	Yes	No	No	Yes	Yes	Yes
Constant	0.0107	-1.037	0.884***	0.492**	1.293	-0.190	-1.022
	(1.526)	(1.697)	(0.194)	(0.202)	(2.242)	(2.444)	(1.697)
Observations	2,758	2,758	3,740	3,740	589	589	2,758
Wald chi2(43)/(8)/(44)	1528.83	1139.31	1128.60	434.25	845.96	657.55	1143.75
Prob > chi2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Log likelihood	-4511.6966	-9778.8654	-6134.7675	-13248.24	-1571.5376	-2898.5852	-9777.0282
Number of Firms	374	374	493	493	25	25	374

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1; * given financial data availability the minimum number of observations turns out to be 13

2.7 Conclusion

This study explored the inventive output of firms active in pharmaceutical patenting. The first observation was that about one quarter of firms contributing to the pool of pharmaceutical inventions have primary activity outside the pharmaceutical industry. This might be especially important as the pool of inventions represent Triadic Patent Families.

Robust support was observed for the non-linear; first negative and then positive; association of science-intensity to the inventive output of firms. Moreover, knowledge diversity was positively associated with quality-weighted inventive output while knowledge heterogeneity negatively influenced both inventive output and quality-weighted inventive output. The latter results are in line with the foundational view of creativity that contends a deep grasp of domain knowledge is important in developing breakthrough innovations. Nuanced subsample analysis revealed that for firms primarily active in the pharmaceutical sector higher knowledge diversity in inventive output did not improve performance. This may indicate the ability to recombine diverse knowledge as a necessary and pervasive capability for inventive activities. Moreover, pre-1995 reliance on codified science was higher, indicating a possible shift to higher science intensity of innovative activities in the past few decades.

In terms of strengths, a relatively new dataset by OECD was used in the study. Moreover, the bottom-up approach of starting from a patent pool lent useful insight on patenting from outside the pharmaceutical sector. Focusing on triadic patent families helped capture more valuable inventions underpinning real innovations. While the study

drew on constructed proxies for complex phenomena such as science utilization in innovation, the conceptual framework can be tested in future studies.

In terms of shortcomings, the construction of the proxy variables in the form of indices and averages may have dampened the explanatory power of underlying variables. This could be accentuated for the originality and radicalness indices that in original form are complex constructs. Additionally, the dataset did not allow distinguishing between examiner and applicant inserted references.

The sample patents were all assigned to the A61K IPC patent class. Therefore models do not control for technology class as it is customary in some studies. There are technical problems in controlling for technology field based on patent classes. The IPC classification is updated periodically and is based on the economic importance; i.e. industry and profession; as opposed to the technical focus of the U.S. classification based on structure and function of the invention (Lerner, 1994). The IPC is more desirable in economic research. However, unlike the U.S. classification, there is no main IPC class to which the patent is assigned to, making controlling for IPC class problematic. It can be argued that patent classes may altogether be a poor choice for controlling pharmaceutical innovation fields. For instance, thalidomide was initially used in the late 1950s and early 1960s for the treatment of nausea in pregnant women. As it later transpired in the 1960s, thalidomide resulted in severe birth defects. The use of thalidomide was banned in most countries at that time; however, it proved useful for leprosy and later, multiple myeloma (Kim and Scialli, 2011). This shows how the application of a drug substance has changed dramatically over time.

Finally, a brief note on future research is useful. Most of the variables used in the study are proxies of complex phenomena. Hence, searching for appropriate instruments could be an interesting foray for quantitative research. Moreover, in depth understanding of the abstract notions of knowledge diversity and heterogeneity as well as the use of scientific knowledge in the inventive process requires qualitative research such as case studies.

CHAPTER THREE: FROM BENCH TO BEDSIDE – DETERMINANTS OF BREAKTHROUGH DRUG DISCOVERY

Abstract

This study adopts a different approach in studying the determinants of breakthrough innovation by focusing on patents underlying successful drug products. Hence, 5381 Orange Book patents were matched with pharmaceutical patents to examine the profile of valuable inventions in the pharmaceutical sector. Results indicate that inventions with fewer applications (i.e. citations across different technology classes) have higher probability of entering the Orange Book. Broader legal scope boosts the probability of being an Orange Book patent. This might be because inventors carve out larger legal protection for more valuable inventions. In terms of knowledge recombination, higher technological knowledge diversity reduces the probability of being an Orange Book patent while technological knowledge distance (i.e. not coming from the patent's technology class) increases the probability of being an Orange Book patent. Regarding collaborative inventions, being assigned to multiple entities does not have a significant impact on the probability of being an Orange Book patent while being assigned to different organizational types reduces the probability of being an Orange Book patent. Subsample analyses reveal differences between drug substance patents and Orange Book patents as well as knowledge recombination before the year 1995.

Keywords: Pharmaceutical Innovation; Patents; Orange Book; Breakthrough Innovations

3.1 Introduction

The pharmaceutical sector is a unique industry with a profound impact on people's health and quality of life. It is substantially more coupled with science, and more regulated than other industries. With an annual growth rate of 4-7%, it is fast approaching the US$1 trillion market size. At the same time, it faces serious challenges in the fundamentals of the industry. Innovation and marketing are the complementary pillars of the industry. Subpar innovation will lead to reduced sales and eventually less resources for innovation (Ding et al., 2014a). Since firms play a vital role in innovation and production of drugs, the broader societal outcomes will be higher prices and fewer therapies for new health challenges such as drug resistant bacteria.

A combination of factors makes innovation in the pharmaceutical sector more challenging and volatile: finite patent protection, long drug development cycles (4–16 years), high failure rates, soaring costs of developing and launching drugs, and huge post-launch market risks (e.g. withdrawal on safety concerns e.g. Vioxx®[26]). While at first sight these issues may deter R&D investments, the pharmaceutical sector leads all other sectors in R&D spending. Consequently, R&D portfolio management is complicated; for instance, in 2010, R&D expenditure of GlaxoSmithKline (GSK) stood at over 6 billion USD in a total of 147 R&D projects across 13 therapeutic areas in different stages of

[26] Rofecoxib (branded as Vioxx, Ceoxx, and Ceeoxx) was an NSAID marketed by Merck and one of the most widely used drugs ever to be withdrawn from the market. It received FDA approval on May 20, 1999 and was withdrawn on September 30, 2004 after it transpired that it increased heart attack and stroke risk in long-term use (Wikipedia).

development (Ding et al., 2014b). Knowledge of high potential innovations' characteristics could help in R&D management processes.

Revolutionary discoveries in diverse disciplines such as life sciences, engineering, informatics, and optimization are utilized to identify naturally occurring compounds, design new ones, or alter the existing ones (Petrova, 2014). In fact, the history of the pharmaceutical industry can be viewed as an evolutionary adaptation to major endogenous and exogenous technological and institutional "*shocks*" (McKelvey et al., 2004). However, in the face of all scientific and technological advances, the annual FDA-approved drug output of the industry has been almost flat for decades (Scannell et al., 2012). Hence, studies characterizing successful innovations can potentially help adjust the innovative efforts for better performance.

In the spirit of Kelley et al. (2013), this research is based on the assumption that an examination of high-potential inventions can improve our understanding about the underlying features of innovations that revolutionize industries and assist in attaining societal goals. Having detailed knowledge of successful innovations would be useful in handling the pressure for continuous innovation.

Given the knowledge-intensity of innovation in pharmaceuticals, exploring the knowledge content of innovations is an important foray of the study. The literature is rife with studies defining valuable, successful, or breakthrough innovation based on forward citation of patents (e.g. top 1% cited patents in own technology field and filing cohort) (e.g. Phene et al., 2006; Squicciarini et al., 2013). This study uses a different approach by

focusing on patents underpinning successful drug products, in other word, inventions with actual applications rather than citations received.

While a different approach, a parallel can also be drawn between this approach and citations-based breakthrough definitions. The main pharmaceutical patent dataset of this study contains 120,594 patents while there are 5,381 total Orange Book patents in the dataset (about 4.4%). However, not all Orange Book patents are from the main pharmaceutical patent class (IPC A61K); focusing on patents assigned to the pharmaceutical patent class (3,870) Orange Book patent occurrence is even rarer at about 3.2%. Hence, the Orange Book patents represent rare, successful inventions worthy of breakthrough label.

Another justification is the existence of studies using specific sets of inventions as breakthroughs. For instance, Fontana et al. (2012) use R&D 100 awards as their source of breakthrough inventions. R&D 100 awards are organized by the Research & Development magazine. Since 1963, every year 100 significant new products available for sale or licensing in the preceding year are introduced and awarded (Fontana et al., 2012).

The FDA's Orange Book or the "Approved Drug Products with Therapeutic Equivalence Evaluations" publication lists the drug products approved on the basis of safety and effectiveness.[27] New drug applications (NDAs) are required to include relevant patent information which is published in the Orange Book after approval.[28] These patents

[27] http://www.fda.gov/Drugs/InformationOnDrugs/ucm129662.htm (Accessed April 11, 2016)
[28] http://www.fda.gov/Drugs/DevelopmentApprovalProcess/ucm079031.htm (Accessed April 11, 2016)

are directly related to approved drugs; hence, will serve as an important benchmark against which to assess other comparable innovations.

This essay attempts to ascertain the differences between successful innovations (i.e. patents listed in the Orange Book) and pharmaceutical inventions not used in drug products. While, strictly speaking, patents represent inventions, the fact that Orange Book patents are directly related to drug products, makes them a very good proxy for innovations.

3.2 Literature Review

A multitude of terms have been used to characterize important technological innovations: radical, revolutionary, breakthrough, discontinuous and disruptive are among the most drawn upon terms. Breakthrough or radical innovations launch new practices that are noticeably better than the prevailing ones in some important aspects. These innovations may have the potential to create new markets, seed follow-up innovations, and create competitive advantage for firms or nations (Arts et al., 2013). Most extant studies have defined breakthrough innovation based on patent forward citations. This brief literature review tackles the issues of the theoretical lens of studies, scope of inventions, and the properties and sources of knowledge.

3.2.1 Theoretical Lens

In terms of theoretical lens, most studies dealing with breakthrough innovations have adopted an approach called "*innovation as a search process*". Innovation as organizational search and learning emanates from attempts to open the black box of the innovation production function. Several types of search mechanisms have been

envisioned such as recombinatory, cognitive and experiential search (Ahuja et al., 2008). The recombinatory search has been exhorted as an ultimate source of novelty by many scholars (Fleming, 2001). A prominent early reference to recombination can be traced to Schumpeter (1939). In defining innovation, he resorts to the production function notion. He states that, in the economic sense production is "*nothing but combining productive services*". He extends this notion to innovation by mentioning that "*innovation combines factors in a new way, or that it consists in carrying out new combinations* …" (p. 27).

Within the recombinant search approach to innovation, a number of recurring themes are prominent. From one perspective the literature can be divided into two strands: those focused on the impact of organizational factors on the chances of developing breakthrough innovation, and those exploring issues pertaining to the properties of knowledge such as the breadth and sources of knowledge recombined in the innovation process (adapted from Kelley et al, 2013; Verhoeven et al., 2016).

The studies on organizational factor highlight the importance of controlling for organizational effects in innovation studies. For instance Baba and Walsh (2010) highlight how Sankyo Pharmaceuticals, despite temporal lead, abandoned the development of statins in response to an adverse observation. This left room for Merck to take the lead. They put down this decision on organizational differences in risk assessment.

3.2.2 Properties of Knowledge

The diversity and distance of knowledge have been tackled under the knowledge recombination theoretical lens. While the literature has operationalized these concepts

very differently, a simple generalization would be that knowledge diversity pertains to the number of different components of knowledge assembled to create the invention. Knowledge distance refers to the degree of differentness of the knowledge components used in the inventive effort. Differentness can be with reference to the focal invention (the focus in this study) or the organizational repertoire.

The distance and diversity of knowledge are related to two competing theories of creativity. According to the "*tension view*", deep knowledge can lead to myopia to the extent that the recombination of distant or diverse knowledge is needed for generating new ideas. On the other hand, according to the "*foundational*" view of creativity, developing a deep understanding of the foundations, assumptions and weaknesses of a particular knowledge domain helps detect anomalies that can lead to breakthrough innovations. Too distant and diverse recombination can be disadvantageous to innovation because only a penetrating attempt can produce breakthroughs (Kaplan and Vakili, 2015). Weisberg (1999) brings up the analogy of "*knowing the territory*" and states that you need to know the discipline to be able to go beyond it.

As an example of research with theoretical implications, Dunlap-Hinkler et al. (2010) explore the impact of firms' prior experience with innovation on the likelihood of success in breakthrough innovation. Examining 1,496 FDA new drug approvals (1993 - 2002), they conclude that success in nongeneric incremental innovation does not have a significant influence on breakthrough chances; previous experience in generics significantly reduced the likelihood of having a breakthrough innovation[29]. They reason

[29] New molecular entities (NMEs)

that a solid foundation in generic incremental innovation thwarts breakthrough performance.

In terms of knowledge diversity, the most common proxy of knowledge diversity is based on patent backward citations that measure the diversity of the technology fields (i.e. patent classes) on which a patent rests. Kaplan and Vakili (2015) identify patents originating new topics in a specific nanotechnology area by developing a text-based measure of novelty (i.e. cognitive novelty). In contrast to the received theory of recombination, they observe that patents seeding new topics are more likely to be associated with local search (i.e. a narrower domain with less knowledge diversity), while economic value (i.e. measured by patent forward citations) requires broader recombination. As a practical lesson, they mention that transforming novel ideas into economically useful ones requires bridging wide recombination and local search.

According to Ardito et al. (2016) for a patent being an established technology (i.e. proxied by patent's backward citations count) has an inverted U-shaped effect on the likelihood of becoming a breakthrough technology (i.e. based on forward patent citations). Knowledge diversity negatively influenced the relationship.

In terms of knowledge distance, Egli et al. (2015) attempt to uncover the patent indicators explaining the diffusion of climate change mitigation technologies. They use patent grants, applications, patents owned by commercial entities, and risk finance as dependent variables (figures form 2000-2010) and construct independent variables based on figures from 1990-2000. Moreover, technologies with more knowledge distance in the

1990's had slower diffusion in the 2000's and were less cited by private (i.e. commercial firms) patent applicants. However, they were more popular with risk finance in the 2000s.

3.2.3 Scope of the Invention

The scope of an invention can be explored from a number of perspectives. The invention application[30], legal, and technological scopes are introduced here.

The application scope is defined here as the utilization of the invention in other inventions. This can be proxied by the number of different technology classes in which the patent is cited (i.e. diverse citations based on citing patents' classes). This is known in the literature as the generality of the patent. For instance, Youtie et al. (2008) use the term breakthrough innovation and general purpose technology (GPT) synonymously. General purpose technologies are pervasive in multiple sectors, are amenable to continual technological improvements, and stimulate complementary investment by adopting sectors (Schultz and Joutz, 2010). Youtie et al. (2008) case study pertains to nanotechnology with ICT comparisons. Using patent data, they observe that nanotechnology shows similar ''pervasiveness'' levels to that of ICT which is considered to be an existing general purpose technology (Youtie et al., 2008). Egli et al. (2015) observe that their measure of industrial generality of a whole patent class (e.g. biofuels) in environmental technologies is robustly correlated with subsequent diffusion of technology by having a positive impact on the dependent variables of patent grants, applications, patents owned by commercial entities, and risk finance.

[30] Application meaning use not the notion of patent application

The legal scope of a patent is defined in terms of its claims. A patent document represents a set of inventive components in the form of claims (OECD, 2009). The claims comprise the core section of a patent and define the legal scope of the invention for which protection is conferred (Archontopoulos et al., 2007). Barring some noise in this measure, such as the tendency to inflate the number of claims by some applicants and changes made during the examination process, they are used as a measure of legal scope (OECD, 2009). Empirical research has reported positive association between number of claims and breakthrough innovation (i.e. based on patent citations) (Kaplan and Vakili, 2015; Singh and Fleming, 2010).

Finally, the technological scope of the invention refers to the patent classes the patent has been assigned by the examiner. Technological scope proxies the different technological areas the patented invention is relevant for. It is also known as the "patent scope" in the literature and is often associated with the technological and economic value of the patent (Squicciarini et al., 2013). In a broad study Kelley et al. (2013) observe that breakthrough inventions (top 1% cited patents) tend to be based on greater technological scope (i.e. the patent scope or the patent classes the patent is assigned to). Lerner (1994) observes that technological scope affects the valuations of privately held venture-backed biotechnology firms with "a one standard deviation increase in average" scope being associated with "a 21% increase in the firm's value".

3.2.4 Physical Sources of Knowledge

Another strand of literature deals with the sources of knowledge components used in the inventive endeavor. There are studies emphasizing the importance of knowledge

from outside the firm's organizational and technological boundaries for innovation (Phene et al., 2003). Some studies have tackled the issues surrounding accessing knowledge from outside the innovative entity under the "*boundary spanning*" label. Without external knowledge integration, gains from internal technology development efforts might not be sustainable (Rosenkopf and Nerkar, 2001). Pharmaceutical firms are increasingly turning to external R&D in the course of restructuring and pipeline optimization. Licensing, acquiring, and alliances are three routes for this (Wuyts, 2014). Higgins and Rodriguez (2006) observe that acquisitions seem to effectively supplement a firm's internal R&D efforts and R&D-focused alliances.

Dunlap-Hinkler et al. (2010) observe that products developed as a result of joint ventures and alliances were more likely to be breakthroughs. Laursen and Salter (2006) report that "*searching widely and deeply*" is related to innovative performance in a curvilinear way (i.e. inverted U-shaped). Search breadth refers to the number of sources of knowledge or information used in the course of the innovation (e.g. suppliers, customers, universities, etc.). Search depth is defined as the intensity of using "*different search channels or sources of innovative ideas*".

To conclude this literature review a number of observations is in order. In terms of the dependent variable used in the studies, citations-based measures (e.g. top 1% cited patents) are common but other variables such as actual FDA new drug approvals, patent applications, grants and risk finance, are occasionally used. Regarding independent variables, two broad themes can be detected: those mainly exploring the impact of organizational-level issues and processes (e.g. alliances) on technology breakthroughs

and those concerned with knowledge structures underpinning breakthrough innovations (e.g. knowledge diversity) (adapted from Kelley et al, 2013; Verhoeven et al., 2016).

3.3 Research Question and Hypotheses

This study seeks to ascertain "*how patents underlying successful drug products differ from other patents?*" Given rising R&D costs and stagnant new drug approvals, any improved knowledge of the inventions that have actually been incorporated into a drug product can have important managerial implications. The unit of analysis is the invention (proxied by the patent). The hypotheses have been grouped into the three categories of application breadth (or generality) of the invention, legal and technological scope of the invention, and knowledge recombination profile.

A preliminary note is necessary before formulating the hypotheses. As mentioned earlier, a patent document contains references to relevant information (or prior art). Patents are issued by an authorized agency granting time-bound exclusive rights to inventions. The granting process includes examination for novelty, inventive activity and industrial applicability. Moreover, the references cited by the inventors can be altered by the examiner. A patent can be cited on other patents as prior art. This discussion highlights three important actors whose actions impact inventive output and quality. The inventors[31] are the formulators of inventions, examiners as gatekeepers and quality controllers, and other inventors citing the patents in their patents are the ultimate arbiters of quality, application and utility of the inventions.

[31] A catch-all for all involved on their side including attorneys drafting the patent text

3.3.1 Application Breadth or Generality

The first hypothesis draws on the notion of general purpose technologies (GPTs). Widespread adoption of a core technology is the upshot of similar decisions on a variety of actors that see a potential in the technology (Youtie et al., 2008). GPTs can be "*process technologies, product-related or organizational transforming technologies*" (Liebenau, 2007).

Feldman and Yoon (2012) examine a number of characteristics of GPTs regarding the Cohen– Boyer patented invention for recombinant DNA (US 4237224 *Process* for Producing Biologically Functional Chimeras). They observe that Cohen–Boyer related patents exhibit characteristics of a GPT: i.e. technological complementarity and "*wide scope of applicability"*.

Only "composition" or "method of use" patents are allowed for listing in the Orange Book. These patents cover "the drug compound, specific formulations of the drug, or methods of treating certain diseases by administering the drug." "*Process patents*" covering "*methods for making*" the drug compound, are not allowed (Barkoff, 2006). Per patent listing regulation:

"*… such patents consist of drug substance (ingredient) patents, drug product (formulation and composition) patents, and method of use patents. Process patents are not covered by this section and information on process patents may not be submitted to FDA.*"[32]

Combining these arguments, we can expect that inventions listed in the FDA's Orange Book or the "Approved ***Drug Products*** with Therapeutic Equivalence

[32] https://www.fda.gov/OHRMS/DOCKETS/98fr/PATENT.pdf

Evaluations", cover more specific inventions hence are less likely to exhibit general purpose technology characteristics especially a wide scope of applicability across technology fields. Hence:

Hypothesis 1: The broader the applications of an invention across technology fields, *ceteris paribus*, the lower the probability of being an Orange Book patent.

A tentative parallel can be drawn between generality and knowledge distance and diversity. Knowledge distance and diversity measure the different technological components used in the invention. This is done by the inventors and examiners. Generality refers to the different technological fields in which the invention has been used (i.e. cited by other patents). Hence, generality materializes in terms of the decisions of other inventors[33].

The pervasiveness (scope of applicability) and continual technological improvements of GPTs can be tracked by patent generality index and forward citations respectively (Schultz and Joutz, 2010).

Squicciarini et al. (2013) construct the generality index based on the first 5-year citation window (since publication of the patents) as follows. It is based on the notion of the Hirschman-Herfindahl Index (HHI) and relies on information concerning the number and distribution of forward citations and the technology classes (IPC) of the citing patents. It is calculated as follows and is defined between zero and one (reproduced form Squicciarini et al., 2013):

[33] Baring examiner citations

Let *X* be the focal patent with y_i patents citing the focal patent, with *i* = 1, …, *N*.

$$\beta_{ji} = \frac{T_{ji}^n}{T_i^n}$$

T_i^n: total number of IPC n-digit classes in y_i

T_{ji}^n: total number of IPC n-digit classes in the j^{th} IPC4 digit class in y_i

j= 1… M_i is the cardinal of all IPC4 digit class in y_i

Generality index is defined as:

$$G_X = 1 - \sum_{j=1}^{M_i} (\frac{1}{N} \sum_{i=1}^{N} \beta_{ji})^2$$

With $\beta_{ji} = \frac{T_{ji}^n}{T_i^n}$ it can be rewritten as:

Equation 5 The Generality Index

$$G_X = 1 - \sum_{j=1}^{M_i} (\frac{1}{N} \sum_{i=1}^{N} \frac{T_{ji}^n}{T_i^n})^2$$

The denominator is equal to $T_i^n * N$

As such the generality of a patent reflects the decisions of other inventors in terms of the application and utility of the invention.

3.3.2 Legal and Technological Scope of the Invention

Two hypotheses are formulated based on the technological and legal scopes of the patent.

We would hypothesize that a broader patent scope is associated with breakthrough innovation designation if our dependent variable was based on forward citation count. However, given the fact that our breakthrough definition is based on

application in a drug product, it is expected to see a more focused legal and technological scope.

A few arguments will be used to illustrate the possible relationship. To begin with, the number of claims is "*associated with the technology or product "space" being protected*" (Lanjouw and Schankerman, 2001). Moreover, "*a product patent claims actual physical objects such as machines or molecules*", while a process patent is about methods or steps for accomplishing a task. "*Process patents normally require a description at a higher level of abstraction*" (Surden, 2011). Therefore, most probably, a drug product can be claimed in fewer statements or claims and would protect a more defined technological space.

A patent examiner may ask the patent applicant to restrict the invention scope (i.e. because "*two or more independent and distinct inventions are claimed in a single application*".[34] If this is required for an Orange Book patent, the patentee will have to choose between "*a drug; a key intermediate for making the drug; methods of making the drug; a drug metabolite; and methods of treating patients using the drug*". Claims pertaining drug patents are supposed to be more valuable; hence, in the event of a restriction requirement, they are often selected over other claims (Andres, 2015). This process will further restrict the number of claims per patent by breaking up broad patents. Hence:

Hypothesis 2: The broader the legal protection of a patent, *ceteris paribus*, the lower the probability of being an Orange Book patent.

[34] http://piersonpatentlaw.com/what-is-a-patent-restriction-requirement-under-37-cfr-1-142/

Turning to the technological scope, Lerner (1994) developed a proxy based on the International Patent Classification (IPC) scheme. He counts the number of first four digits of the IPC classes the patent is assigned to as an indicator of technological scope.

The relationship between technological scope and Orange Book listing is probability similar to the generality index but from the examiner point of view. In other words, the examiner decides on the "*envisioned applicability*" of the invention and assigns the patent to the patent classes. The larger the number of patent classes the patent is assigned to, the broader the scope of applications. According to Lerner (1994) examiners have "a strong incentive" for careful classification of the patent because eventually they return to these classifications for prior art search.

Hypothesis 3: The broader the "*envisioned applicability*" of an invention, *ceteris paribus*, the lower the probability of being an Orange Book patent.

3.3.3 Knowledge Recombination Profile

Knowledge recombination is examined in three dimensions in this study: diversity, distance (or heterogeneity) and boundary spanning.

Knowledge diversity and knowledge distance represent various degrees of differentness of knowledge used in an invention. While knowledge diversity refers to the "breadth" of the technology fields on which a patent relies, knowledge distance represents drawing on knowledge components from fields other than its own (adapted from Egli et al., 2015). Hence, the latter scores higher on the "differentness" scale.

The "*foundational*" view of creativity and the concept of "*better than the Beatles*" are used to formulate the hypotheses on knowledge diversity and distance. The "*better*

than the Beatles" problem refers to the practical situation where many diseases have been satisfactorily tackled, hence a large stock of approved drugs exists and new drugs have only a modest incremental benefit over what is already available (Scannell et al., 2012; Ding et al., 2014a). This is an indication of tightening innovation opportunities; hence, a more in-depth and focused mastery of knowledge is needed to improve chances of breakthrough innovation. These dynamics is also consistent with the "*foundational*" view of creativity calling for deep knowledge of the discipline for innovation (Weisberg, 1999). Evoking the recombinant view, while innovation still requires new knowledge building blocks, their degree of differentness should be such that the inventors can assimilate them and recombine into a differentiated invention. This is a reasonable expectation *if* the new knowledge comes from the vicinity of the knowledge[35] previously experimented with. Hence:

Hypothesis 4: The higher the technological knowledge diversity of a patent, *ceteris paribus*, the higher the probability of being an Orange Book patent.

As an extension of the aforementioned argument, given the fact that technologically distant knowledge comes from fields other than the patent's, its degree of differentness may preclude a seamless recombination; hence:

Hypothesis 5: The higher the technological knowledge distance of a patent, *ceteris paribus*, the lower the probability of being an Orange Book patent.

Consistent with extant literature, the originality and radicalness indices are used to proxy knowledge diversity and distance, respectively. The patent originality index refers

[35] This hypothesis is based on the originality index; hence, in comparison with the radicalness index, includes patent classes from the same class as the focal patent.

to the breadth of the technology fields on which a patent relies and is constructed based on backward citations as follows (Squicciarini et al., 2013):

Equation 6 The Originality index

$$Originality_p = 1 - \sum_j^{n_p} {S_{pj}}^2$$

S_{pj} is the percentage of citations made by patent p to class j out of the n_p IPC four-digit (or seven-digit) patent classes in the patents cited in the patent p.

The patent radicalness index is used to proxy for the technological knowledge distance of a patent's knowledge base. Radicalness (à la Shane) measures the degree to which a patent builds on (through backward citations) a diversified array of technologies; i.e. patents from fields other than own assigned field (Johnstone et al., 2015). It is specified as follows:

Equation 7 The Radicalness Index

$$\text{Radicalness}_P = \sum_j^{n_p} CTj/n_p;\ IPC_{pj} \neq IPC_p$$

CTj is the count of IPC four digit codes of patent j cited in patent p that is not assigned to patent p, out of n IPC classes in the backward citations counted at the most disaggregated level available (up to the 5th hierarchical level). The higher the ratio, the

more diversified the range of technologies underlying the patent (Squicciarini et al., 2013).

"*Boundary spanning*" activities is another dimension of knowledge recombination. The competitive dynamics of drug discovery is also important in this argument. Successful and continuous new drug introductions comprise the essence of sustainable competitive advantage for pharmaceutical firms. The shift from random screening to targeted rational drug design, "*the discovery process has become more systematic*", attracting entrants that compete in a race towards targeting a finite set of publicly known diseases. First to reach the market will enjoy a reputation effect, and without alternatives, market domination (Petrova, 2014). Some industry commentators have faulted the industry for developing too many "me-too" drugs[36] a term that can be traced back to the 1960's. However, the distinction between breakthrough and "me-too" drugs can be misleading as the majority of the latter have been in clinical development before the approval of the class breakthrough drug. Hence, the term "*development races*" better describes new drug development here than "*post hoc imitation*" (DiMasi and Paquette, 2004). A trifecta model of drug development is developing whereby public institutions create fundamental knowledge, biotech firms do applied research and large firms engage in clinical trials, large scale manufacturing, or commercialization. However, there is still rivalry between these organizations in terms of racing for patents (Petrova, 2014).

[36] Minor variations of the original drug using a similar action mechanism (Petrova, 2014)

No matter what mode of innovation (breakthrough or trifecta) is considered, there appears be an incentive to engage in collaborative patenting only in second best or generic technology areas. Hence:

Hypothesis 6: Joint assigned patents, *ceteris paribus*, will have lower probability of being an Orange Book patent.

Patent indicators are used extensively here hence a note on their validity is useful. A few studies have addressed the issue. Based on a sample of 214 influential inventions in biotechnology, Arts et al. (2013) examine various indicators to ascertain how they distinguish these inventions from less important inventions. They contend that multiple, complementary indicators offer a better picture. Moreover, they also observe that "*ex-post*" indictors (i.e. those based on forward citations that indicate impact) perform better than "*ex-ante*" indicators (i.e. those showing dissimilarity or novelty with respect to prior art) in identifying important inventions (67% correct classification versus 79% respectively). Likewise, Verhoeven et al. (2016) develop *ex-ante* measures of technological novelty based on patent classification and citations. Validating the measure against R&D 100 awards and EPO's[37] refused patent applications, they state that inventions detected as novel by the measures are overrepresented in the former group and underrepresented in the latter group.

The study proposed here uses both *ex-post* (generality index and forward citations) and *ex-ante* indicators.

[37] European Patent Office (EPO)

3.4 Data

Patent indicators and the main sampling frame of the study come from the OECD Patent Quality Indicators database (March 2017 edition). The OECD Triadic Patent Families database (March 2017 edition) was used to extract patent family data. The COMETS[38] (Connecting Outcome Measures in Entrepreneurship, Technology, and Science) patent assignee data file was used to extract assignee and organization type data.

In addition to general patent data, the list of Orange Book patents was also needed. Per §314.53 of the "Code of Federal Regulation", an applicant who submits to FDA a New Drug Application (NDA) or an amendment or supplement to it, must submit patent information on "*…drug substance (active ingredient) patents, drug product (formulation and composition) patents, and method-of-use patents*." [39] Requests to remove a patent from the Orange Book can be from the "*NDA holder/patent owner, or from a third party*".[40] One important caveat here is that for patent delisting information was not available for this study. Consequently, we do not know if patents delisted for reasons other than natural patent term expiration were different from other Orange Book patents. However, if there is any such bias, accounting for it might be difficult as delisting dynamics might be at work continuously and impossible to control for. A

[38] "COMETS." Accessed April 11, 2017. http://www1.kauffman.org/COMETS/.
[39] "eCFR — Code of Federal Regulations." Accessed May 31, 2017. https://www.ecfr.gov/cgi-bin/text-idx?SID=95ebb262091c1069eef2566d32e40394&mc=true&node=se21.5.314_153&rgn=div8.
[40] "Orange Book Listings and Delistings". Accessed May 31, 2017. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwie-7-s65rUAhXB3SYKHQWQBmIQFggkMAA&url=http%3A%2F%2Freport.nat.gov.tw%2FReportFront%2Freport_download.jspx%3FsysId%3DC09602410%26fileNo%3D006&usg=AFQjCNH54a8fzVDXt5LxjS5lZUVXiKXAvg&sig2=DbLRpFe3eyOLTkPFL1-s4g

parallel can be drawn with studies based on FDA approved drugs when some approved drugs are withdrawn from the market because of catastrophic adverse effects.

To obtain Orange Book patents, a Freedom of Information (FOI) request was filed with the FDA on June 3, 2016.[41] As a result, patents from 1984 through 2014 Orange Book listings were obtained. In addition to these patents, the current Orange Book online patent file was also downloaded (April 3, 2017)[42] and added to the list.

Dropping duplicates yielded 5830 Orange Book patents. Of these 449 did not match the OECD Patent Quality Indicators database and were not used in the analysis. Likewise, reissue and design patents in the Orange Book were also dropped from the dataset because they were not in the OECD Patent Quality. A reissue patent is a patent issued in lieu of the remainder term of a patent "*deemed wholly or partly inoperative or invalid*"[43]. There were only 105 reissue patents. Design patents (six in the data) cover only aesthetic aspects of a product and are almost never used in innovation studies. Given the negligible percentage of these patents in the sample, there is little concern their exclusion would bias results. The Orange Book online patent file has a drug substance flag, yielding 634 drug substance patents for the sample.

An important issue to explore is the technology field of the Orange Book patents. We would expect them to be assigned to the pharmaceuticals IPC class (A61K). Figure 7 depicts the technology fields of the Orange Book patents. A noticeable issue is that only

[41] It is not clear why FDA requires an FOI request for material that has been published previously. All the material could have been online like the current patent file.
[42] Accessed April 3, 2017. https://www.fda.gov/downloads/Drugs/InformationOnDrugs/UCM163762.zip
[43] Resources, MPEP. "MPEP." Accessed May 31, 2017. https://www.uspto.gov/web/offices/pac/mpep/s1401.html.

about 72% of patents are assigned to the "pharmaceuticals" technology field. More details of what is contained in the other fields can be found in appendix B.

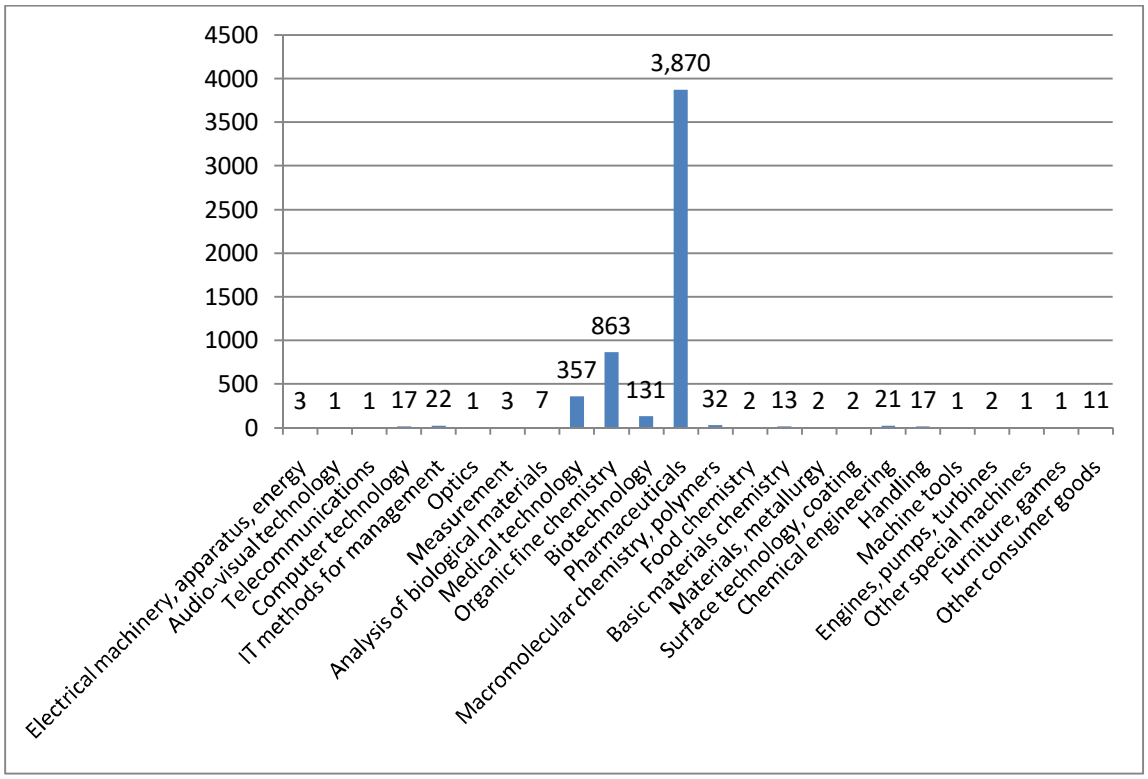

Figure 7 Orange Book Patents by Technology Field (Total= 5381)

Figure 8 depicts the technology fields of drug substance patents. These are more reflective of the type of inventions we would expect to see. These patents, predictably, are from more tightly related fields. An interesting point is that only 56% are from the pharmaceutical patent field.

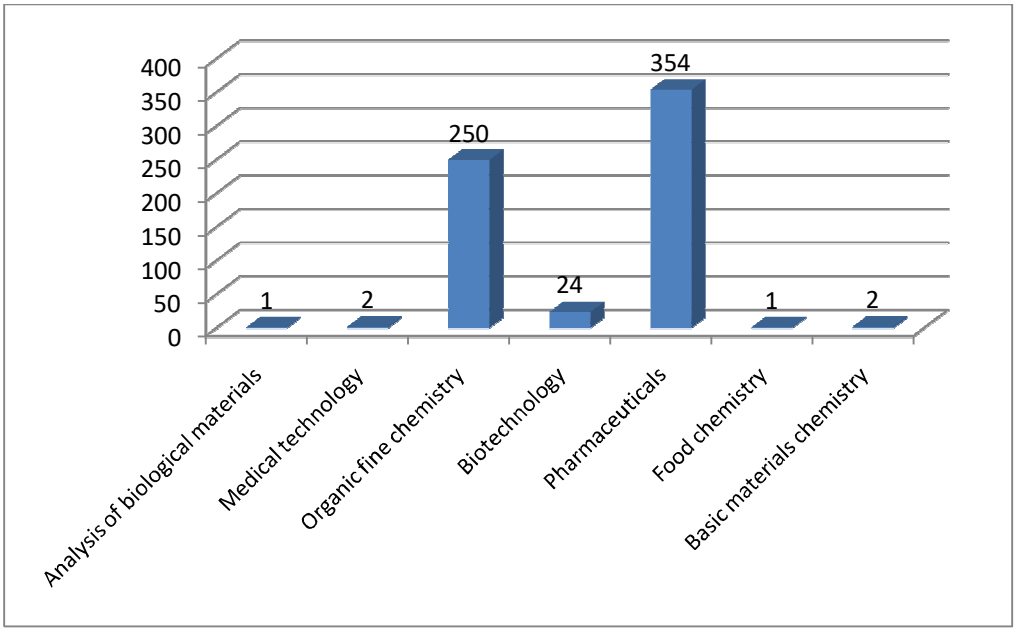

Figure 8 Drug Substance Patents by Technology Field (Total= 634)

Another issue to explore would be to see how these patents are related based on the patent family concept. Such an exploration can reveal cognate patents that reflect the same underlying invention or are closely related. The OECD Triadic Patent Families database (March 2017 edition) was used to extract family information. However, 955[44] patents were not listed in the database (i.e. were not Triadic). This observation is also important since Triadic patents are commonly assumed to be of higher value than non-Triadic patents. The remaining 4,426 patents hail from 2,281 patent families based on figure 9. Almost half (48.5%) of the listed patents have a family member in the Orange book. Another observation is that of the patents coming from same patent family, some are not in the Orange Book.

[44] 955 of the total 5381. In terms of drug substance patents only 35 were not listed.

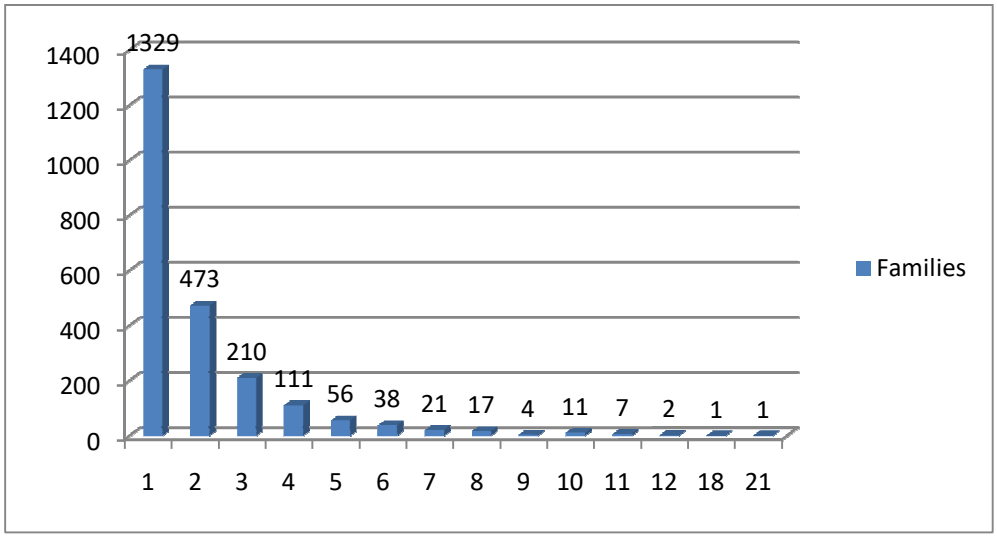

Figure 9 Orange Book Patents by Number of Family (4426 patents in 2281 Families)

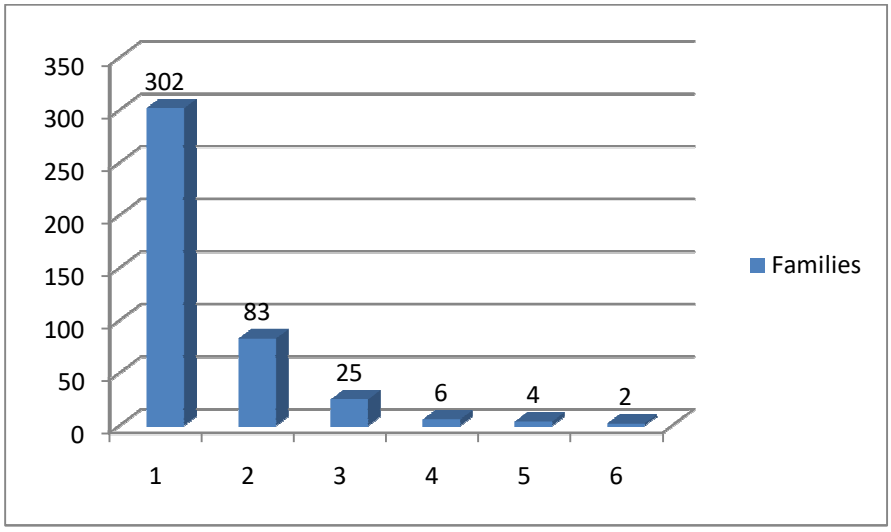

Figure 10 Drug Substance Patents by Number of Family (599 patents in 422 Families)

Figure 10 depicts the family status of drug substance patents. Of the 634 drug substance patents, 35 are not in the OECD Triadic database. The remaining 599 patents come from 422 patent families. One important observation from this exercise is that organizations list patents representing similar inventions. Even for the drug substance patents that are supposed to reflect more concrete inventions, 29.5% of the patents have a family member in the drug substance patent sample. Tentative explanations would be listing patents with different expiration dates or trying to show a larger number of patents in the Orange Book[45]. The family-based breakdown of the sample is important because it indicates that a number of patents either reflect the same invention or closely related inventions.

Based on the COMETS dataset[46], the 3339 Orange Book patents with organization ids are assigned to 778 entities. 66 entities supply at least 10 patents to the Orange Book and 372 entities contribute only one patent. While the majority of the patent assignees are firms, there are other entities such as hospitals and universities among the assignees. Figure 11 depicts the organization type of Orange Book assignees for the 3338 patents with available organization type information. The 371 drug substance patents with organization ids are assigned to 165 entities of which 91 entities have only one drug substance patent in the sample. An observation here is that the patent ownership is more dispersed than all orange book patents. Figure 12 depicts organization type of these

[45] The incentives for listing Orange Book patents have not been addressed in this study.
[46] Based on the unique organization id in the COMETS database. Admittedly, there might be some coding errors here, however, given the lack of a better database for cleaned assignee names, COMETS is used here.

entities. Figure 13 and 14 depicts the top 10 Orange Book and drug substance patenting entities, respectively.

After extracting the Orange Book patents and leaving out duplicates, the resulting patent list was merged with the OECD Patent Quality Indicators database (March 2017 edition). To construct the comparison sample, patents assigned to the IPC patent class A61K were kept (i.e. technology field 16 in the OECD Patent Quality Indicators database) (Squicciarini et al., 2013; Schmoch, 2008). This yields 120,594 patents for the comparison set of which only 81,715 are triadic. Depending on the type of regression run, the sample size varies based on missing values.

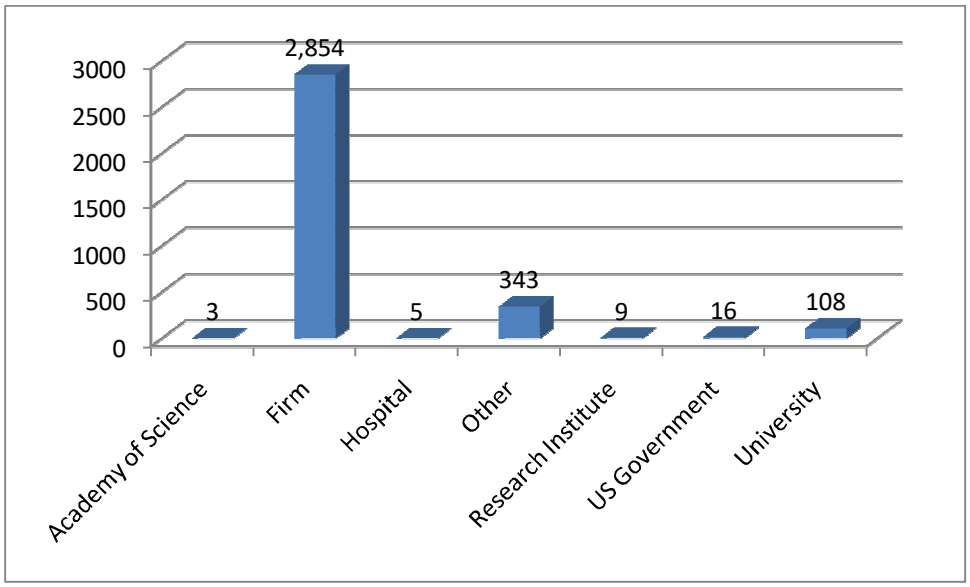

Figure 11 Orange Book Assignee Organization Type (3338 Patents)

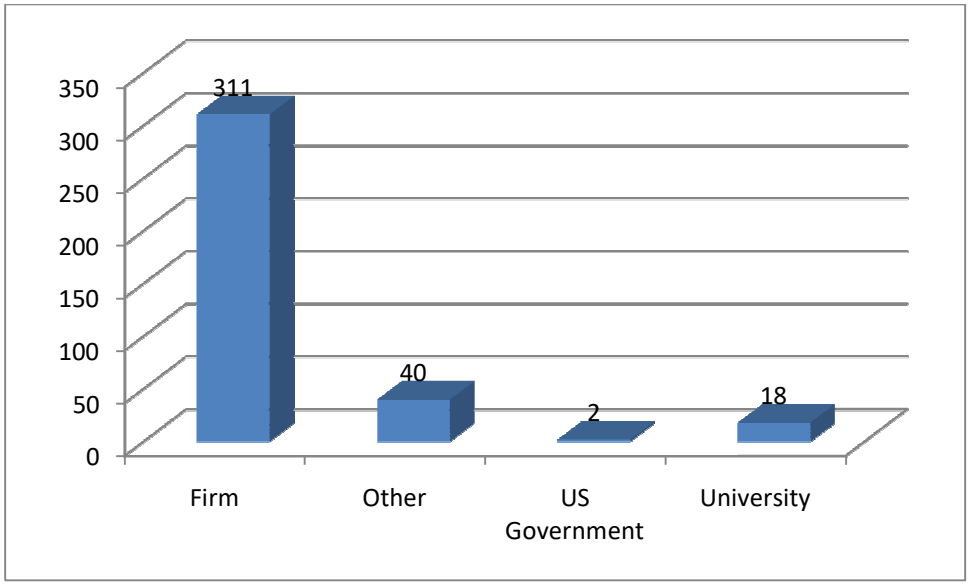

Figure 12 Drug Substance Assignee Organization Type (371 Patents)

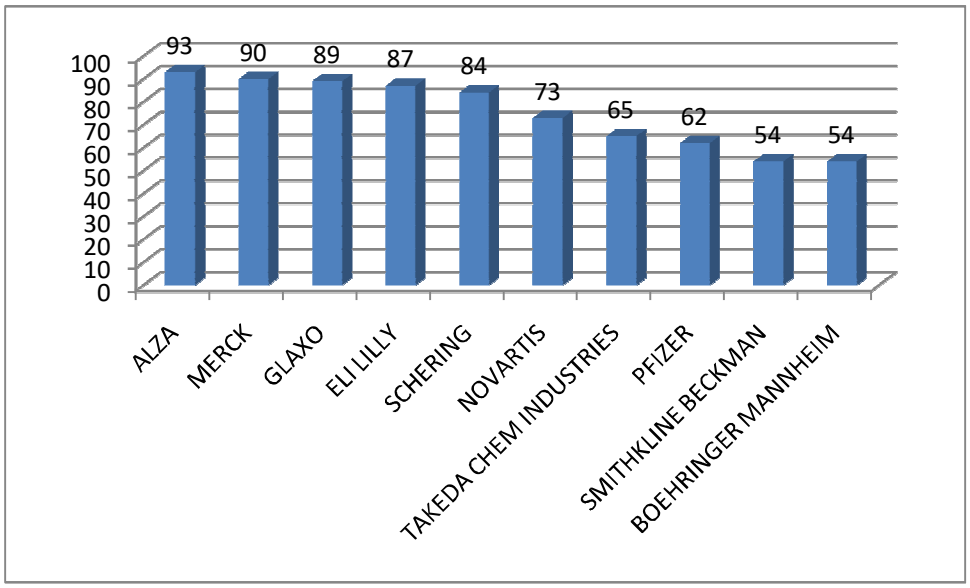

Figure 13 Top 10 Orange Book Patenting Entities (Based on COMETS dataset)

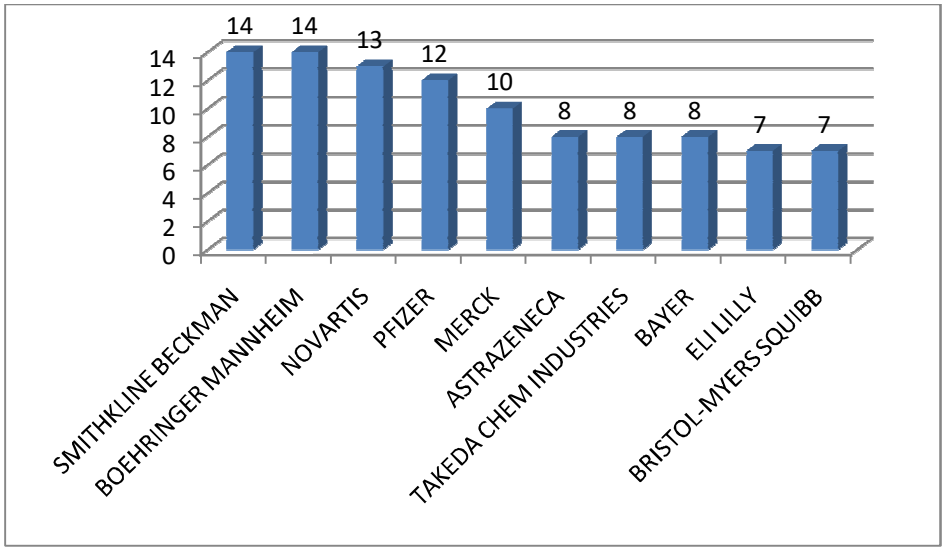

Figure 14 Top 10 Drug Substance Patent Assignees (Based on COMETS dataset)

3.5 Methods

Given the binary nature of the dependent variable (i.e. patent listed in the Orange Book or not), logistic or probit regression are the appropriate estimation methods. A comparison between the probit and logistic regression methods indicates that probit regression is preferred for the data.[47] In probit regression, the probability of a breakthrough innovation is defined as follows:

Equation 8 Probit Breakthrough Probability

$$\Pr(Y = 1|X) = \Phi(X^T\beta)$$

where Φ is the cumulative normal distribution

[47] "Difference of 84.150 in BIC' provides very strong support for current model" [i.e. probit]. This is for the main regression model on all sample.

The regression model can be summarized as follows:

Equation 9 Basic Regression Model

$$Y^{*} = \beta_0 + \beta_{1i} Inventors_i + \beta_{2i} Claims_i + \beta_{3i} Patent\_Scope_i + \beta_{4i} Family\_Size_i + \beta_{5i} Grant\_Lag_i + \beta_{6i} Bwd\_Cits_i + \beta_{7i} NPL\_cits_i + \beta_{8i} 7Year\_Fwd\_Cits_i + \beta_{9i} Originality_i + \beta_{10i} Radicalness_i + \beta_{10i} Generality_i + \beta_{11i} Renewal_i + Y_t + F_t + \epsilon$$

Table 8 depicts the variable definitions and the hypotheses they represent. Estimations will be done using the Stata statistical package version 12.0.

Table 8 Variables and Definitions

Variable	Definition	Reason for Inclusion
	Dependent variable	
Y	Binary dependent variable: $Y = 1$ if the patent is in the Orange Book, otherwise$Y = 0$	
	Explanatory Variables	
$Claims_i$	Number of claims of patent *i*	***Hypothesis 2***
$Patent\_Scope_i$	Cumulative number of distinct 4-digit IPC subclasses (Lerner, 1994)	***Hypothesis 3***
$Original_i$	The originality index of patent *i*	***Hypothesis 4***
$Radical_i$	The radicalness index of patent *i*	***Hypothesis 5***
$Generality_i$	The generality index of patent *i*	***Hypothesis 1***
multi_assignee	Dummy variable =1 if patent assigned to more than one entity	***Hypothesis 6***
diff_org_type	Dummy variable =1 if patent assigned to different organization types (e.g. Univ. and Firm)	***Hypothesis 6***
	Control Variables	
$Inventors_i$	Number of inventors	More inventors have more knowledge and more impact (Keijl et al., 2016)
$Renewal_i$	Count of years during which a granted patent has been kept active (Squicciarini et al., 2013)	Measure of value
$Grant\_Lag_i$	Normalized number of days elapsed between application and granting date (Squicciarini et al., 2013)	Controls for unobservable applicant behavior (Popp et al., 2004); applicants accelerate grant proceedings for their most valuable patents (Harhoff and Wagner, 2009)
Bwd_Cits_i	Patent references of the patent. Backward citations per patent is normalized based on the maximum value received by patents in the same year-and-technology cohort (Squicciarini et al., 2013)	Indicates technology intensity, i.e. knowledge form other patents (Callaert et al., 2006)
NPL_cits_i	Non-patent references. Number of NPL citations divided by cohort maximum (Squicciarini et al., 2013)	Measure of science intensity
$Family\_Size_{\mathrm{i}}$	The family size (i.e. the number of patent offices at which a given invention has been protected) of patent *i* based on OECD calculations. It is normalized by the maximum in cohort and Winsorized to correct for extreme values (Squicciarini et al., 2013)	Measure of economic value
$7Year\_Fwd\_Cits_i$	Forward citation of patent *i* 7 years since publications normalized by cohort maximum (Squicciarini et al., 2013)	Measure of value and follow-up technological improvement
Y_t	Year dummies	Vintage of technology
F_i	Entity dummies	Organizational effects

3.6 Results and Discussion

Table 9 shows summary statistics for the variables. The negative values of the grant lag coefficients are because of normalization (see table 8). Table 10 depicts the pairwise correlation of the variables. After patent and non-patent literature correlation (0.5874), the second biggest positive correlation is for generality and patent scope (0.319). This is in line with the argument that they represent the applicability of the invention from the perspective of other inventors (generality) and the examiner (scope), respectively. Another notable observation is the largest negative correlation between radicalness and patent scope (-0.3665). Which is again intuitive: the more radical the invention, the more limited the envisioned applications by the examiner.

Table 11 depicts the regression results. Five models were run. Model 1 is based on all sample with year dummies, the second model is run on a subsample of patents by entities in the Orange Book with "10 and more patents in the Orange Book". This subsample is meant to control for entity[48] fixed effects. The third model is the same as the second except for using clustered errors for robustness check. The forth model is run on a subsample controlling for the first U.S. patent of the patent family. This is meant to control for possible duplicate patents in a patent family. The last regression is clustered errors model for robustness check of the fourth model. Year dummies are included across the board to control for the vintage of the technology.

[48] "Organization" or "entity" are used interchangeably in this essay. As mentioned earlier, the patents are not just assigned to firms.

There is broad support for the first hypothesis. Broader actual applications of an invention (i.e. generality index), lowers the probability of being an Orange Book patent. The inverse relationship becomes stronger with entity and patent family controls.

The second hypothesis is not supported. In fact, some models indicate the opposite relationship, i.e. in model 5 one more claim boosts the probability of being an Orange Book patent by the miniscule amount of 0.0684% at the 5% level of significance. Hence, either number of claims is not related to Orange Book listing probability or has a negligible effect. A logical explanation is that the number and breadth of claims is determined by the patent drafter's approach and comparable inventions can be drafted in fewer, broad claims or many more, narrower claims[49]. Some studies have used proxies for the length of claims (e.g. number of words or statements); however, such measures were not available for this study.

There is support for the third hypothesis in three models. Model 4 controls for organization and patent family (i.e. possible duplicate inventions) effects. It indicates one IPC class more assigned to the patent reduces the probability of being an Orange Book patent by 1.86% (at the 1% level of significance).

Hypothesis four is not supported in fact the opposite is observed in all models. Higher technological knowledge diversity (i.e. originality index) reduces the probability of being an Orange Book patent. Likewise, hypothesis five is not supported and the opposite is observable with higher technological knowledge distance of a patent boosting

[49] Effective starting late 2007, rule 5/25 limits the number of total claims to 25 and independent claims to 5. This effect is not specifically controlled for in the hypothesis testing but time dummies are included present in models as delineated earlier. Mean claims of all Orange Book patents with filing dates of 2007 and earlier is 20.7 (SD 21.26) and the same number for filing dates of after 2007 is 19.1 (SD 19.1).

the probability of being an Orange book patent. The relationship is consistent in all models. These two hypotheses were formulated based on the foundational view of creativity; however, the observations are more in line with the tension view.

To test hypothesis six, two dummy variables were constructed representing patents assigned to more than one entity ("multi_assignee") and patents assigned to different entities ("diff_org_type" e.g. a firm and a university). The results are depicted in table 12. The dummy variable representing jointly assigned (i.e. "multi_assignee") ownership is not significant. The dummy variable representing the patent assigned to different organizational types (i.e. "diff_org_type") is significant at the 10% level indicating being assigned to different organizational types reduces the probability of being an Orange Book patent by 1.05%. This is in line with the hypothesis. Co-assigned patents come from small-scale collaborative R&D in which dividing the intellectual property is not possible. Moreover, they have certain disadvantages such as slowing the market entry decision, need for complex contractual provisions to cover contingencies, and need for all assignees to be on infringement suits in the U.S. legal system (Kim et al., 2016). These issues lend credence to the regression results, since assignees would like to have sole ownership of important patents if the aforementioned issues are of significance to them. Among drug substance patents, there are only 14 patents assigned to more than one entity and only 2 with different organizational types. Assuming drug substance patents are more important than other Orange Book patents, this is another evidence corroborating the argument.

A note on control variables is also useful. More inventors results in a slight reduction on probability of being an Orange Book patent in only model one. Theoretically, the reason can be that Orange Book patents are the result of smaller and more focused R&D projects or the inventions are split into a number of patents. Having a larger family size (i.e. wider geographical scope) improves the probability of being an Orange Book patent. The relationship exists across all models. Grant lag coefficient is significant in two models indicating Orange Book patents are granted faster if we do not control for organization effects. The coefficient is negative because due to the index construction a higher index reading is associated with less lag.

More patent citations improve the probability of being an Orange Book patent in four models. Non-patent literature is significant in only one model (which does not control for organization and patent family effects) with a negative coefficient. It might be that commercially valuable inventions are first patented; hence, not much directly related scientific prior art would be there to underpin them.

Seven-year forward patent citations boost the probability of being an Orange Book patent in four models. As mentioned earlier, this is in line with the literature using forward citations as a measure of value. In terms of renewal, intuitively, keeping the patents in effect (renewed) boosts the probability of being an Orange Book patent (significant in all models).

A final point is that the size of the coefficients is really small; hence, the impact of the issues discussed here are small. The biggest impacts pertain to “generality” and

"originality" in the models controlling for both organization and family effects (models 4 and 5).

A caveat on the construction of the subsample of Organizations with ten or more patents in the Orange Book is that it may represent a specific set of organizations. However, this was done to enable organization fixed effect controls. Given the large number of entities in the main sample, fixed effects on the main sample could lead to biased estimates.

Table 9 Summary Statistics for All Patents

Variable	Obs	Mean	Std. Dev.	Min	Max
Inventors	83390	3.142	2.267	1.000	33
Claims	125935	15.832	14.336	1.000	397
Patent scope	125975	2.761	1.390	1.000	15
Family size	125975	10.132	8.438	1.000	56
Grant lag	125975	1122.712	672.110	-252.000	8117
Bwd cits	125975	19.890	27.686	0.000	268
NPL cits	125975	20.025	27.906	0.000	196
FWD cits 7yr	125975	12.942	25.390	0.000	1090
Generality	102469	0.551	0.220	0.000	0.958
Originality	120091	0.820	0.163	0.000	0.988
Radicalness	120202	0.289	0.239	0.000	1.000
Renewal	125975	8.818	4.414	0.000	29
Filing Date	125975		8.944	1976	2016

* Showing only three decimal points

Table 10 Pairwise Correlation Matrix

	Inventors	Claims	Patent scope	Family size	Grant lag	Bwd cits	NPL cits	FWD cits 7yr	Generality	Originality	Radicalness	Renewal
Inventors	1.000											
Claims	0.051	1.000										
Patent scope	0.164	0.031	1.000									
Family size	0.203	0.093	0.148	1.000								
Grant lag	0.039	0.059	0.097	-0.044	1.000							
Bwd cits	0.097	0.198	0.055	0.170	0.165	1.000						
NPL cits	0.027	0.142	0.132	0.070	0.202	0.587	1.000					
FWD cits 7yr	0.095	0.157	0.084	0.262	0.005	0.301	0.202	1.000				
Generality	0.133	0.054	0.319	0.089	0.099	0.109	0.164	0.183	1.000			
Originality	0.114	0.093	0.193	0.088	0.147	0.279	0.246	0.108	0.231	1.000		
Radicalness	-0.137	0.001*	-0.367	-0.216	0.029	0.066	0.078	-0.054	0.003*	0.174	1.000	
Renewal	-0.047	0.053	-0.045	0.088	0.128	-0.178	-0.146	0.082	-0.004*	-0.058	0.011	1.000

All significant at the $p<0.01$ except for those marked *

Table 11 Probit Regression Results (Dependent Variable: Orange Book Patent=1; Average Marginal Effects)

	(1)	(2)	(3)	(4)	(5)
Variables	**Year Dummies**	**Year/Org. Dummies†**	**Year Dummies, Org. Clustered Errors (88)†**	**Year/Org. Dummies_ First Patent†**	**Year Dummies, Org. Clustered Errors (88)_ First Patent†**
Inventors	-0.000780**	-0.00167	0.000792	-0.00240	-0.00111
	(0.000352)	(0.00166)	(0.00289)	(0.00276)	(0.00457)
Claims (Hyp 2)	0.000338***	0.000350	0.000496**	0.000181	0.000684**
	(0.0000421)	(0.000218)	(0.000198)	(0.000364)	(0.000328)
Patent scope (Hyp 3)	-0.00807***	-0.00736*	-0.00827	-0.0186***	-0.0129
	(0.000750)	(0.00396)	(0.00610)	(0.00635)	(0.0106)
Family size	0.00377***	0.00736***	0.00633***	0.0116***	0.00935***
	(0.0000856)	(0.00136)	(0.000583)	(0.00136)	(0.000883)
Grant lag	-0.0000124	-0.0000163	-0.0000253**	-0.0000547	-0.0000656***
	(0.00000166)	(0.00000795)	(0.0000128)	(0.0000134)	(0.0000208)
Bwd cits	0.000510***	0.000411**	0.000485	0.000507*	0.00116*
	(0.0000376)	(0.000189)	(0.000308)	(0.000290)	(0.000603)
NPL cits	-0.000286***	-0.0000443	-0.0000501	0.000403	0.000584
	(0.0000398)	(0.000195)	(0.000342)	(0.000325)	(0.000638)
FWD cits 7yr	0.000234***	0.000239**	0.000241	0.00113***	0.00168***
	(0.000023)	(0.000113)	(0.000152)	(0.000250)	(0.000244)
Generality (Hyp 1)	-0.0456***	-0.0640***	-0.0930***	-0.151***	-0.206***
	(0.00390)	(0.0228)	(0.0189)	(0.0345)	(0.0271)
Originality (Hyp 4)	-0.0257***	-0.0697**	-0.0721*	-0.150***	-0.187***
	(0.00521)	(0.0296)	(0.0393)	(0.0468)	(0.0696)
Radicalness (Hyp 5)	0.00801**	0.0596***	0.0630**	0.0861***	0.137***
	(0.00395)	(0.0211)	(0.0262)	(0.0326)	(0.0429)
Renewal	0.00934***	0.0148***	0.0179***	0.0219***	0.0280***
	(0.000446)	(0.00335)	(0.00297)	(0.00415)	(0.00464)
Log likelihood	-10713.995	-1431.9638		-987.19745	
Log pseudolikelihood			-1641.5188		-1181.7711
LR chi2	4623.78	1053.15		966.35	
Wald chi2			1493.45		1021.20
Pseudo R2	0.1775	0.2689	0.1689	0.3286	0.2415
Sensitivity	2.63%	18.75%	5.92%	40.50%	29.68%
Specificity	99.82%	98.91%	99.18%	96.06%	96.74%
Correctly classified	95.05%	90.66%	89.83%	86.20%	84.28%
Linktest (_hatsq P>|z|)	0.000	0.251	0.116	0.021	0.162
Hosmer-Lemeshow (Prob > chi2)	0.2175	0.7355	0.8888	0.0541	0.3246
Observations	66,537	5,912	6,067	3,146	3,245

Average Marginal Effects; Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1; For sensitivity/specificity analysis: predicted Pr(D) >= .5;
† Only single assignee patents used; Entities with 10 and more patents in the Orange Book

Table 12 Jointly Assigned Patents (Dependent Variable: Orange Book Patent=1)

Variables	(1) Multi Assignee-Year Dummies	(2) Different Org.-Year Dummies
Inventors	-0.000736**	-0.000751**
	(0.000354)	(0.000353)
Claims (Hyp 2)	0.000338***	0.000338***
	(0.0000421)	(0.0000421)
Patent scope (Hyp 3)	-0.00806***	-0.00806***
	(0.000750)	(0.000750)
Family size	0.00377***	0.00377***
	(0.0000856)	(0.0000856)
Grant lag	-0.0000124***	-0.0000124***
	(0.00000166)	(0.00000166)
Bwd cits	0.000508***	0.000508***
	(0.0000376)	(0.0000376)
NPL cits	-0.000282***	-0.000280***
	(0.0000399)	(0.0000399)
FWD cits 7yr	0.000233***	0.000233***
	(0.000023)	(0.000023)
Generality (Hyp 1)	-0.0455***	-0.0455***
	(0.00390)	(0.00390)
Originality (Hyp 4)	-0.0257***	-0.0257***
	(0.00521)	(0.00521)
Radicalness (Hyp 5)	0.00808**	0.00811**
	(0.00395)	(0.00395)
Renewal	0.00935***	0.00936***
	(0.000446)	(0.000446)
multi_assignee (Hyp 6)	-0.00436	
	(0.00387)	
diff_org_type (Hyp 6)		-0.0105*
		(0.00546)
Log likelihood	-10713.351	-10712.06
LR chi2	4625.07	4627.65
Pseudo R2	0.1775	0.1776
Sensitivity	2.63%	2.63%
Specificity	99.82%	99.82%
Correctly classified	95.05%	95.05%
Linktest (_hatsq P>|z|)	0.000	0.000
Hosmer-Lemeshow (Prob > chi2)	0.3318	0.3815
Observations	66,537	66537

Average Marginal Effects; Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1; For sensitivity/specificity analysis: predicted Pr(D) >= .5

3.6.1 Drug Substance Patents

A subsample analysis was done on the drug substance patents flagged in the online Orange Book patent file. Table 13 depicts the models run on all sample and other Orange Book patents. These patents are more representative of the type of inventions that could potentially change the actual drug output of the industry. Moreover, strictly speaking, the hypotheses were formulated to characterize these inventions.

Model 1 indicate that, compared with other patents, broader legal protection (claims) boosts the probability of being a drug substance patent. Other variables of interest cease to have any significance. This might be because of the large pool of heterogeneous patents that the small number of drug substance patents (643) is compared against.

Focusing on the models run on Orange Book patents, four models indicate that hypothesis one does not hold. Broader application of an invention across technology fields (i.e. generality index) boosts the probability of being a drug substance patent. This might be related to the wide potential applicability of an active chemical substance.

Hypothesis two regarding the breadth of the legal protection of a patent still does not hold. There is evidence for the opposite, i.e. broader legal protection (i.e. number of claims) boosts the probability of being a drug substance patent in two models. These models do not control for family effects. If we take out duplicate family patents the effect dissipates.

Hypothesis three is no longer supported, with broader "*envisioned applicability*" of an invention boosting the probability of being a drug substance patent in models 3 and 4 controlling for family effects. This is consistent with the change in actual application

impact in this subsample analysis (hypotheses one). On the conceptual level, this might be evidence that the claim made here that the generality index and patent scope represent the same fundamental issue of applicability of the invention from the perspective of different parties (i.e. other inventors and the examiner).

Interestingly, hypothesis four is supported in models 3 and 5 (al the 5% level of significance with large coefficients) that do not control for organization effect. This means higher technological knowledge diversity (i.e. originality index) increases the probability of being a drug substance patent when we disregard organization fixed effects[50]. Likewise, model 2 and 3 support hypothesis five, in other words, higher technological knowledge distance reduces the probability of being a drug substance (at the 10% level and 1% level respectively) patent if we do not control for patent family effects. Partial support for these two hypotheses indicates the general logic of hypothesis formulation was sound. These observations yield tentative support for the foundational view.

[50] It might be organization fixed effects make estimates volatile, hence, the clustered models have significant coefficients (models 3 and 5).

Table 13 Probit Regression (Dependent Variable Drug Substance Patents; Average Marginal Effects)

	(1)	(2)	(3)	(4)	(5)
Variables	**Year Dummies**	**Year/Org. Dummies†-**	**Year Dummies, Firm Clustered Errors (600)†-**	**Year/Org. Dummies_ First Patent†**	**Year Dummies, Firm Clustered Errors (435)_ First Patent†**
Sample	**All**	**Orange Book**	**Orange Book**	**Orange Book-1st patent**	**Orange Book-1st patent**
Inventors	0.000602***	0.0315***	0.0225***	0.0360***	0.0349***
	(0.000112)	(0.00593)	(0.00344)	(0.00745)	(0.00531)
Claims (Hyp 2)	0.0000616***	0.000930*	0.000660**	0.00108	0.000771
	(0.0000151)	(0.000547)	(0.000331)	(0.000808)	(0.000590)
Patent scope (Hyp 3)	-0.000311	0.00704	0.00477	0.0368**	0.0254**
	(0.000319)	(0.0119)	(0.00740)	(0.0171)	(0.0111)
Family size	0.000646***	0.00174	0.00204***	0.00425**	0.00272**
	(0.0000408)	(0.00124)	(0.000729)	(0.00190)	(0.00110)
Grant lag	-0.000000521	0.0000265	0.0000138	-0.000082**	-0.0000614**
	(0.000000678)	(0.0000229)	(0.0000135)	(0.0000359)	(0.0000249)
Bwd cits	0.0000115	-0.00191***	-0.00154***	-0.00200**	-0.00176***
	(0.000015)	(0.000577)	(0.000330)	(0.000854)	(0.000528)
NPL cits	-0.0000208	-0.0000928	-0.0000117	0.000690	0.000716
	(0.0000155)	(0.000535)	(0.000376)	(0.000924)	(0.000606)
FWD cits 7yr	0.0000267***	-0.000366	0.0000683	0.00152**	0.00117***
	(0.00000831)	(0.000317)	(0.000209)	(0.000630)	(0.000375)
Generality (Hyp 1)	0.000554	0.439***	0.248***	0.478***	0.305***
	(0.00222)	(0.0911)	(0.0482)	(0.109)	(0.0703)
Originality (Hyp 4)	-0.00168	0.0722	0.206**	0.0696	0.302**
	(0.00323)	(0.115)	(0.103)	(0.141)	(0.128)
Radicalness (Hyp 5)	0.000297	-0.130*	-0.122***	-0.0693	-0.109
	(0.00197)	(0.0743)	(0.0467)	(0.104)	(0.0667)
Renewal	0.00179***	0.0191**	0.0148***	0.0376***	0.0342***
	(0.000234)	(0.00888)	(0.00479)	(0.0136)	(0.00866)
Log likelihood	-1911.7904	-448.11945		-234.59784	
Log pseudolikelihood			-778.198		-480.41456
LR chi2	855.31	379.40		338.74	
Wald chi2			345.12		386.33
Pseudo R2	0.1828	0.2974	0.2151	0.4193	0.3218
Sensitivity	0.00%	44.19%	19.03%	72.12%	60.51%
Specificity	100.00%	93.09%	98.04%	87.63%	90.73%
Correctly classified	99.29%	82.27%	86.18%	81.92%	81.48%
Linktest (_hatsq P>|z|)	0.052	0.937	0.356	0.821	0.086
Hosmer-Lemeshow (Prob > chi2)	0.2346	0.5914	0.1411	0.5575	0.8387
Observations	55139	1207	2344	614	1,150

Average Marginal Effects; Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1; For sensitivity/specificity analysis: predicted Pr(D) >= .5; † Only single assignee patents used

3.6.2 Pre-1995 Knowledge Recombination

Another subsample analysis was conducted to see how the impact of the applicability of the invention and knowledge recombination has changed over time with a 1995 cutoff point. This is motivated by the 1990s changes such as the emergence of biotechnology-based therapies. Also, for the first time, R&D productivity hit one drug per billion R&D spending (Scannell et al., 2012). To this end, a dummy variable was constructed to represent the filing years of 1995 and before that. Per table 14, this dummy variable was interacted with respective variables of interest and run on a subsample of first family patents to control for possible family effects.

Results indicate that before the year 1995, higher generality index boosts the probability of being an Orange Book patent (significant at the 1% level). Likewise, higher originality index before 1995 increases the probability of being an Orange Book patent (significant at the 1% level). However, higher radicalness index reduces the probability (significant at the 10% level).

This analysis indicates that the knowledge recombination profile of Orange Book patents as well as the diversity of their applications (proxied by the generality index) has changed with regard to the 1995 cutoff point. However, exploring the robustness of this observation or the reasons behind this is beyond the scope of the present study.

Table 14 Pre-1995 Knowledge Recombination (Probit: Orange Book Patent=1)

Variables	(1) **Pre-1995 Generality - Clustered Errors (5335) -First Patent**	(2) **Pre-1995 Generality - Clustered Errors (5335) - First Patent**	(3) **Pre-1995 Generality - Clustered Errors (5335) -First Patent**
Inventors	-0.000573	-0.000627	-0.000734
	(0.000593)	(0.000590)	(0.000593)
Claims (Hyp 2)	0.000209***	0.000210***	0.000216***
	(0.0000731)	(0.0000732)	(0.0000733)
Patent scope (Hyp 3)	-0.00814***	-0.00817***	-0.00825***
	(0.00136)	(0.00136)	(0.00136)
Family size	0.00428***	0.00429***	0.00429***
	(0.000293)	(0.000296)	(0.000295)
Grant lag	-0.000001	-0.000000471	-0.000000722
	(0.0000025)	(0.00000251)	(0.00000252)
Bwd cits	0.000574***	0.000585***	0.000558***
	(0.000083)	(0.0000837)	(0.0000829)
NPL cits	-0.000265***	-0.000262***	-0.000282***
	(0.0000962)	(0.0000945)	(0.0000953)
FWD cits 7yr	0.000508***	0.000495***	0.000498***
	(0.0000577)	(0.0000577)	(0.0000578)
Generality (Hyp 1)	-0.0796***	-0.0353***	-0.0358***
	(0.00820)	(0.00548)	(0.00548)
pre_1995_general	0.0806***		
	(0.0107)		
Originality (Hyp 4)	-0.0375***	-0.0799***	-0.0372***
	(0.00756)	(0.0111)	(0.00776)
Radicalness (Hyp 5)	0.0159**	0.0155**	0.0254***
	(0.00703)	(0.00704)	(0.00899)
pre_1995	-0.0173***	-0.0343***	0.0288***
	(0.00593)	(0.0110)	(0.00418)
Renewal	0.00471***	0.00475***	0.00494***
	(0.000492)	(0.000498)	(0.000496)
pre_1995_original		0.0735***	
		(0.0134)	
pre_1995_radical			-0.0200*
			(0.0103)
Log pseudolikelihood	-4486.1491	-4501.5759	-4513.841
Wald chi2	1125.34	1162.46	1182.23
Pseudo R2	0.2009	0.1981	0.1960
Sensitivity	3.66%	3.52%	3.74%
Specificity	99.80%	99.77%	99.80%
Correctly classified	95.28%	95.25%	95.28%
Linktest (_hatsq P>|z|)	0.001	0.000	0.000
Hosmer-Lemeshow (Prob > chi2)	0.4419	0.3366	0.8021
Observations	29,596	29,596	29,596

Average Marginal Effects; Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

3.7 Conclusion

The aim of this study was to see how patents underlying successful drug products differ from other pharmaceutical patents. A series of analysis was conducted based on the Orange Book patents as well as the drug substance patents therein.

To begin with, a few descriptive observations are noteworthy. Drawing on the patent family concept, it was observed that nearly half of Orange Book patents and about 30% of drug substance patents have other patent family members in the Orange Book. This indicates that a smaller set of inventions underpin the Orange Book drugs Moreover, ownership of drug substance patents is more dispersed in comparison with all Orange Book patents. Finally, Orange Book patent come from diverse technology fields while some are closely related to pharmaceutical products (e.g. organic chemistry) some are less intuitive.

Results indicate the actual breadth of applications of an invention (generality) lowers the probability of being an Orange Book patent. Likewise, broader "*envisioned*" applicability (patent scope) of the invention reduces the probability of being an Orange Book. These observations indicate Orange Book patents have more focused applications than other patents.

Broader legal scope boosts the probability of being an Orange Book patent. This might be because inventors carve out larger legal protection for more valuable inventions.

In terms of knowledge recombination, higher technological knowledge diversity reduces the probability of being an Orange Book patent while technological knowledge distance of a patent increases the probability.

Regarding collaborative inventions, being assigned to multiple entities does not have a significant impact on the probability of being an Orange Book patent while being assigned to different organizational types reduces the probability of being an Orange Book patent. This might be because organizations tend to collaborate on less valuable inventions and conduct more important projects by themselves for competitive purposes.

A subsample analysis on drug substance patent versus Orange Book patents indicates different dynamics. Broader applications (i.e. generality) boosts the probability of being a drug substance patent. This may indicate active chemical substances have broader applications than other inventions. Similarly, broader "envisioned" applicability of an invention boosts the probability of being a drug substance patent.

In terms of knowledge recombination profile, higher technological knowledge diversity (i.e. originality index) increases the probability of being a drug substance patent. However, higher the technological knowledge distance reduces the probability of being a drug substance patent.

The combined observations regarding knowledge recombination indicate when focusing on the Orange Book patents, the tension view on creativity is more relevant while the drug substance patents within the Orange Book show traces of the foundational view.

A subsample analysis was conducted on the way breadth of application (generality) and knowledge recombination impacted the probability of being an Orange Book before 1995. Results indicate that for pre-1995 patents, increasing breadth of application (generality) boosts the probability of being an Orange Book patent more than

post 1995 patents. For knowledge recombination, for pre-1995 patents, more knowledge diversity boosts the probability of being an Orange Book patent while more distant recombination reduces the probability of being an Orange Book patent. Hence, over time three has been changes in the profile of patents that end up in the Orange Book.

A note on shortcomings and future research is in order. For a start, marginal effects are mostly very small. Moreover, the research was based on the assumption that Orange Book patents represent a set of valuable inventions against which other patents can be benchmarked. However, there is little information on the politics and incentives of listing a patent in the Orange Book by applicants as well as the FDA. The existence of multiple patents of the same family and listing of some design patents are issues pointing to the need for more information in this regard.

Another issue is that some patent indicators are proxy measures and are meant to represent complex concepts such as knowledge diversity and distance. The Originality index (knowledge diversity proxy) counts all bits of knowledge based on IPC patent class regardless of being in the same class as the citing patent or not. Hence, it is a noisy measure of familiarity with knowledge components and there is some indicator overlap with radiclaness. It would be desirable to find other measures to corroborate the observations. Some indices such as originality and radicalness are complex mathematical constructs and difficult to interpret. Standardized, and up-to-date assignee names are hard to come by and introduce unnecessary measurement errors in quantitative, large sample studies.

CHAPTER FOUR: DRUG DISCOVERY INNOVATION: A SYSTEMIC VIEW

Abstract

An expert opinion survey was conducted to explore the barriers and drivers of drug discovery innovation. Some top drivers and barriers to innovation are negatively influenced or caused by the "molecular reductionist" drug discovery paradigm. The barriers to innovation show traces of several systemic level failures. Lack of change in the fundamental rules of the game has created a "lock-in/path dependency failure" in which the innovation system has failed to adapt expeditiously. Deficiencies in firm capability development have led to "transition failures". Moreover, hard (i.e. regulatory) and soft (i.e. cultural) institutional failures, along with "regulatory capture" are observable. Respondents possess nuanced knowledge of broad R&D spending and drug approval trends. They consider the overall drug approval rate and R&D spending to be stagnant.

Keywords: Innovation System; Drug Discovery; Eroom's Law, Survey

4.1 Introduction

The pharmaceutical industry is fundamentally dependent on innovation. Cutting edge research, new knowledge creation, new drug development, and improving existing drugs are the driving forces behind the industry. The occasional success in developing a

new therapy for an untreated condition is the industry's defining hallmarks (Petrova, 2014).

The pharmaceutical industry can be characterized as a system or network because of the multiplicity of actors in the innovative endeavors; e.g. firms, universities, research organizations, financiers, regulatory entities and consumers. Moreover, in the course of the past few decades the global pharmaceutical industry has experience significant changes in technology (e.g. the emergence of biotechnology), demand (e.g. cost-containment imperatives) and institutions (e.g. patent law) (Mckelvey et al., 2004) that make a systemic analysis desirable.

However, the industry faces numerous challenges. For over a decade the industry has been scrutinized for its unsustainable drug discovery and development model, feeble innovative output, focus on incremental rather than radical innovation, excessive regulation, and lack of venture capital investment. At the same time, solutions are few and far between (Tait, 2007). For instance, no effort in the past 60 years has had a meaningful impact on the innovative output. The industry needs to invent a new R&D model. According to Jean-Pierre Garnier, the former chief Executive Officer of GlaxoSmithKline, "*R&D productivity is the number one issue*". Without addressing this issue, probably, no other solution will work (Munos, 2009).

The combination of these issues calls for a holistic perspective in the analysis of the pharmaceutical sector. This study adapts the "*innovation systems*" approach to address the underlying challenges of innovation in the pharmaceutical sector.

4.2 Conceptual Framework

This study draws on the "innovation systems" conceptual framework; hence, a brief introduction of the framework is in order. It is worth mentioning that the innovation systems literature is vast and the aim of this section is to introduce what is relevant for the current study. First systemic levels and dimensions of analysis are introduced, then two broad and complementary perspectives in innovation systems (i.e. structural vs. functional) analysis is delineated.

4.2.1 Systemic Levels of Analysis

Innovation systems can be conceptualized at different levels of analysis, e.g. national, regional, sectoral, or technological; however, creation, diffusion, and use of knowledge underpin all conceptualization levels. The initial conception of the approach was at the national level and other levels of analysis were inspired by the initial works (Carlsson et al., 2002; Lundvall et al., 2009).

Innovation systems can also be analyzed from different dimensions. A common dimension is the physical or geographical aspect. Sometimes the level of interest is a specific country or region. Many policies and regulatory frameworks materialize at the national level; therefore, the national level is a salient level of analysis. In other instances, the aspect of interest may be a specific sector or a particular technology. With the advent of information technology most economic activities may have an international dimension as well. Another dimension is time. A snapshot of a dynamic innovation system will inevitably be different depending on the time horizon of the analysis (Carlsson et al., 2002).

In defining the national innovation systems (NIS or NSI) a distinction can be made between a narrow definition and a broad definition. The narrow definition focuses on science and technology, and encompasses those institutions that are directly involved in promoting the acquisition and dissemination of knowledge and are the main sources of innovation. The broad conception includes learning, innovation and competence-building at various levels of aggregation (Lundvall et al., 2009). According to the broad definition, the "narrow" institutions are embedded in the wider context of the socio-economic system, including economic policies, and are influenced by the broader contextual institutions in the rate, direction and relative success of innovative endeavors (Freeman, 2002). Hence, a national innovation system can be defined as "*that set of distinct institutions which jointly and individually contribute to the development and diffusion of new technologies and which provides the framework within which governments form and implement policies to influence the innovation process.*" In other words, "*it is a system of interconnected institutions to create, store and transfer the knowledge, skills and artifacts, which define new technologies*" (Carlsson, 2006). The national dimension of the system does not emanate solely from the innovation policy aspect but also from shared language, culture, national policies and legal and regulatory frameworks that influence the innovative milieu (Carlsson, 2006).

Beginning from the early 1990s, the interest at the regional level promulgated related concepts such as "learning regions", "innovative milieus", "industrial districts", and "local productive systems". In addition to the general systems of innovation literature, the regional innovation system (RIS) also draws on the regional science

tradition that emphasized both the role of proximity and location specific rules and norms. While there is no generally accepted definition of an RIS, it is usually conceptualized as "*a set of interacting private and public interests, formal institutions, and other organizations that function according to organizational and institutional arrangements and relationships conducive to the generation, use, and dissemination of knowledge*" (Doloreux and Parto, 2005). The assumption is that these forces incentivize firms in the specific region to develop forms of capital in line with norms, values, and interactions within the community and reinforce the regional innovative capabilities. The regional innovation system concept sets forth detailed analysis of the interplay of innovation, learning and the economic performance of particular regions (Doloreux and Parto, 2005).

According to Malerba (2002) "*sectors*" offer a crucial level of analysis for the study of innovation and production. There are two main approaches to studying sectors. The first one revolves around the industrial economics literature that puts little emphasis on the role of non-firm organizations, knowledge production and learning activities. The second approach is much more heterogeneous and provides detailed empirical insights into the workings of the sectors, mostly from a single dimension such as competencies, production features and innovation; hence, lacking an integrated sectoral analysis. The sectoral innovation systems (SIS) approach is an attempt at providing a multidimensional, integrated and dynamic view of sectors. It offers a dynamic, coherent and multidimensional take on sectors. A sectoral system is "*a set of products and the set of agents carrying out market and non-market interactions for the creation, production and*

sale of those products". The system encompasses a specific knowledge base, technologies, inputs and demand (Malerba, 2002).

The complexity of technology development and the required multilevel interaction can be cast under the innovation systems concept, thus forming the technological innovation system (TIS) approach. Technology development is contingent upon interrelated processes with roots in various economic fields of activity and a multiplicity of actors and institutions. Moreover, the broader context of the technology includes adjacent sectors (figure 15) (Wirth and Markard, 2011). The original conception of the technological system was in terms of a network of agents interacting under a particular institutional infrastructure that are involved in the generation, diffusion and utilization of a specific technology. This definition facilitated the specification of the system at different levels of analysis such as technology in the sense of a knowledge field, a product or an artifact, a set of related products, and artifacts with a particular function (e.g. health care or transport) (Carlsson et al., 2002).

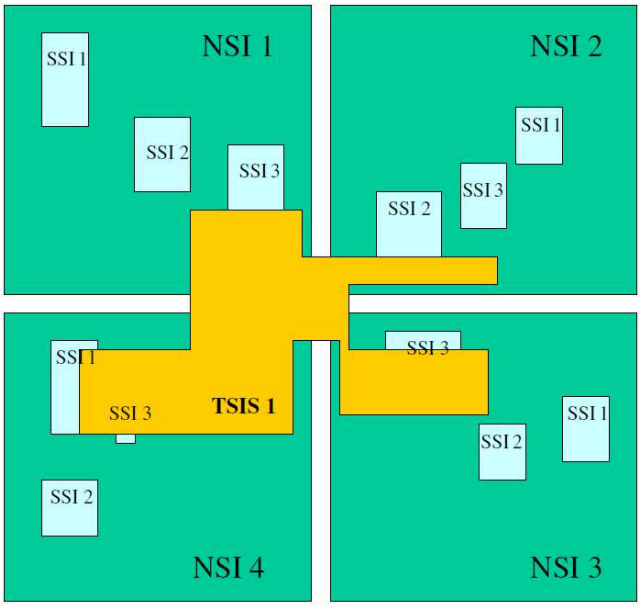

Figure 15 National, Sectoral, and Technological Innovation Systems: Boundary Relationships

Source: Hekkert et al. (2007)

TSIS: Technology Specific Innovation System; SSI: Sectoral System of Innovation; NSI: National System of Innovation

At the national level the complexity of the innovation system may be quite extreme with huge number of actors, networks and institutions. However, with the reduced number of structural components at the level of a technology-specific innovation system, a dynamic analysis is more realistic (Hekkert et al., 2007).

Drawing on the aforementioned arguments, this study is closer to a sectoral system study with the physical dimension largely bound to the U.S. and a cross-sectional time dimension.

4.2.2 Structural versus Functional Perspectives

The systems perspective can further be divided into two complementary perspectives of the functional perspective and the structural approach (see e.g. Bleda and del Río, 2013).

Policy analysis under a structural perspective focuses on the problems related to the structural make-up of the innovation system. In other words, deficiencies in system components and their interaction is the focus of attention rather than their impact on system performance (Bleda and del Río, 2013). Structural elements are relatively stable over time hence reflect the static aspect of the system. They change faster at the early stages of an innovation system formation; as the system gets established, changes become slow and can only be perceptible from a historical perspective (Suurs et al., 2010).

The structural elements of an innovation system are depicted in table 15. Technology is regarded as a structural component as well as the upshot of the system. In other words, technology development and diffusion is the ultimate goal of the system but at the same time technology is introduced by actors and evolves in the system (adapted from Hellsmark and Jacobsson. 2009). It is worth noting that Edquist and Hommen (2008) add the notion of "constituents" as an intermediate concept to the structure of the system. The constituents of an innovation system include both components and the relations among them. The components and the relations form a "whole" different from their "individual" properties. Differing constituents yield different institutional set-ups (adapted from Edquist, 2005).

Table 15 Structural Entities of an Innovation System

Entity	Definition
Technology	Consists of artifacts (e.g. machinery), coded knowledge (e.g. patents) and tacit knowledge (embodied in people)
Actors	Individuals, firms, and organizations influencing technology development
Networks	Are non-market relations between actors and can be for learning (e.g. university-industry links) or political (i.e. aimed at changing institutions) purposes
Institutions	Norms, beliefs, routines, rules, standards, etc. that shape and regulate the relationships and interactions in the system

Source: Hellsmark and Jacobsson (2009)

Another strand of literature focuses on what actually "happens" (i.e. processes) in the innovation system (Edquist, 2004) rather than its mere structure. Functions or activities are the main processes of an innovation system that contribute to the overarching goal of the system which is the development, diffusion and utilization of innovations (Sisko et al., 2013). These functions also yield a basis for performance evaluation and the comparison of system dynamics across systems (Bergek et al., 2008).

There is an interaction between the structural configuration of a system and the system's functional profile (Suurs and Hekkert, 2009). Individuals and organizations carry out the activities and institutions provide incentives or obstacles towards the attainment of these functions. Additionally, grasping the relationship between components and activities is necessary in the comprehension and explanation of innovation processes (Edquist, 2004). Table 16 depicts the system functions commonly evoked in innovation system studies.

Each system function can be attained through different mechanisms. Moreover, it is also possible to come up with activities that contribute negatively to a system function (Suurs et al., 2010). The analysis of system functions helps locate "*inducement*" and "*blocking*" mechanisms. Inducement mechanisms promote the development of a TIS (e.g. price change in favor of a technology), while blocking mechanisms are market or systemic failures (e.g. institutional deficiencies) that thwart the development of the innovation system (Sisko et al., 2013). Features of the structural components as well as the larger contextual issues may also be at fault with blocking mechanisms. Hence, the working of a TIS is only partly driven by internal dynamics of the system (Bergek et al., 2008). For instance, in the emerging TIS of "IT in home care", examples of inducement mechanisms are growth potential and government R&D policy; blocking mechanisms are absence of standards and poor demand articulation. Mapping the relations between these mechanisms and the functional profile of the system is useful in decision-making (Bergek et al., 2008). Table 17 depicts a number of systemic failures posited by various authors. Transition, Transition and learning failure have been used to refer to the same phenomenon in the literature (Klein Woolthuis et al., 2005).

This study draws on the functional perspective, the concept of blocking and inducing mechanisms to explore the drivers and barriers to drug discovery in the U.S. pharmaceutical industry. Moreover, the structural aspects of the system are also touched upon through a survey instrument.

Table 16 Functions of a Technological Innovation System (TIS)

Function	Definition
Knowledge development and diffusion	Usually placed at the center of a TIS; it is concerned with the knowledge base of the TIS (globally) and the performance of the local TIS is measured against
Influence on the direction of search	Refers to mechanisms influencing the direction of search in terms of competing technologies, applications, markets, business models, etc.
Entrepreneurial experimentation	An innovation system without vibrant experimentation will stagnate. Variety of experimentation also matters, e.g.: number of new entrants; number of different types of technology applications; the breadth of technologies used
Market formation	For the overall TIS, the market proceeds through "nursing markets" providing a "learning space" to a "bridging market" and finally developing into mass markets.
Legitimation	Social acceptance and compliance with relevant institutions; the formation of new industries requires legitimacy.
Resource mobilization	For the evolution of a TIS, mobilization of a range of different resources (e.g. competence/human capital; financial capital; complementary assets) is needed
Development of positive externalities	The generation of positive external economies is a key process in the formation and growth of a TIS. External economies or free utilities may be pecuniary or non-pecuniary.

Source: Bergek et al., 2008

Table 17 Examples of Systemic Failures

Systemic failure	Definition
Infrastructural failures	Physical infrastructure (such as IT, telecom, and roads) and the science and technology infrastructure needed for the functioning of the actors in the innovation system
Transition/ Capabilities/ Learning failures	Inability of firms to adapt to new technological developments
Lock-in/path dependency failures	Inability of complete (social) systems to adapt to new technological paradigms
Hard institutional failure (formal institutions)	Failures in the framework of regulation and the general legal system
Soft institutional failure (informal institutions)	Failures in the social institutions such as political culture and social values
Strong network failures	Missing out on new outside developments because of close links between actors
Weak network failures (dynamic complementarities' failure)	Lack of linkages between actors leading to insufficient use of complementarities, interactive learning, and new idea generation

Source: Klein Woolthuis et al. (2005)

4.3 Historical Evolution of the Pharmaceutical Sector

While a thorough investigation of the historical evolution of the industry is beyond the focus of this study, a brief review is useful.

The evolution of the industry in the early periods not only molded the institutions but also shaped firms' organizational capabilities that have implications to this day (Henderson et al., 1999). Previous studies have discussed the history of the modern pharmaceutical industry in terms of a number of time periods or epochs (e.g. Henderson et al., 1999; McKelvey et al., 2004; Malerba and Orsenigo, 2015).

The pharmaceutical industry came into being in the late nineteenth century as a subset of the emergent chemical sector (Malerba and Orsenigo, 2015). The first epoch, corresponding to the early stages of the industry spanning the period 1850 to 1945 and is characterized by a lack of in-house R&D and no tight linkages to science. Mass production of pharmaceutical began in the late nineteenth century in the U.K. and the U.S. Little new drug development occurred and the research, if any, was based on primitive methods. However, starting the 1930s, the emerging sectoral system contained firms, universities, and to some extent regulatory players. Universities provided basic science and trained chemists. Since early days some companies (e.g. Merck and Pfizer) were innovators while others (e.g. Bristol-Myers and Warner-Lambert) focused on imitation and inventing around (McKelvey et al., 2004). The structural aspects of the system set in during this period.

The second epoch, or the "*random screening*" period was the golden age of pharmaceuticals that ran from 1945 to early 1980s (or 1945 to 1990s according to Henderson et al., 1999). The random screening approach entailed randomly assessing natural and chemical compounds in test tubes and live laboratory animals for potential therapeutic properties. This method was devised because of lack of specific and detailed knowledge on disease mechanisms (McKelvey et al., 2004). A turning point in terms of patent protection occurred in 1946 with the granting of a patent for streptomycin[51]. Before that antibiotics were denied patents because they were deemed "naturally occurring substances". This marked the beginning of patent regime tightening (Malerba

[51] First antibiotic effective against TB

and Orsenigo, 2015). In addition to patent protection, firm-specific tacit knowledge and skills of screening also functioned as a mechanism to protect economic returns (McKelvey et al., 2004). The patent institution and firm-specific tacit knowledge that help perpetuate the interests of established players might have made the system susceptible to lock-in/path dependency failures. To sum up, this period was driven by R&D which led to many drug discoveries. Firms enjoyed high profitability and many medical and pharmaceutical knowledge developments occurred in this period (McKelvey et al., 2004).

The third epoch heralded a knowledge revolution that brought about the "*guided search*" learning regime. Advances in molecular biology increased understanding of the roots of diseases as well as the way drugs work. Molecular genetics and recombinant DNA technologies also joined the revolution in mid-course and opened new avenues for innovation. However, new knowledge did not have any automatic influence on the competitiveness of existing firms as they incorporated it into their conventional small molecule discovery. For instance, they automated screening for new drug targets. The emergence of new biotechnology firms (NBFs) was the most apparent crystallization of the changes (McKelvey et al., 2004). In summary, while new knowledge and actors entered the innovation system, main actors adopted and incorporated the new knowledge into their extant routine. Hence, continuing the "lock-in/path dependency failures" preconditions.

Finally, Malerba and Orsenigo (2015) label the first decade of the new century the "*winter of discontent*". In the 1990s the pharmaceutical industry witnessed good

economic and financial performance, and enjoyed respect in the civil and policy circles. However, the turn of the century marked a dramatic change of fortune with declining innovativeness, public perception reversal emanating from drug withdrawals, intellectual property (IP) disputes, and drug price hikes (Malerba and Orsenigo, 2015). In short, this period is a manifestation of numerous systemic failures that has intertwined causes hence making innovation system studies in this sector highly desirable.

4.4 Institutional Set-up

The innovation systems concept is an institutional approach "par excellence". Serious scholars in the realm of technology development have always been aware of the impact of institutions in molding technology advancement efforts (Nelson and Nelson, 2002). The institutional set-up can harbor an important selection mechanism of products and organizations. In this line, a couple of studies highlighting the importance of broad institutional issues in pharmaceutical innovation are reviewed here.

The pharmaceutical industry was born in Switzerland and Germany, in part, due to their strength in university research and scientific training in related fields. In the U.S., the WWII government investment in commercial penicillin development and chemical structure analysis shaped the industry. The successful commercial development of penicillin showcased the commercial potential of drug development leading to pharmaceutical companies' development of internal R&D capabilities and large R&D investments. Generally speaking, four crucial areas have been instrumental in the institutional set-up: public support for health research; intellectual property (IP); product approval procedures, and healthcare system and reimbursement structure. Despite

differences in nature and amount, nearly every government supports health-related research. With drugs being a regulated product, approval procedures have a huge impact on development cost and firm competitive position. Finally, healthcare and reimbursement systems vary widely between countries. These impact the potential of capturing rent from innovation. Fragmented health care markets and subsequent low buyer bargaining power allowed better innovation rent capturing in the U.S. (Henderson et al. 1999).

Henderson et al. (1999) explore the impact of the molecular biology revolution on the pharmaceutical industry evolution. They mention that different national innovation systems may suit the promotion of different types of innovations. Competence destroying innovations (e.g. the case of biotechnology) call for the formation of new organizational and institutional forms; hence, they tend to emerge in locations favorable to institutional flexibility and variety. In contrast, in competence enhancing innovations; e.g. rational drug design; the relative differences among countries can be linked to the strength of existing institutional arrangements such as strong links to universities or other issues impacting new technology access.

From a dynamic approach, Mckelvey et al. (2004) observe changing actors, relationships, and networks over time through the lens of sectoral innovation systems (SSI). The actors and their relationships are embedded in and affected by contextual factors such as public policy and legal systems. A few features of the system they highlight are as follows. First, system development was not conscious and the self-organization started from extant institutions and organizations. Second, despite

fragmentation, there is also integration; for instance, new biotechnology firms (NBFs) could not flourish without public funding of academic research and the contracts and demand from large firms. Third, while the system is not entirely coherent and at rest, it is self-sustaining, with agents engaging in complementary functions. Fourth, the trends cannot be simply compartmentalized as processes of deepening division of labor or processes of horizontal or vertical integration. Division of labor and integration are taking place at the same time. Moreover, agents are changing their functions and position in the networks. For instance, universities not only specialize in their core activities (i.e. teaching and research), but they also diversify downstream into commercializing their new products which requires the creation of new incentives and organizational forms. Finally, the SSI changes over time in response to different exogenous shocks and internal learning and selection processes. For instance, the thalidomide case[52] led to tough product approval procedures that changed R&D costs, industry structure, drug prices, and competitiveness of firms and industries (Mckelvey et al., 2004).

To sum up, innovative output and specialization pattern of the industry can be influenced by the host national innovation system based on unique institutional incentives and barriers.

4.5 Research Question

According to Sisko et al. (2013) a TIS analysis can be conducted in three steps. Firstly, the structure of the system in terms of comprising components is identified. Then,

[52] Thalidomide was widely used in the late 1950s and early 1960s for the treatment of nausea in pregnant women. As it later transpired in the 1960s, thalidomide resulted in severe birth defects. The use of thalidomide was banned in most countries at that time; however, it proved useful for leprosy and later, multiple myeloma (Kim and Scialli, 2011).

the performance of the functions of the system is evaluated. Finally, inducement and blocking mechanisms are identified from the preceding steps that lay the ground for crafting policies. Consistent with this logic, steps two and three can be conflated into a single research question for the present study: "*what aspects of the pharmaceutical innovation system drive or hinder innovation in drug discovery?*" The study does not seek to conduct a broad mapping of the innovation system but it seeks to pursue the more focused objective of exploring possible reasons behind the productivity paradox in the pharmaceutical sector. Hence, it is more focused on the functional aspect of the system and identification of possible systemic failures. However, structural elements are not completely ignored.

As such, the present research is not a traditional hypothesis testing endeavor but more of an exploratory study that elucidates crucial aspects of the drug discovery innovation system. Moreover, it complements the preceding two essays by bringing in the voices of those involved in innovative endeavors. This is a crucial effort since quantitative analysis of patent data may miss important underlying dynamics of innovation. As an example of a qualitative study on systemic issues, Swan et al. (2007) conduct a three-year exploratory study in the U.K. and U.S. biomedical sectors to identify factors facilitating and impeding innovation projects across contexts. The first phase of their study involved an interview-based survey with a range of individuals representing key stakeholder groups who had significant experience of working in early-stage biomedical innovation projects. The second phase consisted of longitudinal case studies of innovation projects representing different approaches to organizing biomedical

innovation. Likewise, Sisko et al. (2013) in a study of the inducement and blocking mechanisms in the Finish life sciences innovation system, draw on 33 qualitative interviews with senior managers and decision-makers.

4.6 Data and Methods

This study draws on both primary and secondary sources of data and information. According to Sauermann and Roach (2013) surveys are important sources of data in innovation studies. Hence, a concise online survey of inventors and experts in the drug discovery field was conducted to produce firsthand valuable insights on many aspect of the innovation system. Given the complexity of the issue at hand and low response rates, this primary data is augmented with existing studies, datasets, statistics and commentaries on the pharmaceutical innovation system. The secondary sources not only complement the primary sources, but they also can offer reference points to interpret, validate, and qualify the primary data where necessary.

Given the nature of the issue at hand, the study does not draw on quantitative methods and statistical inference. Descriptive statistics will be drawn upon as appropriate.

4.6.1 Survey Instrument

To collect primary data, a concise survey instrument was developed drawing largely on Scannell et al. (2012) and the innovation system literature. Appendix C depicts the instrument. There are 12 substantive questions[53] on the instrument. Two questions deal with the "background" of the respondents in terms of field of specialization and

[53] One is consenting to the terms, the other asks for relevant comments to the study making total 14 questions

position. Fields of specialization were taken from the FDA therapeutic areas[54] used in reporting approved drugs.

Question 4 asks the respondent to rank up to five "drivers of innovation" from among eight options presented to them. The options (depicted in table 18) were developed from the literature and were presented in a dropdown menu format. The respondent could also suggest and rank own options.

[54] FDA therapeutic areas will be used: "New FDA Approved Drugs By Medical Area | CenterWatch." Accessed July 10, 2016. http://www.centerwatch.com/drug-information/fda-approved-drugs/therapeutic-areas.

Table 18 Drivers of Innovation: Response Options

Driver	Justification	Approx. Match to Functions
R&D investment	Basic resource input into innovation; Crux of Eroom's Law	Resource mobilization
Basic science (e.g. scientific publications)	Knowledge-base of pharmaceutical innovation	Knowledge development and diffusion
Skilled R&D scientists	Basic requirement of innovation especially in a science-based area	Resource mobilization
Good R&D management	Good "*ba*" (Nonaka, 1998); appropriate internal direction for search	Influence on the direction of search (here, the emphasis is on "internal guidance" from corporate strategy and not influence form policy incentives)
Collaborative R&D with other outside entities (e.g. universities)	Trifecta model (Petrova, 2014); Kaitin and DiMasi (2011)	Resource mobilization (in terms of accessing complementary competences); Knowledge development and diffusion; Development of positive externalities
The diversity of knowledge available to the inventors (e.g. enzymology, toxicology, etc.)	"*Tension*" view on creativity; (Kaitin and DiMasi, 2011)	Resource mobilization
The depth of specialized knowledge available to the inventors	"*Foundational*" view on creativity (Kaitin and DiMasi, 2011)	Resource mobilization
Market size of the drug	Market pull arguments of innovation; Acemoglu and Linn (2004)	Market formation

R&D investment was offered as a driver of innovation based on the Errom's Law of rising corporate R&D spending and almost flat output as articulated by Scannell et al. (2012). Basic science reflects the science-intensity of pharmaceutical innovation. For instance, as mentioned previously, Ward and Dranove (1995) divide R&D into the three

stages of government-funded basic research, publication in medical journals, and industry-funded applied R&D.

The "Skilled R&D scientists" option is offered as another basic ingredient of innovation. The inclusion is especially important given the new paradigm of drug discovery based on automation (e.g. high throughput screening of drug targets) and taking the human judgment out of the R&D search process. In the words of Scannell et al. (2012) if the logic is that "automation, systematization and process measurement have worked" in other industries, what is the point in relying on the random efforts of "chemists and biologists".

Good R&D management was included in the option list because of the importance of context and incentive structure in the performance and outcome of human resources. For instance, Nonaka and Konno (1998) contend that knowledge is embedded in shared spaces called "*ba*", a word borrowed from Japanese philosophy. *Ba* can be physical (e.g. office), virtual (e.g. teleconference), mental (e.g. shared experiences or ideals) or a combination of these. *Ba* is a platform for advancing individual or collective knowledge. They offer the SECI model of knowledge creation consisting of socialization (i.e. sharing tacit knowledge), externalization (i.e. expression of tacit knowledge in comprehensible form), combination (i.e. transformation of explicit knowledge into more complex forms) and internalization (i.e. conversion of explicit knowledge into organization's tacit knowledge). They contend that support of knowledge creation required understanding of the delicate characteristics of the *ba*. Likewise, according to Johnson (1996) resolving problems, competing effectively and diversifying a company requires building on human

resource capabilities. For an effective technical organization, core competence is not a stagnant research area but represents the cutting edge of the field and striving to exploit it towards organizational goals. In his words "*good R&D management is tolerant of failures and near misses and determined to try again*".

Collaboration in R&D was raised because of the studies arguing for a new model of drug discovery based on a network of innovation stakeholders—including large and small pharmaceutical and biotechnology companies, academic research centers, contract research organizations, public–private partnerships, and patient groups— sharing the risks and rewards of the innovations. Some believe there is already a shift towards this (Kaitin and DiMasi, 2011; also see Petrova's (2014) trifecta model).

The inspiration for options raising the "diversity of knowledge" and the "depth of specialized knowledge" available to the inventors comes from the "*tension*" and "*foundational*" views on creativity. The former asserts that deep knowledge can lead to myopia to the extent that the recombination of distant or diverse knowledge is needed in order to see new ideas, while the "*foundational*" view touts a deep understanding of the knowledge domain as a prerequisite to generate breakthrough innovations (Kaplan and Vakili, 2015).

Finally, the "market size of the drug" reflects the intuitive importance of the existence of a market as a driver of innovation which is also empirically examined (see e.g. Acemoglu and Linn, 2004).

Six questions tried to construct a profile of the innovation system structure. Table 19 depicts the questions and their implications for the structure and function of the innovation system.

Table 19 Questions Dealing with the Innovation System Structure

Question	Justification	Structural Entity	Approx. Implications for System Functions
List three most important public R&D funding entities (e.g. NIH)	Capture most important public funders; can have policy implications by helping identify most important funders	Actor	Resource mobilization
List three most important private R&D funding entities (i.e. funding outside R&D such as a venture capital company or big companies)	Capture most important private funders. Interesting to see large firms vs. venture capital role in the eyes of practitioners	Actor	Resource mobilization
List three most important producers of basic research (e.g. a specific university)	Can have policy implications in terms of identifying successful entities/programs	Actor	Knowledge development and diffusion
List three most enabling legislations or regulations	Has clear policy implications by identifying successful initiatives	Institution	Influence on the direction of search
List three most burdensome legislations or regulations (i.e. hindering innovation)	Has clear policy implications by identifying legislative barriers to innovation	Institution	Influence on the direction of search
List top three new companies in your area of expertise and reason for inclusion in this list (e.g. for new technology development; for new market creation; etc.)	Identifies most innovative entities; Can have policy implications by tracing the sources of their success	Actor; Technology (if identified by the respondent)	Entrepreneurial experimentation

One question tried to put the Eroom's Law (see Scannell et al, 2012) to the test. This attempt was done very subtly by asking respondents to judge corporate R&D spending and new drug approval trends (i.e. if decreasing, stagnant, increasing or they have no information). However, they were asked to choose their own time frame (i.e. past five years, past decade, past few decades) for assessing the trends. New drugs were meant to include new biologics and R&D spending was supposed to reflect inflation adjusted.

One question asked respondents to rank nine barriers to innovation with the option to add own choices. The barriers offered for ranking and their justification is depicted in table 20.

Table 20 Barriers to Innovation: Response Options

Barrier	Justification	Functions
Availability of good drugs for many diseases	"*Better than the Beatles*" problem: a large stock of approved drugs exists and new drugs have a modest incremental benefit over what was already available (Scannell et al., 2012).	Market formation
Over-cautious regulation for safety	Ratcheting up of regulatory burden in response to past drug failures and reduced risk tolerance of regulatory agencies (Scannell et al., 2012).	Influence on the direction of search; Resource mobilization
Inflated R&D wages	Meant to reflect cost pressure on R&D spending	Resource mobilization
Designing drug substances with a single or narrow therapeutic benefits	Over time there has been a shift form looking broadly for therapeutic benefits in active agents to designing molecules for precise effects (Scannell et al., 2012).	Market formation
Complex clinical trials	"*Better than the Beatles*" problem and cautious regulators increased the complexity of medical practice (Scannell et al., 2012).	Resource mobilization
Reduced quality of published science	Knowledge is the main input of R&D; Poor basic science will negatively impact the government-funded basic research, publication in medical journals, and industry-funded applied R&D cycle (adapted from Ward and Dranove, 1995).	Knowledge development and diffusion
Patented or proprietary research tools	Patents on research tools such as sequencing methods and "reach-through" licensing practices in which upstream research tool owner seeks control and royalties from downstream applications (Burk and Lemley, 2003) may have negative impact on innovation.	Resource mobilization; Knowledge development and diffusion
Lack of inter-organizational collaboration in R&D	Given importance of division of labor and cooperation in the trifecta model, lack of collaboration can be an important barrier to innovation.	Development of positive externalities
Companies pursuing the same drug targets	Flip side of automation and new drug discovery approach of molecular reductionism (see e.g. Scannell et al., 2012). With the shift from random screening to targeted rational drug design, "*the discovery process has become more systematic*" (Petrova, 2014). There are "*development races*" for new drug development that are the reason for having "me-toos" (DiMasi and Paquette, 2004).	Market formation; Resource mobilization (R&D for same targets that lead to inflated industry level R&D spending); Influence on the direction of search (lack of guidance or guiding towards one target may lead to this problem)

Respondents were given the chance to share any comments related to "boosting innovation and lifting barriers to innovation". Finally, they were asked to rate their responses in terms of generalizability across other therapeutic areas.

The recruitment/invitation email, with the survey link, contained the George Mason University's Internal Review Board (IRB) approval number, giving it an authoritative air. The instrument was preamble by a consent form approved by the IRB informing the respondent of the anonymity of the responses, risks, benefits, the voluntary nature of the survey, and contact information of the researcher. The SurveyMonkey platform was used for the survey but was customized with the university logo. Different features such as dropdown menus were used to make the survey user-friendly.

4.6.2 Respondents

The unsolicited online survey invitation was sent to potential respondents extracted from drug substance patent inventors, managers/founders of new firms (from angel.co with "drug discovery" and "drug development" tags as of May 2017), speakers on two drug discovery/development conferences (one held one upcoming), one small pharmaceutical consultant network with open member profiles, and LinkedIn professional profiles with "drug discovery" skill tags. Email invitations were sent to those individuals whose email address could be found on the internet. Some emails were constructed based on institutional email formats hence more of these invitations did not go through because the email addresses were not correct. Drug substance patents came from the online Orange Book patent files downloaded June 22, 2016 and April 3, 2017

were used to extract respondents. The patent files have a drug substance flag. About 3200 names[55] from 710 patents were sifted through. About 27% of names were duplicates.

In deciding who is eligible for participation, the LinkedIn and/or institutional profile of the individual were examined for evidence of work in the field of drug discovery and U.S. residency. U.S. residence or experience is emphasized because "*institutional/ country-specific factors*" provide crucial incentives and "*selection processes*" that "*shape*" the pharmaceutical innovation system of each country (McKelvey, et al., 2004). Moreover, having U.S. specific responses would be crucial for policy-relevance.

The inclusion criteria for drug patent inventors were low, i.e. they automatically qualified for having contributed to a drug-substance invention and being a U.S. resident at the time of the invention[56]. The construction of the sample was itself subject to learning and adjustment; it was more strictly applied for LinkedIn respondents because "drug discovery" skills might have been loosely endorsed by others. R&D project/research/group management or leadership experience was one of the main criteria for inclusion in the sample for LinkedIn members. Academic and non-profile sector individuals were mostly selected for experience in industry or depth of their experience. Given scale of manual work, there is some margin of error hence the instrument upfront asked respondents to mention their position. LinkedIn search was done among experts based in the U.S. It seems the search list retuned people based on proximity in my

[55] Estimates are based on rows of data in my excel sheet. Given non-standardized names on patents, working with inventor names is not easy.
[56] City and State of inventors are mentioned on the U.S. patent.

network. However, from the first round of survey invitations, a couple of highly networked and experienced people sent me a LinkedIn connection invitation which helped with the search. In other words, this made me embedded in some very relevant networks. Premium membership was required for the searches which at the time of search (July 23, 2017 through September 1, 2017) required searching the directory for "drug discovery" skills tag.

An interesting observation is that out of about 721 LinkedIn profiles examined as potential respondents, 31 were also in the drug substance patent inventors. This is another concrete indication that LinkedIn profiles can be very relevant for extracting potential respondents.

There is no known accepted population of respondents for this survey and there were numerous biases in constructing the potential respondent list, most important of all would be the ability to find a relevant email address. Some of these biases emanate from limited resources; for instance, if we accept LinkedIn as one source of potential respondents, not all LinkedIn profiles with "drug discover" skills were examined. Moreover, the drug patent inventor list was constructed over a long period of time (stating June 2016) with several months of break. Some email addresses might have become publicly available after the email address search on the internet. Overall, the practice should be viewed as an exercise in contacting experienced people in the field for their valued positions and observations regarding innovation in the sector, much like a qualitative study based on elite interviews.

Many studies in the innovation systems genre are not quantitative. For instance, to study the antecedents of the NIS concept, Sharif (2006) conducts interviews 12 interviews and holds 5 informal conversations. Suurs et al. (2010) draw on digitalized media, historical reports, and expert interviews to study the formative stage of a technological innovation system in the energy sector. Finally, Swan et al. (2007) scrutinize innovation mechanisms at the project level in the biomedical field in the U.K. and U.S. using 97 interviews (44 U.K.; 53 U.S.).

Table 21 depicts the details of the potential respondents and response rates. The invitation to LinkedIn, conference and forum potential respondents were addressed by first name (in few exception titles such as Prof., Dr. was used). The rest of the invitations were generic (i.e. starting: "Dear respondent"). With the exception of the upcoming conference speakers, all addressed invitations have much higher response rate. However, higher response rate might be due to informing them that questions are independent and they can skip any question they find burdensome. Moreover, timing was broken down into realistic, pessimistic and a fast track set of questions that were deemed potentially interesting for the respondents (e.g. drivers and barriers to innovation and R&D trend). The upcoming conference speakers' zero response rate is noteworthy and could be because the conference was not held yet.

One reminder was sent seven or eight days after the original invitation. Forum faculty received a reminder nine days after initial invite and 41 LinkedIn potential respondents had not received a reminder at the time of extracting this data (9/4/2017).

The reminders were not individually addressed. First batch of invitations were sent on July 12, 2017 and the last were sent on August 31, 2017[57].

A note on response rate is in order. According to Sauermann and Roach (2013) online surveys often exhibit lower response rates of around 10–25%. Also, according to Pew Research Center (May, 2017) the telephone poll response rates has been declining for decades and currently stands at about 9%. The total response rate of around 9% is closer to the lower end of the spectrum reported by Sauermann and Roach (2013) but given the fact that the study was pitched as "*part of a doctoral study*", higher organizational position of respondents compared with the researcher (i.e. experienced R&D managers, etc. versus a doctoral candidate), survey subject matter (i.e. industry-level innovation related issues pitched to technical people), one round of reminders, and vacation time (there were a number of automated holiday email responses), the lower response rate might be justified.

[57] The survey may be open beyond this date and still collecting possible responses.

Table 21 Potential Respondents and Response Rates

Source	Potential respondents (Emails sent)	Bounced Emails(approximate)	Responses (click through)	Response Rate (click through) %	Adjusted Sent Emails*	Adjusted Response Rate (click through)* %
New Firms	66	1	4	6.06	65	6.15
New Firms (generated)	19	3	1	5.26	16	6.25
Drug substance (Known, Executive)	194	14	7	3.61	180	3.89
Drug substance (Known, Academic)	37	4	5	13.51	33	15.15
Drug substance (Known, Other)	218	17	6	2.75	201	2.99
Drug substance (Generate, Executive)	140	53	7	5.00	87	8.05
Drug substance (Generate, Accademic+Other)	170	52	2	1.18	118	1.69
Consultant Network	31	4	4	12.90	27	14.81
Forum Faculty	40	0	9	22.50	40	22.50
Upcoming Conference Speakers	27	1	0	0.00	26	0.00
LinkedIn (Executive)	365	35	42	11.51	330	12.73
LinkedIn (Academic)	82	2	20	24.39	80	25.00
Total	1389	185	107	7.70	1204	8.89

* Adjusted for bounced emails (i.e. emails that did not go through); Some respondents only consented but did not respond to any questions; hence, the term "click through".

Overall 107 consented and started the survey but 15 of these did not respond any substantive questions. Two background questions were asked from the respondents. The first asked them to identify up to three specializations. Figure 16 depicts the field of specialization of those who responded to this question and figure 17 depicts a word cloud from SurveyMonkey export tool. In the words of one of the respondents, "*medicinal*

chemistry" is the field developing drugs and was missing from the list. The reason is that the list of specialization came from the FDA drug categorization rather than disciplinary specialization. However, the cloud indicates the respondent pool is highly relevant with "medicinal chemistry" taking center stage and "drug discovery" and "drug development" respectively taking prominent positions in the word cloud. As depicted in figure 16, oncology, pharmacology/toxicology, and "infections and infectious diseases" are respectively the most prevalent main specialization of the respondents.

Figure 18 depicts the organizational position of the respondents. A noteworthy issue is that 81.5% of potential respondents are corporate R&D managers, entrepreneurs, or consultants. People holding these positions are in the best position to judge the drivers and barriers to corporate innovation. The existence of four CEOs among the respondents is also noteworthy. While not all of the people went ahead with all sections of the survey questions, this analysis offers a general profile of the people reached at for the survey.

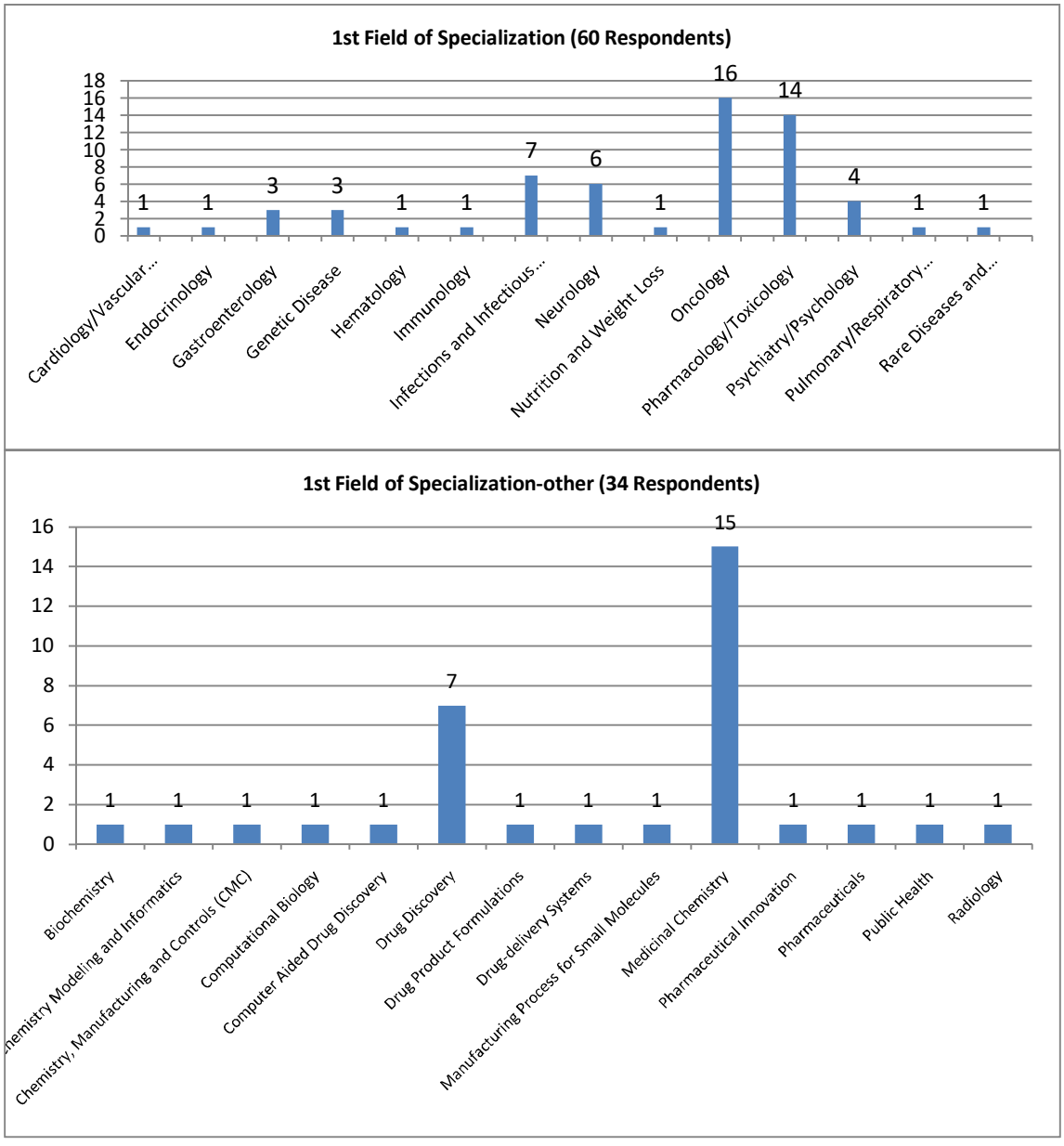

Figure 16 Main Field of Specialization of Respondents (Responses are cleaned and standardized for bottom panel) [Q: What is your main field of specialization? (Choose up to three in order of expertise if applicable or add yours)]

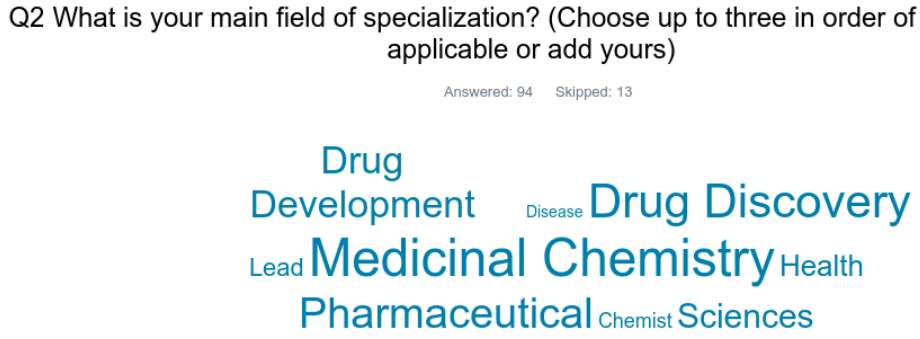

Figure 17 Word Cloud of Fields of Specialization (Based on most important words and phrases respondents used to describe their specialization from SurveyMonkey Export Tools)

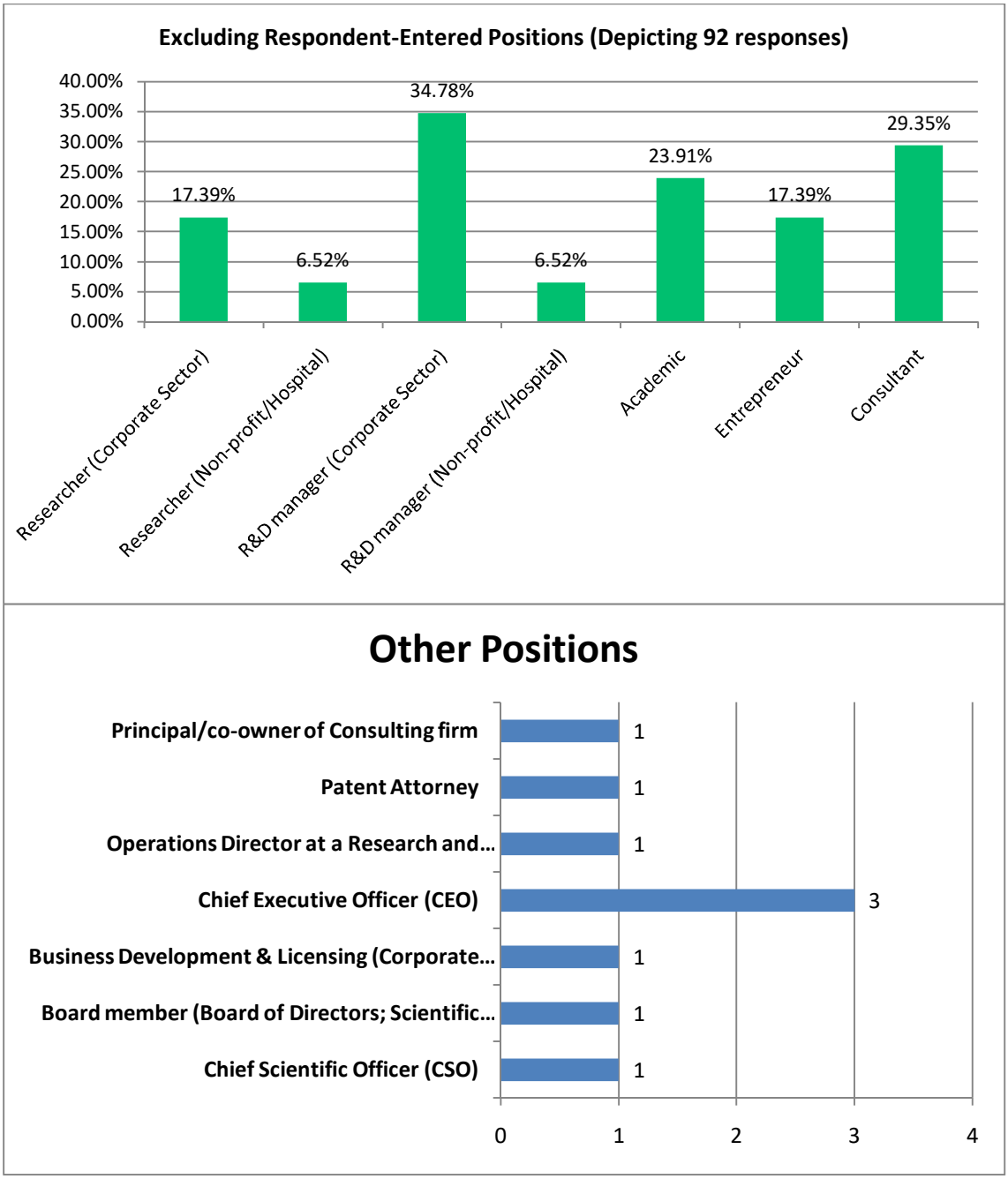

Figure 18 Organizational Position of Respondents [Q. What is your current position? (Optional; select all that applies)]

4.7 R&D and Innovation Trends

The term "Eroom's Law" coined by Scannell et al. (2012), refers to the paradoxical phenomenon of declining efficiency of R&D since 1950. They estimate that the number of FDA-approved drugs[58] per billion US dollars of R&D spending halved about every 9 years since 1950. Similarly, Munos (2009) points out that the new-drug output has been essentially constant despite efforts to boost the numbers. The reason may simply be the limitations of the prevailing R&D model.

It would be instructive to see what the subjective views of industry insiders would be about these observed R&D and innovation trends. This is especially important as their perceptions and beliefs will be the prime driver of decisions impacting these trends and outcomes. To this end, question 11 was devised to have responders judge R&D and drug approval trends. An innovative approach was used by allowing them to choose the time frame within which they could make a judgment (i.e. judging trend over five years, a decade, decades). Overall, 54 responders has some input for this exercise.

Figure 19 depicts the results of the exercise. Overall, a slim majority of the responders vote along the lines of the Eroom's Law in terms of new drug approval trend. Regarding R&D spending, the majority view is that the spending is stagnant. While this is not in line with the Eroom's Law long term trend, R&D spending figures of the recent years, indeed, confirm this trend. This is especially more cogent as the majority tended to judge the trend over the five-year time period.

Understandably, most responders were more comfortable to pass judgment on shorter periods of five years and a decade for both R&D and drug approval trends.

[58] New molecular entities and new biologics

Longer term judgments would require more effort on their parts. Some respondents would not have as much high-level experience of the industry with longer time periods.

These results point to the fact that the overall judgment of the respondents has: firstly, largely been reflective of what has actually been going on in the industry, secondly, their information is generally current and finally, there appears to be not much bias in their rendition of what the facts and figures show.

This broad reporting may mask very specific reasons for the responses hence a closer look at the comments some responders left is instructive. Table 22 depicts the comments along with an attempt at linking them to systemic failures. Regarding drug approval, some comments hinted at improved approvals for a few recent years. Two comments hinted at the "*better than the Beatles*" problem with the move into more difficult drug discovery areas. In a similar vein, another comment mentions that there are more approvals for "*small*" drugs. Difficulty and cost of clinical trial has been mentioned in two comments. Problems of big companies in conducting creative R&D and the doubtful sustainability of their business model have also been mentioned. Finally, one comment raises the important issue of better translational research.

In terms of R&D spending, high cost of R&D and clinical trials, less emphasis on in-house R&D and more on acquiring early-stage products, and less spending on research than development, are important issues raised in the comments.

Table 22 Responder Comments for R&D Spending and Drug Approval Trends (16 responders)

Category	Responder Comment (with some rewording)	Note	System Failure(s)
Drug Approval	New drug approvals decline with the move into more difficult areas.	Hints at "better than the Beatles"	Transition failure
	We are seeing more approvals but for "small" drugs with narrow and orphan indications	Narrow targets?	Transition failure/ Hard institutional failure
	9 out of 10 drugs fail in clinical trial – need for better translational research	Translational research weakness	Transition failure/ Hard institutional failure
	95% of small molecule drugs fail in the clinical trials CAR-T cells are interesting but very expensive for patients	Hint at new technology or paradigm	Transition failure/ Hard institutional failure
	2015 was a strong year and 2016 was less successful		
	Large companies have enjoyed decreasing ability to conduct creative R&D; one major reason is senior management's lack of insight	R&D management issues	Transition failure/ Soft institutional failure
	New drug approval increase is only for last year	Only recent year increase in drug approvals, may be temporary	
	The percentage of prescriptions filled by patented drugs has decreased to about 10%.	Hints at "better than the Beatles"	Transition failure
	New drug approvals increasing slowly	Improvement (if any) is of recent years	
	At least ten years are needed from the lab to approval; hence, approvals are based on ten year-old innovations	There might be a large lag between discovery and approval. A long time needed to see	Hard institutional failure
	New drug approvals in Hematology-oncology has accelerated		

		if Eroom's Law has changed.	
R&D Spending	R&D spending in big companies is stagnant, but there is an increase in venture capital based discovery/development	Data shows discovery investment By VC more flat; development a bit up	Transition failure/ Soft institutional failure
	R&D costs have risen exponentially; hence, the science current dollars buy is less than ten years ago.		Hard institutional failure
	In-house R&D spending is lower because companies seek to source early-stage assets from start-ups, universities and innovation centers.	Hint at "acquisition of pipeline"	Transition failure/ Soft institutional failure
	Research spending is decreasing but development (clinical trials) spending is increasing because of clinical trial costs.	Hint at reduced discovery spending and increased development spending	Hard institutional failure
Other	R&D return on investment declines with the move into more difficult areas	Maybe Return driven industry may have less incentive to invest in discovery	Transition failure
	Trends cast doubt over the sustainability of the pharmaceutical business model		Transition failure/ Soft institutional failure/ Hard institutional failure
	There are many "deferred funding" notices which are ominous signs		Hard institutional failure

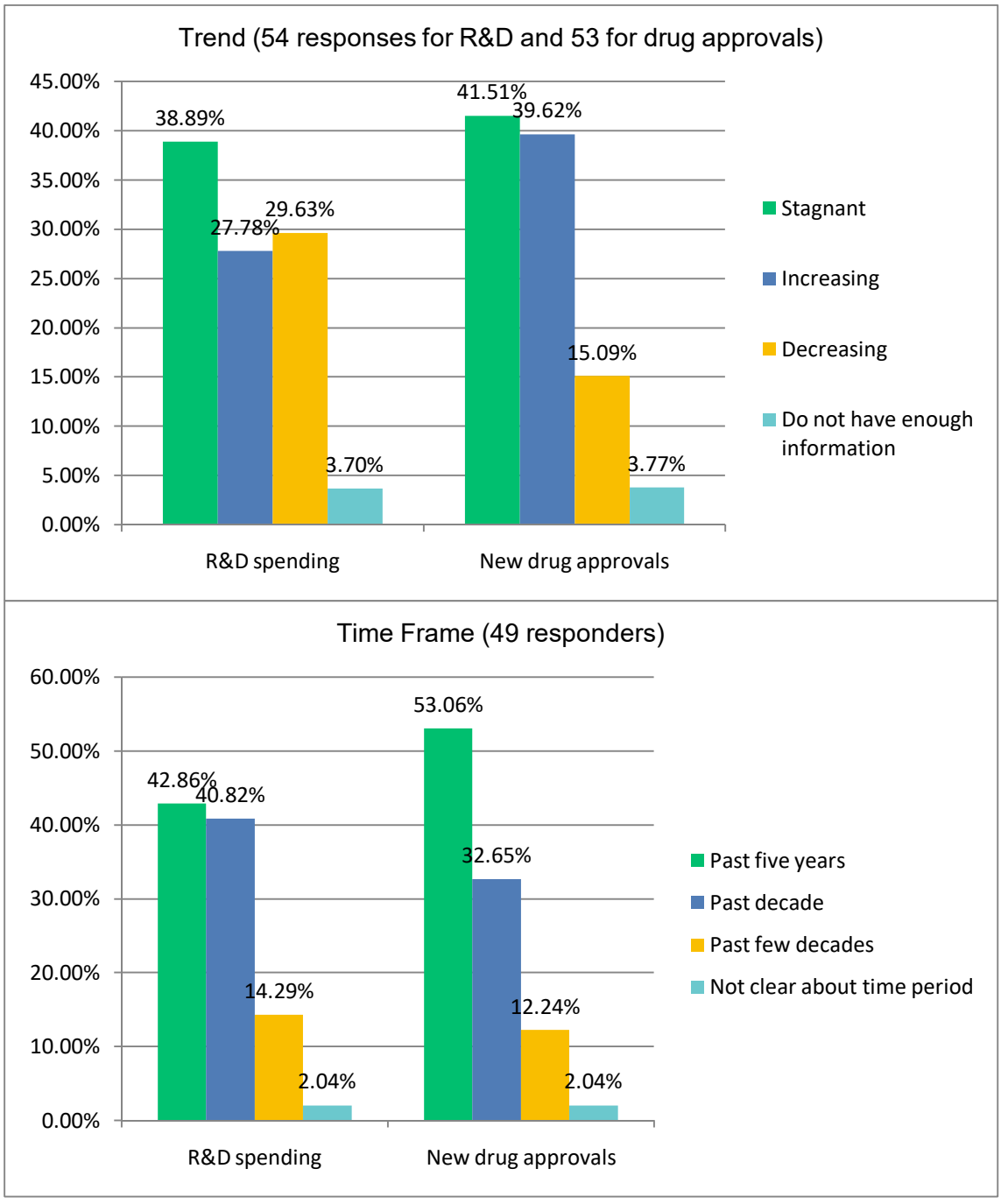

Figure 19 R&D Spending and New Drug Approval Trends

4.8 Drivers of Innovation

This section discusses some important drivers of innovation in the pharmaceutical sector. A close examination of these issues can help identify systemic failures responsible for the R&D productivity decline.

4.8.1 Prioritizing Drivers of Innovation: Survey Insights

Respondents were offered eight drivers of innovation extracted from the literature and asked to choose and rank up to five of them. They could also suggest and rank their own options. 92 respondents completed this question.

Figure 20 depicts the ranking of different drivers of innovation. A useful approach to analyzing the data is to see within each rank what drivers were the top choices. R&D investment, skilled R&D scientists, and basic science respectively got the top three most responses for the top driver of innovation. There is a variety of estimates on the cost of drug development, with some estimate embroiled in controversy and claims of overestimation. For instance, Adams and Brantner (2006) attempted to replicate DiMasi and colleagues ("The Price of Innovation" J. Health Econ. 2003; 22 (2):151-85) estimate of $868 million; however, their estimates came out between $500 million to more than $2 billion, depending on the therapy or the developing firm. They caution against the use of a single estimate by policymakers.

Table 23 depicts the drivers entered by responders. "*Unmet medical need*" was mentioned six times that somehow related to the existence of a market. "Innovation culture" and "easier approval process" are other most frequently cites drivers. Few responders attempted to enter other choices. This may indicate the choices offered to

them were acceptable and comprehensive. Alternatively, it may indicate they were not willing to expend time to proffer their own drivers.

Another important way to analyze the results would be to tally the total responses[59] for each driver hence summarizing overall importance. Figure 21 depicts this attempt. Drivers suggested by the respondents were cleaned and included in the tally. This approach renders "skilled R&D scientists" the top ranked driver followed by "R&D investment" and "good R&D management". These results are intuitive especially the closeness of votes for "R&D investment" and "skilled R&D scientists" is notable and may be indicative of their mutual dependence for synergistic outcome. "Basic science" and "collaborative R&D" rank fourth and filth respectively. Another interesting observation is that the depth of specialized knowledge received more votes than the diversity of knowledge. A note of caution here is that these are all basic ingredients of innovative activities and all should be present in a synergistic way to lead to a positive outcome.

The fact that "skilled R&D scientists" is the top voted driver of innovation is at odds with the dominant paradigm of drug discovery based on automation (e.g. high throughput screening of drug targets) and "basic research–brute force". Such mentality is observable in some industry commentaries. For instance, Andrew Grove, former Intel CEO, draws parallels between e-commerce efficiencies and advocates for "*an "e-trial" system along similar lines*" to speed up clinical trials (Grove, 2011). Others have criticized such parallels because of the "*simplicity and predictability of semiconductor*

[59] Like tallying votes

physics versus "biology's mysteries" (Scannell et al., 2012). These discussions can be symptomatic of transition failure and soft institutional failure.

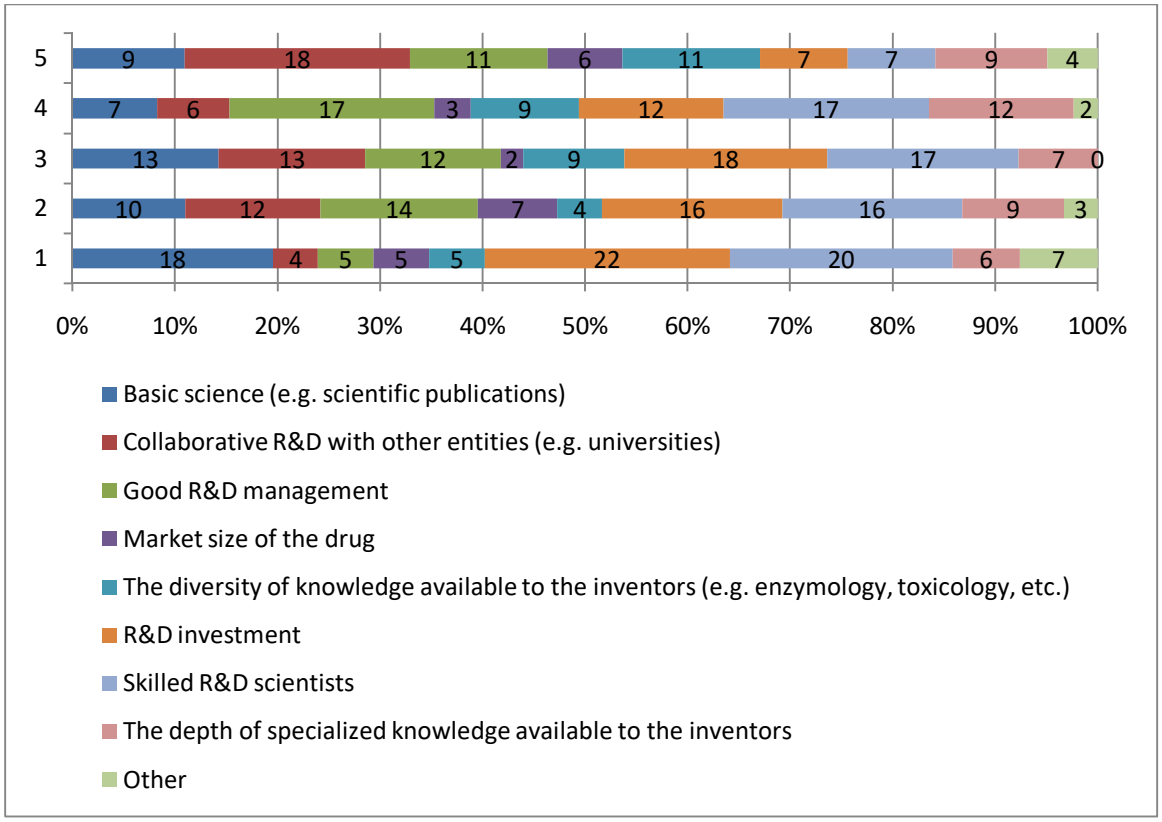

Figure 20 Ranking Drivers of Innovation (Y-axis is ranks; data labels are responses)

Table 23 Responders' Choices for "Other" Category (15 responders)

Driver of Innovation	Frequency	Rank in which mentioned
Unmet Medical Need	6	1, 2
Deep understanding of disease-related biology	1	1
Innovation Culture	3	1
Easier Approval Process	3	2, 4, 0
Patent Extension	1	4
Orphan Diseases	1	5
Manufacturability	1	5
Importance of Indication	1	5
Target Novelty	1	0

* 0 denotes not in and rank and entered as an additional driver; responses have been cleaned and aggregated

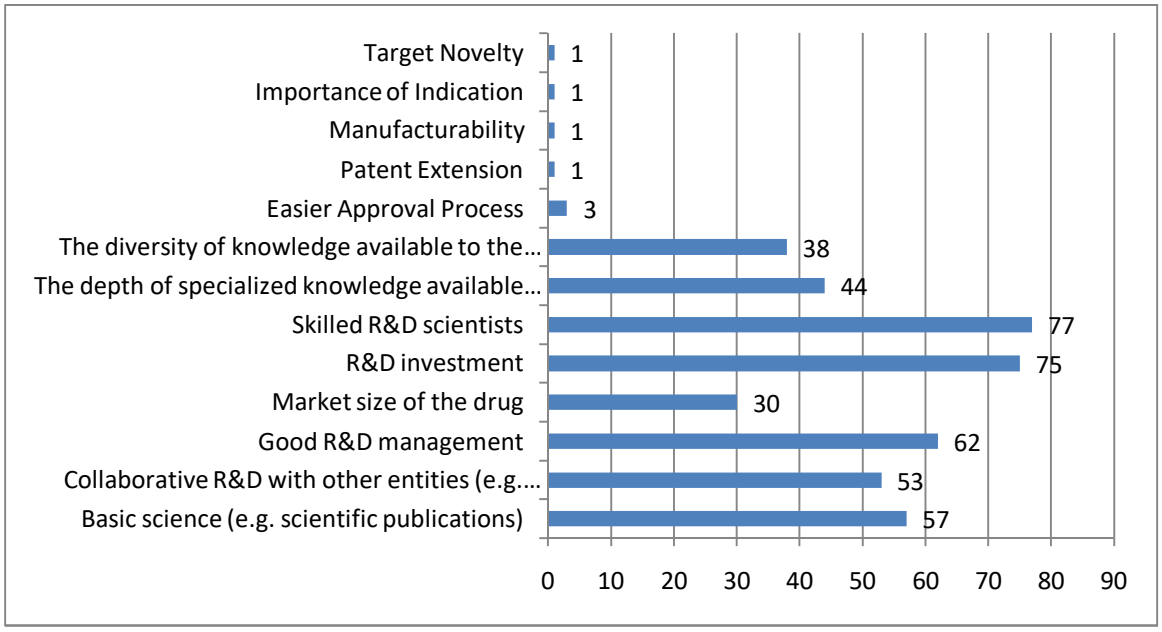

Figure 21 Drivers of Innovation Based on Overall Responses (Total=443)

4.8.2 Resource Mobilization

Innovation requires the synergistic deployment of a number of resources. This section looks at the possible problems emanating from resource availability or utilization.

a) R&D Funding

R&D investment was deemed as the second most important driver of innovation by survey respondents (figure 20) hence merits a closer look. R&D funding can be examined based on corporate, public and venture capital sources. The crux of the Eroom's Law (see chapter one) is that the corporate funding has been on the increase since the 1950s without concomitant increase in new drug approvals. Hence, it can be assumed that the corporate R&D funds are not scarce.

Two survey questions asked the respondents to identify the most important public and private sources of R&D funds. NIH and venture capital were the top ranked choices (see Appendix C tables 1a and 2a for the full list). Consequently, a close examination of these sources is in order.

The knowledge base of the biopharmaceutical companies comes much more from government spending than from business finance (Lazonick and Tulum, 2011). Figure 22 depicts NIH funding for the 1950-2019 period. The year 2003 witnessed a peak in NIH funding as a share of GDP even after adjusting for the Biomedical Research and Development Price Index (BRDPI) inflation. NIH funding grew by 15% between 2003 and 2010 but real funding decreased 1.5% in this period based on inflation adjustment and decreased 11.4% after BRDPI inflation adjustment. As evident, the NIH budget started to stagnate after 2003 and went into decline in 2010. At the same time, the cost of biomedical research has increased rapidly (Boadi, 2014).

Gant application success rate is another important measure to examine. Figure 23 shows the success rate of NIH Research Project Grant (R01)[60] equivalent and research project grant applications. The upshot of the rise in applications and reduced or flat approval rates is reduced resources for experimentation at early stages of technology development.

Figure 24 depicts grants by career stage of applicants. Post-2003, a sharp downward trend for both new and established investigators is observable. However, the

[60] The original and oldest grant mechanism of NIH that funds pojects in line with the NIH mission ("NIH Research Project Grant Program (R01) | Grants.nih.gov." Accessed August 14, 2017 https://grants.nih.gov/grants/funding/r01.htm.).

downward trend is reduced and seems to flatten beginning 2011. While the gap between established and first-time applicant has declined, per figure 25, the "cumulative investigator rate"[61] has also declined from 43% to 31% (Lauer, 2016).

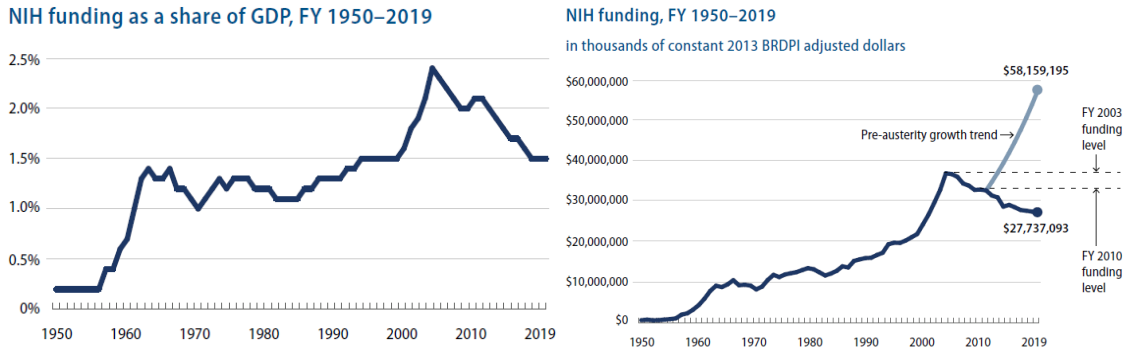

Figure 22 NIH Funding, FY 1950–2019

Source: Boadi (2014); NIH funding figures through FY 2014 are based on total budget authority. Projected NIH funding figures for FY 2015 through FY 2019 are based on data from the Congressional Budget Office.

[61] "*the likelihood that unique investigators are funded over a 5 year window*" (Lauer, 2016).

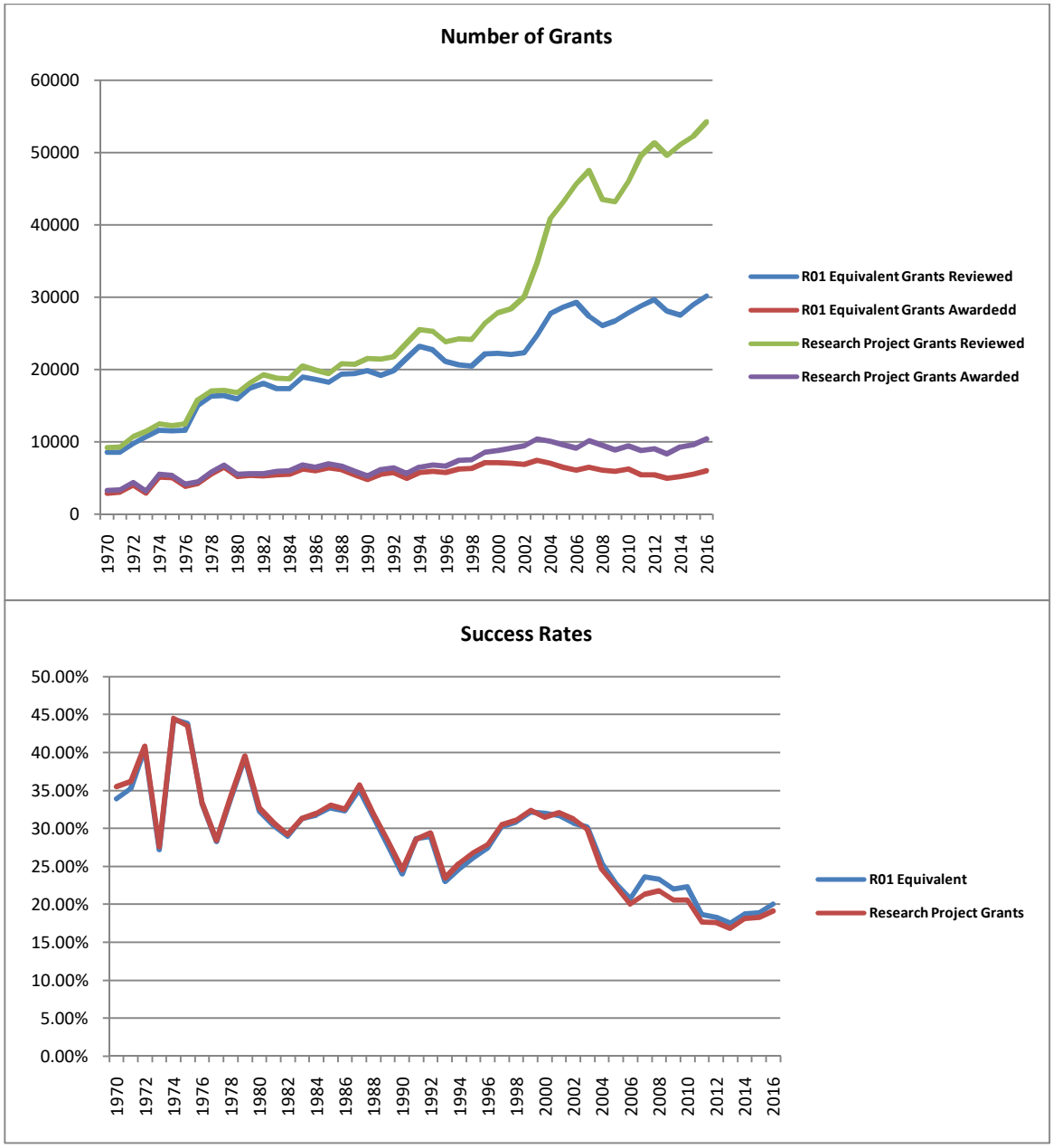

Figure 23 NIH R01 Equivalent and Research Project Grants Applications (Top Panel: Reviewed and Awarded; Bottom Panel: Success Rates)

Source: Extracted from https://report.nih.gov/displayreport.aspx?rid=665

*Excludes awards made with American Recovery and Reinvestment Act (ARRA) funds. **R01-equivalent** awards include R01, R23, R29, R37 and RF1 activity codes. Research projects are defined as activity codes DP1, DP2, DP3, DP4, DP5, P01, P42, PN1, PM1, R00, R01, R03, R15, R21, R22, R23, R29, R33, R34, R35, R36, R37, R50, R55, R56, R61, RC1, RC2, RC3, RC4, RF1, RL1, RL2, RL9, RM1, UA5, UC1, UC2, UC3, UC4, UC7, UF1, UG3, UH2, UH3, UH5, UM1, UM2, U01, U19, and U34.

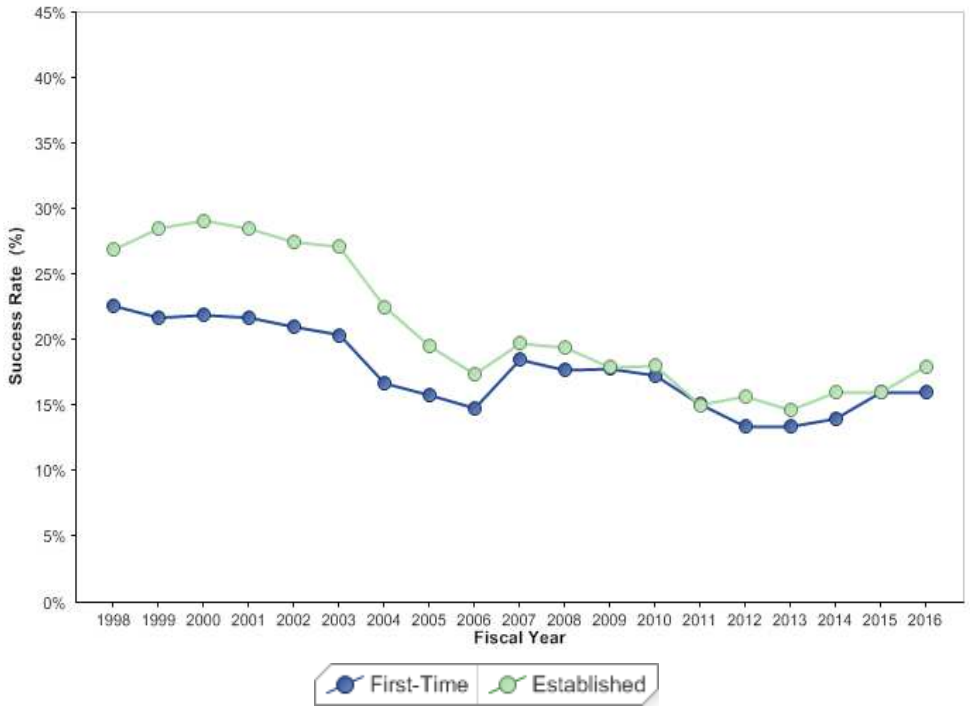

Figure 24 Success Rates of New R01-Equivalent Grants by Career Stage of Investigator

Source: NIH Data Book: https://report.nih.gov/nihdatabook/index.aspx

* New or type 1 application is submitted for funding for the first time.; First-time investigator is a the Contact Principal Investigator who is a first time is a first time investigator.

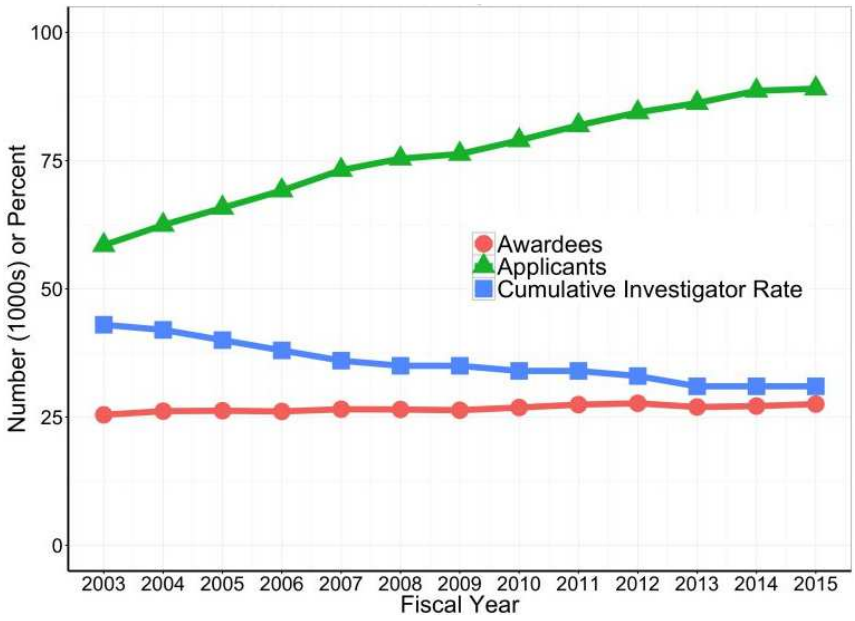

Figure 25 NIH Research Project Grant Applicants, Awardees and Cumulative Investigator Rate

Source: Lauer (2016)

Turning to venture capital, it is worth mentioning that venture capital refers to "*equity or equity-linked investment*" in privately held companies. Funds are raised from other individuals or entities. VCs assume an active role in the investments (e.g. as a director, advisor, or manager of the firm). As a related institution, angel investors are individuals and entities that mainly invest their own funds in new start-up firms (Williams, 2013). Some reports claim "*angel investors are the first stop in a new era of drug development"*. One reason is that big pharmaceutical companies are increasingly

shunning the riskier phases of product development and seek safer bets with products at a more developed stage[62].

Figure 26 depicts venture capital investments in relevant industries reported in the MoneyTree™ healthcare sector. The overall trend of the investments shows a slight uptick in recent years, mostly driven by biotechnology investments. Per figure 27, more investment goes into development rather than discovery. The picture indicates preference for drug development than drug discovery. This might be logical since the perceived drug discovery risk/return profile might be higher in comparison with drug development.

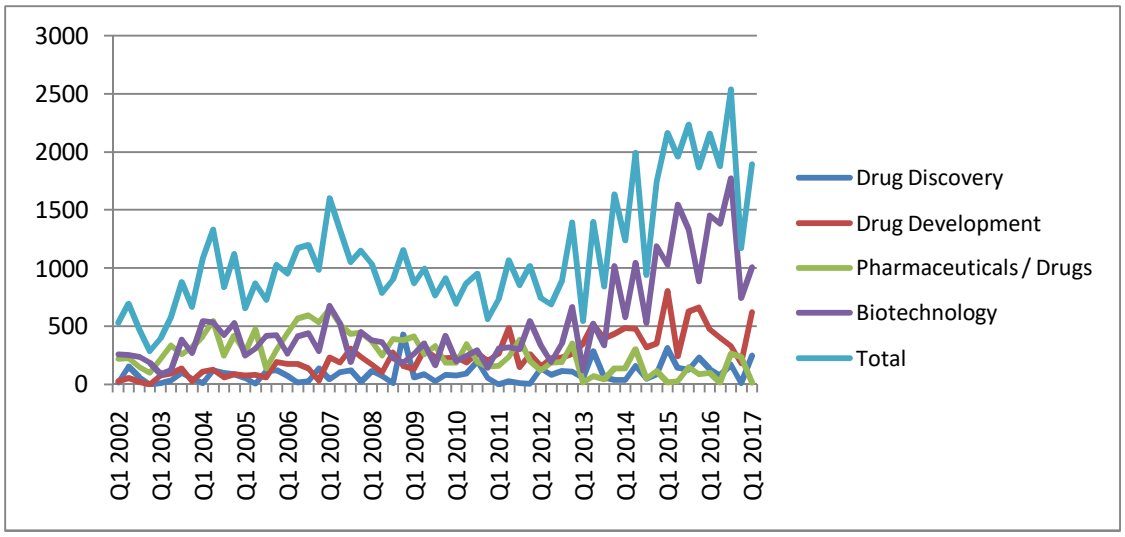

Figure 26 Quarterly Venture Capital Investments Q1 2002-Q1 2017 (in millions; deflated by GDP price index, base year 2009)

Source: Extracted from PwC/CBInsights MoneyTree™ data explorer[63]

[62] "Angel Investors Are the First Stop in a New Era of Drug Development — NewsWorks." Accessed September 14, 2017. http://www.newsworks.org/index.php/local/the-pulse/103720-angel-investors-are-the-first-stop-in-a-new-era-of-drug-development.

[63] Accessed July 1, 2017 from http://www.pwc.com/moneytree

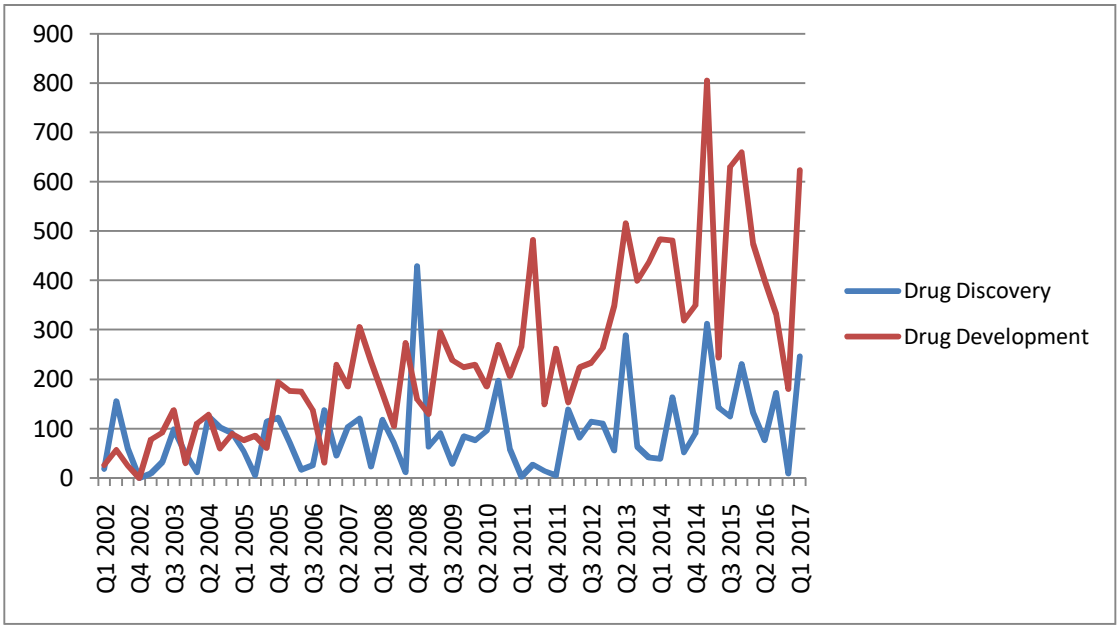

Figure 27 Drug Discovery/Development Quarterly Venture Capital Investments Q1 2002-Q1 2017 (in millions; deflated by GDP price index, base year 2009)

Source: Extracted from PwC/CBInsights MoneyTree™ data explorer[64]

[64] Accessed July 1, 2017 from http://www.pwc.com/moneytree

A brief conclusion of this analysis is that corporate R&D productivity has been on the decline. At the same time, VC investments prefer later stage investment with less risk. Finally, NIH funding which lays the foundation of the biopharmaceutical knowledge base has been on the decline and cannot keep up with rising applications.

b) R&D Employment

Skilled R&D scientists ranked as the top important driver of innovation based on vote tally of the survey (figure 20). Per 2013 figures, the pharmaceuticals and medicines (NAICS 3254) industry ranked second in terms of domestic R&D employment with 117,000 employees (software publishers ranked first). However, in terms of compensation (figure 28) it ranks first by a noticeable margin (Shackelford and Moris, 2016).

In terms of overall employment rate, between 2009 and 2013, the U.S. biopharmaceutical sector eliminated at least 156,000 American jobs. This included cutting R&D departments, reducing sales teams, and eliminating redundancies in post-merger workforces. A few explanations have been offered for this. First, there is the intuitive impact of mergers and acquisitions that lead to duplicate organizational positions (e.g. two marketing managers). The second trend is reduced sales force because of changing customer base. While in the past physicians had a large say in prescriptions, with healthcare reform, payers and government entities hold more sway. Patent expiration and marketing ethics are other issues at play. For instance, GSK was criticized for "*hawking*" the blockbuster asthma medication Advair that could have led to overutilization and deaths. The final trend is preference for pipeline acquisition that

reduces the need for in-house R&D. Many firms are focusing on deals or licensing agreements to exploit existing or development stage drugs.[65]

Presuming these trends hold, R&D employment reduction coupled with "*pipeline acquisition*" rather than internal R&D could potentially explain the flatter R&D spending post-2010. This may superficially tackle the symptoms of the Erooms's Law (i.e. reduced R&D spending coupled with even flat innovative output can improve the look of R&D productivity). However, the underlying feeble R&D productivity of big biopharmaceutical firms will still hold.

[65] "3 Major Trends Driving Layoffs in Biotech and Pharma | BioPharma Dive." Accessed September 6, 2017. http://www.biopharmadive.com/news/3-major-trends-driving-layoffs-in-biotech-and-pharma/399484/.

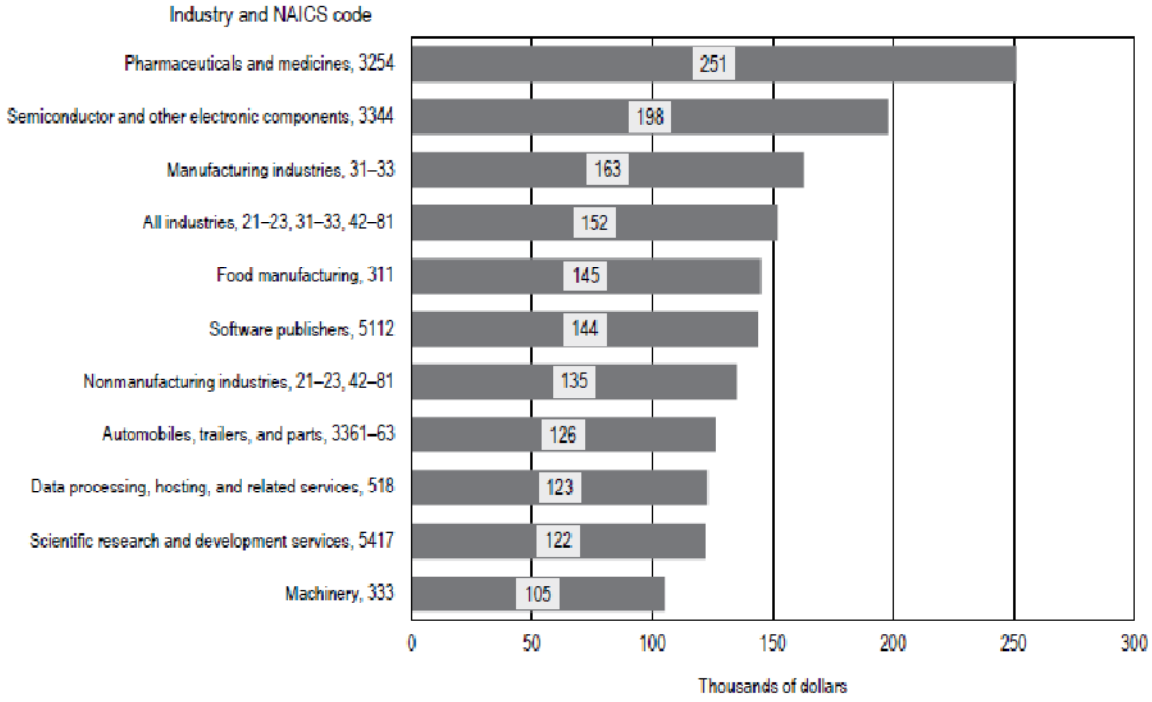

Figure 28 Annual Employee Compensation per Domestic Full-time Equivalent R&D Employee in 2013
Source: Shackelford and Moris (2016)

4.9 Barriers to Innovation

A survey question was devised to construct a list of the most important barriers to innovation. These can help identify systemic issues underlying the Eroom's Law.

54 responders had some form of input for this exercise. Figure 29 depicts the rankings of the most important barriers and figure 30 tallies the votes for each barrier of innovation put to the test. Per figure 29, the top ranked barriers to innovation is availability of good drugs for many diseases. This is related to the "better than the Beatles" problem identified in the literature and restricts the space for innovation. The top

ranked barrier in rank 2 is "complex clinical trials". The literature has documented the increasingly complex, time-consuming and costly clinical trial procedures over the past few decades (see e.g. Scannell et al., 2012). This barrier also appears as the most prominent in rank 3.

In terms of tallies of responses for all barriers, "complex clinical trials" is the top-ranked barrier. Pursuing same drug targets by companies and "*designing drug substances with a single or narrow therapeutic benefits*" are tied for the second highest ranked barrier. The "lack of inter-organizational collaboration in R&D" wins the third highest responses as a barrier to innovation.

Respondents were able to enter their own barriers to innovation. Table 24 depicts their responses. In comparison with the drivers of innovation there are more entries by respondents here. This can reflect the higher complexity of barriers or the responders' that were patient enough to proceed to the last questions, were also willing to devote more time to leaving a thorough response. Entries can be categorized into cost, funding, legal, management, policy, and R&D related issues. A review of noteworthy issues raised is useful. In terms of funding, lack of public funding for early translational research has been raised and the fact that NIH funding is not enough for this purpose. In terms of legal issues, lack of drug pricing regulation is mentioned twice. Regarding overall managerial issues, risk aversion and short-termism in drug discovery and R&D is the single most frequently mentioned theme. It is also detectable in the overall national policy level issues.

In the R&D function category, the issue of "disease etiology" can be detected in three comments. It seems despite advances in the science of human biology, there is still much room for advancement. Finally, the fact that the "reductionist/target-based" is flawed has been raised by two respondents.

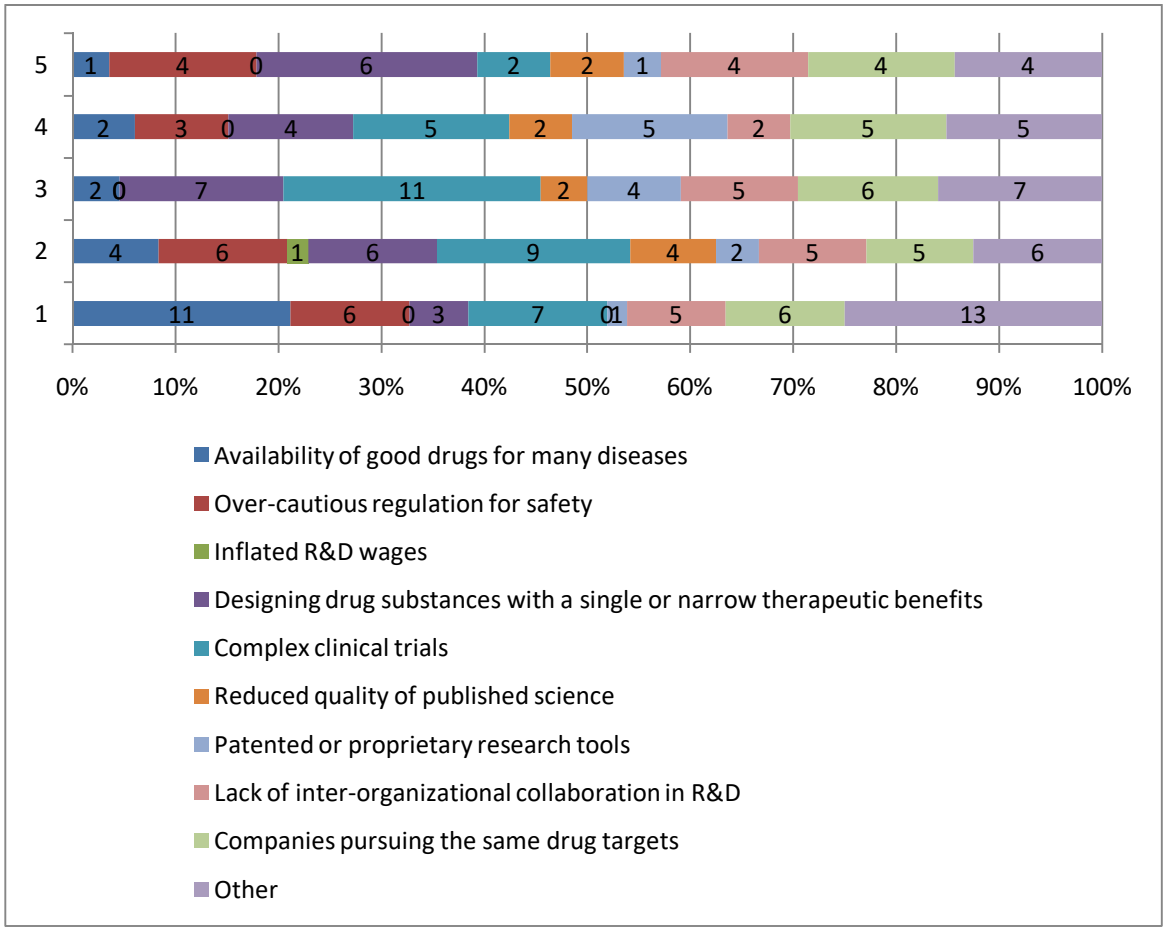

Figure 29 Ranking Barriers to Innovation (Y-axis is ranks; data labels are responses)

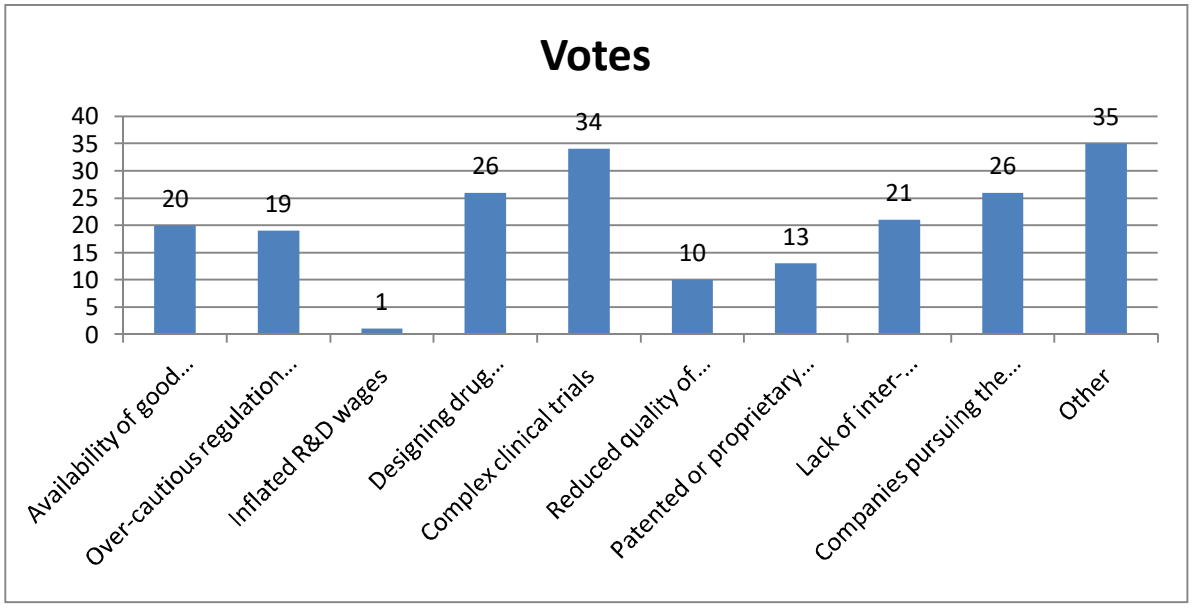

Figure 30 Barriers to Innovation Based on Overall Responses (Total=205)

Table 24 Responder Proposed Barriers to Innovation (29 responders)

Category	Subcategory	Comment	Frequency	Rank	System Failure
Cost	Cost	Huge Cost in Novel Drug Development	1	2	Transition failure/ Hard institutional failure/ Soft institutional failure
Funding	Science/ Translational Research	Lack of Public Funding for Translational Research	1	1	Hard institutional failure
Funding	Shortage	Funding	1	1	Hard institutional failure/ Soft institutional failure
Funding	Short-termism	Investor Short-termism	1	4	Soft institutional failure
Human Resource	Pay/Strategy	Clinical Scientists Displacing Drug Discovery Scientists	1	3	Soft institutional failure
Human Resource	Quality	High Quality Scientist Shortage	1	5	Hard institutional failure
Law	Bureaucracy	Rampant Bureaucracy	1	4	Hard institutional failure
Law	Pricing	Lack of Legal Drug Pricing/ Counter-productive Financial Incentives	2	1, 5	Hard institutional failure
Law	Tort	Legal Environment for Tort Cases	1	3	Hard institutional failure
Management	Feedback	Management Decisions without Scientist Input	1	4	Soft institutional failure
Management	Risk/Short-termism	Risk Aversion/Short-termism in Drug Discovery/R&D Investment/Big Companies/Management	8	1, 2, 4, 3, 1, 4, 3, 4	Soft institutional failure
Management	Strategy	Corporate Lack of Agility/ Commitment	2	3, 5	Soft institutional failure/ Transition failure
Management	Strategy	Poor R&D Leadership in Big Companies	1	1	Soft institutional failure/ Transition failure
Management	Strategy	Excessive Merging	1	4	Hard institutional failure/ Soft institutional failure
Management		Innovation Happens in Cash-strapped Universities and start-up; Big Companies Doctor R&D Funds	1	3	Soft institutional failure/ Hard institutional failure
Policy		Unhealthy Incentives/ Federal Level Bias Against innovative Approaches/Researchers	2	1, 5	Hard institutional failure

Policy/Management	Short-termism	US Economic Model [Short-term Profits?]	1	4	Soft institutional failure/ Hard institutional failure
R&D	Risk	High Cost and High Risk	1	3	Transition failure/ Soft institutional failure/ Hard institutional failure
R&D	Science	Weak Knowledge of Disease Etiology/ Human Biology	3	1, 1, 5	Transition failure
R&D	Science	Target Identification Challenges	1	1	Transition failure
R&D	Science	Unpredictability of Clinical Efficacy	1	1	Transition failure
R&D	Science	Drug Toxicity Test Failure	1	1	Transition failure
R&D	Science	Lack of Pharmacodynamic Biomarkers for Clinical Trial Decision Making	1	2	Transition failure
R&D	Science	Drug Leads Optimization Challenges	1	2	Transition failure
R&D	Science	Drug Candidate Safety Assessment Problems	1	5	Transition failure
R&D	Science/Translational Research	Lack of Translational Models	1	1	Transition failure
R&D	Science; "Better than the Beatles"	Easy Targets Addressed	1	1	Transition failure
R&D	Short-termism	Short required turn-around time for projects	1	1	Soft institutional failure
R&D	Short-termism	Volatile R&D Priorities (caused by management turnover)	1	2	Soft institutional failure
R&D	Short-termism	Incremental Innovation Focus	1	5	Soft institutional failure
R&D	Strategy	Reductionistic/Target-based Medicinal Chemistry Flawed	2	2, 3	Soft institutional failure
R&D	Strategy	Regimentation of Drug R&D (Scientists not Free to Explore New Hypotheses).	1	2	Soft institutional failure
R&D	Strategy	Research vs Clinical Trial Trade-off (Research often stopped for clinical trial leading to low innovation for next trial wave)	1	3	Soft institutional failure
R&D	Strategy	Focus on Process instead of Science	1	4	Soft institutional failure

4.10 Systemic Failures and Eroom's Law

This section will try to seek some explanations for the decline in the R&D productivity based on the survey responses, extant literature and data. The systemic failure framework addressed earlier is an appropriate framing device. Admittedly, given the broad, intertwined issues dealt with here, some issue may straddle multiple categories.

4.10.1 Lock-in or Path Dependency Failures

Lock-in or path dependency failures refer to problems of a complete social system to adapt to new technological paradigms. Lock-in involves a "*complex composition of causes*" including other failures such as network failures and capability failures (Klein Woolthuis et al., 2003). While the aim here is not to consider the whole societal level issues contributing to the drug discovery productivity problem, this label is still appropriate for exploring issues at the supra-firm level.

a) Early-stage Technology Development

Early-stage technology development and experimentation is thwarted because of restricted early stage funding from all major players and other systemic failures. Per figure 31, major funders in the early stages of a venture are government, VC and big pharmaceutical companies. As depicted in the previous section, NIH budget has been on the decline while overall grant applications have been rising. The upshot of this trend is stagnant or reduced success rates in getting NIH grants. This is counterintuitive since flat drug approval rates per Eroom's Law would call for more early stage experimentation. This is especially problematic because VC tends to invest in drug development rather

than discovery. Moreover, the survey identified numerous issues with firm R&D strategy such as short-termism and risk-aversion.

A gap in funding for "very early" translational research was also identified by survey respondents which could be a legitimate policy intervention point for the NIH.

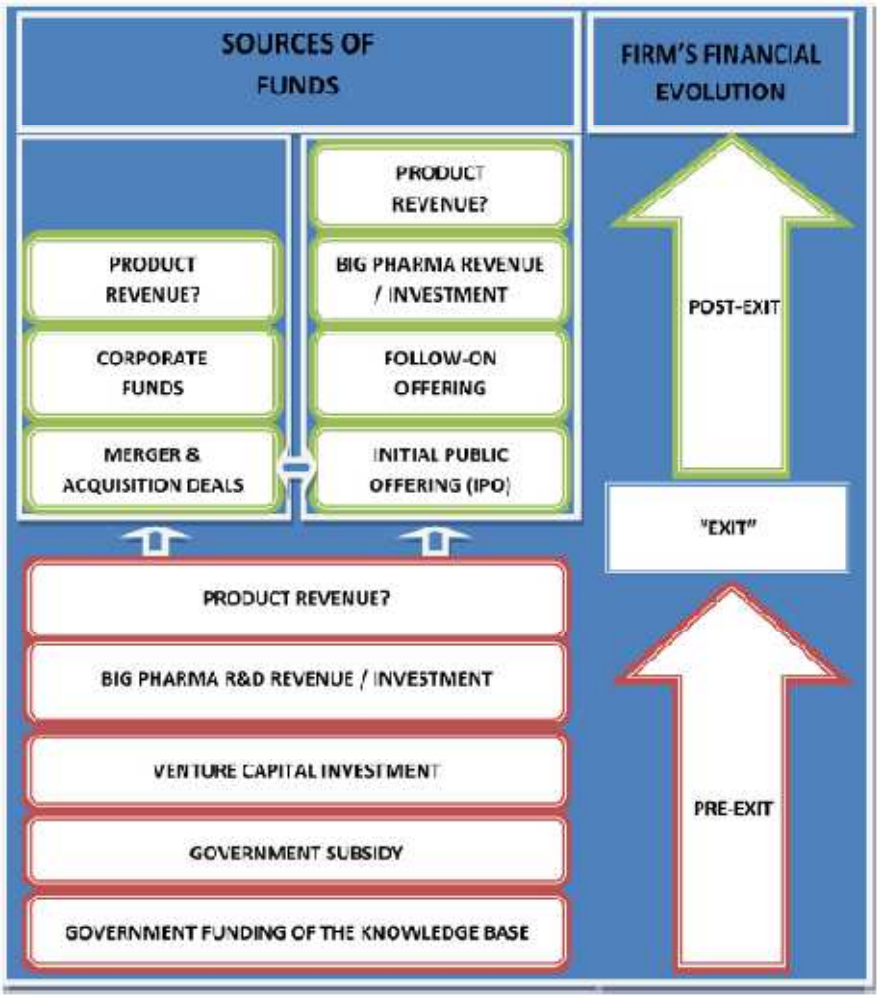

Figure 31 Funding Model of the U.S. Biopharmaceutical Startups
Source: Lazonick and Tulum (2011)

b) Restricted Entrepreneurial Experimentation

An innovation system faces uncertainty regarding technologies, applications and markets. Entrepreneurial experimentation is a source of uncertainty reduction by inquiring into new technologies and applications as a learning process to sift through what works and what does not. Lack of vibrant experimentation will lead to stagnation. The number and variety of experiments can be mapped regarding the number of new entrants and diversifying incumbents, number of different types of applications, and technologies used (Bergek et al., 2008).

The industry is still dominated by firms established before WWII. In contrast to the impact of microelectronics on computing and related industries, radical technological upheavals in the biomedical sector seem to be reinforcing rather than undermining the incumbents (Cockburn and Henderson, 1999).

In the U.S., biotechnology heralded in the first wave of new entrants. The first biotech startup was Genentech established in 1976 and presented the model for other new biotechnology firms (NBFs) (Mckelvey et al., 2004). The majority of the firms could not turn into fully integrated drug producers and lack of complementary competencies, especially in clinical trial, marketing and distribution, and dealing with regulatory agencies has thwarted the growth of these companies (Cockburn and Henderson, 1999; Mckelvey et al., 2004). With few exceptions; e.g. Genentech and Amgen; most of these companies have functioned as a research company, doing contract research for or with established pharmaceutical companies, or supply intermediate products (Mckelvey et al., 2004).

While Disruptive technology in life sciences has breached the barriers to entry, new entrants have had to tune their innovation strategies to the prevailing model run by the MNCs. Regulatory and market barriers to entry have precluded them from competing on their own terms and developing an alternative innovation trajectory (Tait, 2007).

According to Munos (2009) analysis[66], there are more than 4,300 companies active in drug innovation, but only 261 entities (6%) have registered at least one NME since 1950; of which, only 32 (12%) have been in existence for the entire span of 59 years. With 229 (88%) of organizations either failing, being merged, acquired, or getting into M&A deals, there has been substantial turnover in the industry. Of the 261 organizations, only 105 exist today, whereas 137 have disappeared through M&A and 19 were liquidated. The 32 evergreen firms comprises of 23 smaller companies with unique innovative foci. A number of them (Novo Nordisk, Ferring, Grifols, Ucb, Endo and Purdue) are focused on a particular disease area or therapeutic strategy; some are not solely focused on drugs and have moved into products and services too (Bausch and Lomb, and Allergan); some are home-country bound (Takeda, Santen, Eisai, Angelini and Orion) while others are conglomerates (Boehringer–Ingelheim, Solvay, Baxter and Carter–Wallace); and some focus on generics (Teva and Mission Pharmacal) (Munos, 2009).

Another important observation by Munos (2009) is the close correlation between expected NME output, calculated based on a Poison estimation, and the number of companies. He argues, if the NME output of drug companies has been constant (figure

[66] Analyzing data 1950-2008

32), the only way to increase the overall industry output would be to increase the number of companies. The relationship is nonlinear and explains 95% of variation in expected NME output. However, there has been a surge of M&A activity starting more than a decade ago. Hence, so far while new drug discovery technologies or platforms have been developed, the whole system does not appear to be benefiting from them.

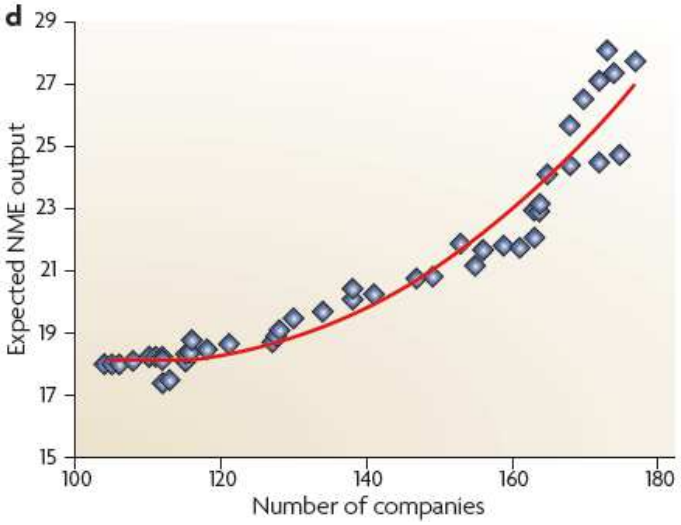

Figure 32 The Dynamics of Innovation
* Red line is the number of companies
* The expected NME output and the number of companies are closely correlated in a nonlinear relationship that explains 95% of the changes in expected NME output by changes in the number of companies.

4.10.2 Hard and Soft Intuitional Failure

Two questions in the survey dealt with enabling and burdensome legislations or regulations (see appendix C for full details). An important observation was that some items were mentioned as both a burden and enabler. Most notable are the FDA core legislations, Hatch-Waxman Act, and patent law. This observation points to the

complexity of regulating this sector, the double-edged nature of some policies, and probably ill-conceived or poorly implemented regulations.

Numerous issues straddle the hard and soft institutional failures. It means both laws on the book and social and political culture impact them. For instance, abuse of policy incentives is due to poor conception/implementation but also predatory culture of policy targets. A number of them are explored here.

a) "Pisano Puzzle"

In addition to the Eroom's law, the biopharmaceutical (BP) sector is beset by another idiosyncrasy, i.e. the "Pisano puzzle". The "Pisano puzzle" refers to the US biotech boom in the pre-2008 economic crisis despite 10–20 year time-frame for product development and the lack of profitability of the industry as a whole. Lazonick and Tulum (2011) justify the term "Pisano puzzle" by posing the question of "why would Money from venture capitalists and big pharma flow into an industry in which profits are so hard to come by?" They conclude that the U.S. biopharmaceutical business model relies on a knowledge base funded by the NIH. Moreover, various complementary government subsidies especially the Orphan Drug Act (ODA) provisions are also drawn upon. M&A and IPOs[67] allow venture capitalists and established firms that invest in BP startups to extracts returns on their investments long before the commercialization of a drug. In many cases the startup never develops a commercial drug. The existence of a speculative stock market allows for IPOs of startups.

[67] Initial Public Offering (of the company stock)

At the expense of R&D, stock buybacks and dividends have been used by established corporations to distribute cash to shareholders. Buyback of stock raises stock prices with the prime beneficiary being corporate executives that gain from exercising their stock options (Lazonick and Tulum, 2011).

b) Abuse of Policy Incentives

Policy incentives are in place to induce innovation in certain areas but they have been open to abuse.

The Orphan Drug Act (ODA) of 1983 offered incentives for the development of treatments for particular types of diseases (i.e. conditions with small market size and those with a prevalence of less than 200,000) through clinical trial support (grants and contracts), tax credit of 50% for clinical testing costs, and a seven-year exclusive right to market the orphan drug for the approved use. (Kesselheim, 2011). Examining 2002 orphan drug designations with 352 FDA-approvals Wellman-Labadie and Zhou (2010) report "*commercial and ethical abuses*" with 9% of orphan drugs having blockbuster[68] status, at least 14 discontinued products recycled as orphan drugs, 32% of orphan designations relating to cancer indicating focus on lucrative niches rather than unaddressed or under-addressed diseases, and multiple orphan designations for the same active agent.

Avery (2008) has examined a couple of loopholes in the original Orphan Drug Act. "*Reverse payments*" refers to the situation where the patent holder would pay the challenger for a settlement that would keep the challenger out of the market. Another

[68] Popular drug with at least $1 billion annual sales

tactic has been to launch "*authorized generics*" in anticipation of generic entry into the market. An authorized generic drug is made by the original drug holder with the distribution outsourced to a licensee with own packaging and FDA identification number; or basically the branded drug marketed as generic at lower price. Even the reform provisions in the Medicare Modernization Act of 2003 failed to stem these practices (Avery, 2008).

Pediatric patients respond to drugs differently than adults; hence, prescription of drugs without clinical trials may be underdosed, ineffective, or even dangerous for pediatric patients. To incentivize pediatric studies, the FDA Modernization Act of 1997 offered six months of market exclusivity time, starting at the end of the drug's patent-protected period. After the enactment, many companies started pediatric trials (Kesselheim, 2011). The Food and Drug Administration Amendments Act (FDAAA) of 2007 renewed and amended the original Act (Rivera and Hartzema 2014). This policy has also been beset by misconduct. Overcompensating manufacturers, getting pediatric exclusivity for popular among adults, methodological flaws and poor quality of trial (Kesselheim, 2011), and blockbuster drugs acquiring pediatric exclusivity (Rivera and Hartzema, 2014) are among the problems.

To recap, in the words of Kesselheim (2011), while these incentives attract interest from drug developers, they are also prone to misuse.

c) Regulatory Capture

The Regulatory capture theory, attributed to George Stigler, refers to a situation in which regulatory agencies become dominated by the industries they were charged with regulating.[69] There is evidence for this phenomenon in the pharmaceutical industry.

In reauthorizing the Prescription Drug User Fee Act (PDUFA) in 1997, the congressional debate was expanded to modernization of the entire FDA. The main mission of the FDA changed from "ensuring that drugs are safe and effective" to include "promoting the public health by promptly and efficiently reviewing clinical research and taking appropriate action ... in a timely manner." To balance the dual and conflicting roles, FDA has been asked to consult a range of interest groups including pharmaceutical companies. Some believe this is an implicit exhortation to cooperate with sponsors for timely approval of drug applications (Zelenay, 2005).

Staying on the PDUFA Act, Moynihan (2002) has tracked the fate of GlaxoSmithKline's alosetron (Lotronex), a drug for irritable bowel syndrome. Once, hailed as a potential blockbuster, it was voluntarily withdrawn in late 2000 in the wake of serious adverse events, including deaths[70]. He observes that the drug has been approved, withdrawn, and approved again. Outside critics claim it is a case of regulatory capture because per PDUFA Act of 1992, companies pay fees for drug approvals. A former insider (Dr Paul Stolley) believes there is detrimental corporate influence in the FDA (Moynihan, 2002).

---

[69] "Regulatory Capture." Accessed October 15, 2017. http://www.investopedia.com/terms/r/regulatory-capture.asp.

[70] This is not meant to track the fate of the drug as it seems it is in use for only women with severe IBS. "*Due to the serious GI adverse reactions associated with this drug, treatment should be restricted to female patients for whom the benefit-to-risk balance is most favorable*" (https://www.drugs.com/lotronex.html).

Regarding orphan drug development, Rzakhanov (2008) observes that orphan drug development tends to be dominated by biotechnology firms (74% of entities) especially larger and more successful biotechnology firms (i.e. with better market value, higher R&D expenditures, and cash reserves).

On the research side, there was controversy surrounding the Tufts Center for the Study of Drug Development (CSDD) estimate of the cost of bringing a drug to market. CSDD assessed the total cost at $2.6 billion by including in the calculations "*an estimate of $1.2 billion in returns that investors forego on that money during the 10-plus years a drug candidate spends in development"*. There was criticism of this serving as a potential excuse for the industry to justify high drug prices.[71]

The ability to perpetuate the rules of the game by old players might be the reason why new drug discovery technologies and system shocks have not imparted change in the tripartite of the innovation system itself, its markets, and the regulatory system to bring about a disruptive change.

4.10.3 Transition Failures

Responders alluded to numerous problems relate to firm level strategy and capabilities (e.g. table 24) in the course of the survey. For instance, managerial decision making without scientists input, poor R&D leadership, and lack of corporate agility and commitment to innovation reflect the broader strategic management issues. Examples of more technical issues are target identification challenges, weak knowledge of disease etiology/ human biology, and toxicity test failures. Problems in these areas will have a

[71] "Tufts Study Finds Big Rise In Cost Of Drug Development | Chemical & Engineering News." Accessed September 18, 2017. http://cen.acs.org/articles/92/web/2014/11/Tufts-Study-Finds-Big-Rise.html.

detrimental effect in the R&D productivity. The issue of current R&D model limitations has also been raised by Munos (2009).

While the aforementioned issue pertain to big established firms, it was already discussed (in the lock-in failure section) that new biotechnology firms could not turn into fully integrated drug producers and lack of complementary competencies, especially in clinical trial and marketing, thwarted their growth (Mckelvey et al., 2004). This effect can be related to the "regulatory capture" discussed earlier. For instance, the rules of the game perpetuated by the established firms might deter the development of certain capabilities by new entrants. For instance, the CEO of Novo Nordisk recently mentioned that the documentation submitted to FDA for two new insulin therapies contained millions of pages that if printed and stacked would be taller than the Empire State Building (Scannell et al., 2012). This particular case illustrates how smaller firms may not have the resources to meet the regulatory burden. On the other hand, this may be another example of regulatory capture: cautious and excessive regulation to deter new entrants without the experience and resources to meet the requirement.

4.11 Recommendations and Generalizability of Survey Responses

Table 25 depicts the responders' comments for "boosting innovation and lifting barriers to innovation". 30 respondents had some input for the question. A few comments are worth mentioning upfront. Firstly, one responder believed the options offered in the survey did not reflect the most important issues. The responder's comments and articulations in other sections of the survey lead to the tentative conclusion that he/she preferred to see more issues about the human side of the innovation and creativity and

broader systemic issues (e.g. “counter-productive incentive” ensuing from the flawed drug payment system). Another responder commented that the definition of innovation for this question was not clear as innovation can be without clear purpose (e.g. academic) or more application-oriented that drives drug discovery. The assumption for the question was dug discovery innovation as emphasized in the consent section and in other questions.

Comments in table 25 have been tentatively assigned labels for ease of analysis. Some may clearly fall between two stools or be multifaceted. The barriers include comments along the lines of funding, governance, market, and pricing. In terms of funding the risk aversion issues is raised not only in VC funding but also in NIH practices. Along the same line, avoidance of early-stage funding has been mentioned as an issue. The need for translational research funding is emphasized. Moreover, problematic review procedures disadvantaging real innovators and new investigators and restricted appeal mechanisms for NIH grant applicants have been emphasized in one comment.

In terms of governance shortcomings, in addition to predatory rent-seeking behavior, the issue of high university overhead and competition from professional grant writing has been raised as constraints for faculty research funding. Moreover, mergers have been hinted as a problem in terms of reducing diversity of approaches being experimented in the market.

In terms of market, it is mentioned that current incentives push the efforts towards orphan drugs to charge higher prices. Finally, comments on pricing seem to be at

different directions one calling for easier payments by insurance companies the other emphasizing the unsustainable and predatory pricing practices.

Turning to drivers of innovation, in terms of financial resources, more government funding for research and incubator projects, and more incentives for drug discovery to companies, are raised. Collaboration along the academia, government, and industry, and public-private lines, and pre-competitive stage has been called for. In terms of human resources, the need for a critical mass of R&D scientists and an organizational culture conducive to innovation are emphasized. Finally, various aspects of the market (monetary value, patient medical need, potential size) are emphasized.

Some comments seem to be raising conflicting issues. For instance, one calls for more exclusivity the other calls for more government role or socialization of the practice. These may ensue for the complex nature of innovation with incentives having both positive and negative consequences. Frequently, the right balance needs to be struck to get the desired results. Moreover, responders depending on their organizational position or experience may advocate for or against an instrument or approach.

Table 25 Comments Regarding Boosting Innovation and Lifting Barriers

Category	Subcategory	Comments	Frequency
Barrier	Funding	Public Funding for Translational Research is Missing	1
Barrier	Funding	Society Should Be Willing to Foot the Cost of Drug Development	1
Barrier	Funding	Relax SEC Rules for Company Funding	1
Barrier	Funding	Big Companies Kill Small Companies with their Finacial Advantage	1
Barrier	Funding	Reinstating NIH Grant Appeal Procedures	1
Barrier	Funding	Open Grants to New Investigators (Increase Transparency in Grant Process)	1
Barrier	Funding	Restoring SBIR Grants and Reducing Delays	1
Barrier	Funding	Risk Aversion in Funding (VC and NIH)/VC Does not Fund Early-stage Biopharma	2
Barrier	Funding	Government Funding For Small Companies (Private Funding of Development Goes after Innovations that Support the Desired ROI)/ (e.g. More SBIR)	2
Barrier	Governance	Focus on Quick Financial Rewards a Real Barrier (focus should be on goals of improving human health and maintain a profit).	1
Barrier	Governance	No Real Commitment to Innovation	1
Barrier	Governance	High University Overhead; Faculty Disadvantaged by Professional Grant Writers	1
Barrier	Governance	Reduce Mergers to Have More Diverse Approaches	1
Barrier	Governance	Displace “Share Price or Share Holder Value” As Top Priority	1
Barrier	Governance	Socialize Medicine	1
Barrier	Governance	Predatory VC after All Value (at the Expense of the Real Innovator)	1
Barrier	Market	Current Regulation Pushes Towards “Orphan Markets To Charge Higher Prices”	1
Barrier	Pricing	Facilitate Drugs Reimbursement by Insurance Companies	1
Barrier	Pricing	US Innovation Based on Predatory, Unsustainable Pricing	1
Driver	Collaboration	More Collaborative Research; Academia, Government, and Industry/Public-Private	3
Driver	Collaboration	R&D and Drug Development Communication	1
Driver	Collaboration	More “Pre-competitive Collaborations” Needed	1
Driver	Funding	More Government Incentives for Drug Discovery to Companies	1
Driver	Funding	Increase Budget for Government Research Funders such as NIH and NSF	1

Driver	Funding	More Funding of Incubators	1
Driver	Funding	More Basic and Applied Science Funding	1
Driver	Governance	Social Pressure on Companies to Boost Domestic R&D Spending	1
Driver	Governance	Innovation Requires the Right Mix of Time, IP, Leadership and Patience	1
Driver	Human Resources	Scientist Ability and Freedom for Innovative Thinking & Cross-fertilization of Ideas	1
Driver	Human Resources	Increase R&D Employees	1
Driver	Market	Better Knowledge of Unmet Patient Needs	1
Driver	Market	Increased Exclusivity for Innovative Therapies (e.g Expanded Hatch-Waxman)	1
Driver	Market	Need for Continual Innovation based on Health Paradigm Shifts and Emerging Infectious Agents	1
Driver	Market	"Potential Market Value Drives" Investments Because Clinical Trials are Expensive	1
Driver	Market	Innovation can be General (e.g. Academic Research) and not Directed to Discovery- Big Companies Need an Economic Reason to Bridge the Gap	1

A final question asked the respondents to assess the generalizability (or applicability) of their responses across other therapeutic areas. The position of the question at the very end was meant to capture their assessment after completing all questions. Figure 33 depicts the responses. Over 90% of responders (49 responders) considered their responses being generalizable to a large extent. Only 4 responders think their responses are generalizable "to some extent". The only negative response came with a comment reminding that their field of work is "*independent of therapeutic areas*". Hence, it can be assumed the negative response is meant to be in the affirmative.

The generalizability of responses may alleviate the need to break down the responses across therapeutic (or maybe institutional) lines. The high generalizability of the responses can be indicative of the success of the survey instrument in capturing general drug discovery innovation drivers and barriers.

From the responder side, given the fact that responders mostly report more than one specialization, they may have knowledge of the common drivers and barriers of innovation across therapeutic areas.

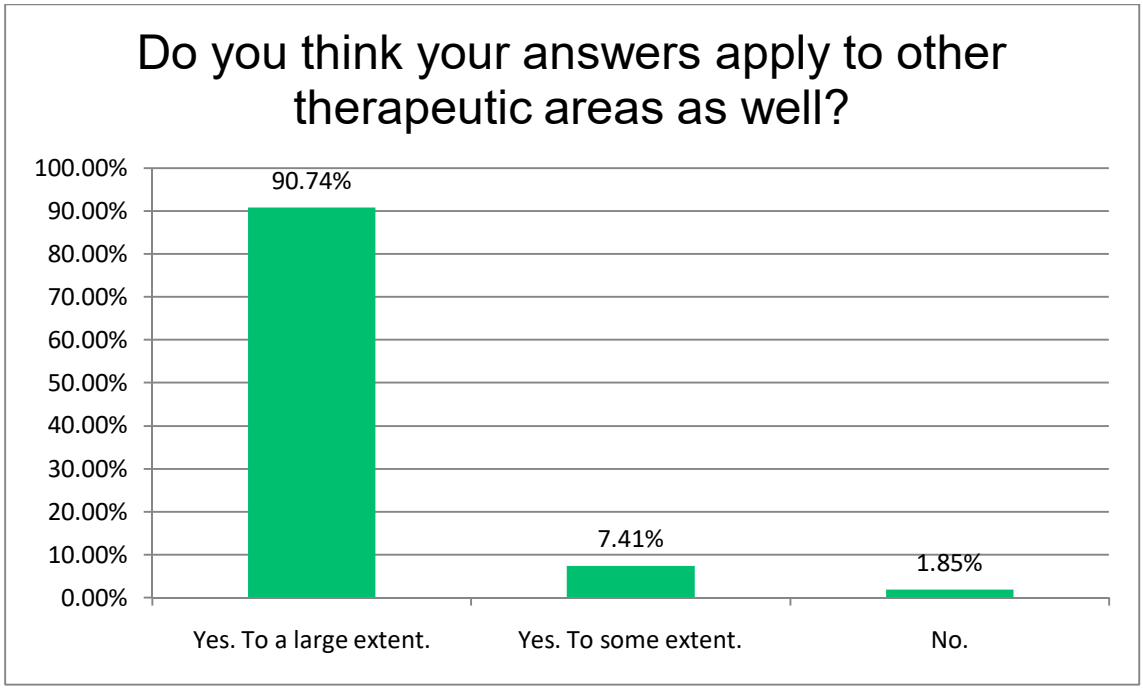

Figure 33 Generalizability of Responses (Total=54)

4.12 Discussion

Table 26 depicts a mapping of selected arguments and observations onto the functions of innovation systems. This brief discussion seeks to trace some higher level systemic failures form the survey findings and extant literature.

The fact that there has been no major change in the major NME applicants leads to "lock-in/path dependency failures". Moreover, since most new biotechnology firms did not develop complementary capabilities (e.g. marketing or regulatory approval skills) and found a niche in the shadows of older pharmaceutical players, can be held as signs of "transition failures". These two failures are closely connected as the former refers to adaptation problem at the broader system level while the latter connotes adaptation shortcoming at the firm level.

Evidence of hard institutional failure is present in the abuse of market exclusivity incentives, artificially high drug prices, and the seeming control of rules of the game by big companies. As a case in point, the legislative failure in the realm of pricing has long been discussed. According to Lazonick and Tulum (2011), in 1990 President George H.W. Bush vetoed a Congressional bill aimed at keeping drug prices down. Recent debates have focused on using Plan D of Medicare to negotiate prices. All along, the counterargument by the pharmaceutical companies has been that regulating prices will cut profits, subsequently reduce R&D resources and diminish innovation.

On the drug price and innovation link, the argument of the big pharmaceutical companies is that without the potential revenues, risking large sums for drug development would not be desirable. A question in response could be how much do big companies contribute to innovation? For example, the risk of developing a promising hepatitis C drug was taken by Pharmasset; however, only after the Gilead's acquisition, pricing strategy changed based on corporate greed.[72] Lack of price control coupled with exclusivity incentives (e.g. Orphan Drugs) and patent protection leads to concentration of competition in certain areas. With the "*winner take all*" phenomenon, secrecy and duplication follows. After the conclusion of the "development race", there is a winner and a number of "*me-too*" drugs as a result of parallel development races[73]. First to reach the market will enjoy a reputation effect, and without alternatives, will enjoy market domination (Petrova, 2014). In a similar vein, Pammolli et al. (2011) believe the R&D

[72] "Yes, We Can Lower Sky-High Drug Prices — Other Countries Have Done It - LA Times." Accessed October 26, 2017. http://www.latimes.com/business/lazarus/la-fi-lazarus-drug-prices-20170725-story.html.
[73] According to DiMasi and Paquette (2004) the majority of "mee-too" drugs have been in clinical development before the approval of the class leading drug.

productivity decline is related to the concentration of R&D investments in areas with unmet therapeutic needs and high risk of failure.

Another evidence of hard institutional failure could be the pursuit of "same drug targets" by companies as a problem. A form of industry-level policy making and prioritization effort could potentially reduce this duplication. Risk-aversion, short-termism, and predatory VC and merger behavior can be symptomatic of soft institutional failure (i.e. political culture and social values). In a similar line, "complex clinical trials" emerged as the top barrier based on tallies of responses. Getting a proper balance between safety regulation and manageable clinical trials can be an important regulatory debate.

Moreover, the emphasis on collaboration along multiple dimensions of public-private and academic-government-industry triple helix can indicate a weak network failures problem. A weak network failure refers to insufficient linkages between companies with complementary capabilities and other knowledge infrastructure such as universities. This situation precludes interactive learning and new idea formation (Klein Woolthuis et al., 2005).

The issue of regulatory capture is also raised in the literature. For instance, on the industry side, Rzakhanov (2008) reports that orphan drug development tends to be dominated by biotechnology firms (74% of entities) especially larger and more successful biotechnology firms (i.e. with better market value, higher R&D expenditures, and cash reserves). On the research side, for instance, there was controversy surrounding the Tufts Center for the Study of Drug Development (CSDD) estimate of the cost of bringing a

drug to market. CSDD assessed the total cost at $2.6 billion by including in the calculations "*an estimate of $1.2 billion in returns that investors forego on that money during the 10-plus years a drug candidate spends in development"*. There was criticism of this serving as a potential excuse for the industry to justify high drug prices.[74] Finally, even starting at the congressional level, some research indicates regulatory capture. For instance, the Prescription Drug User Fee Act (PDUFA) expanded the main mission of the FDA from "ensuring that drugs are safe and effective" to include "promoting the public health by promptly and efficiently reviewing clinical research and taking appropriate action ... in a timely manner." To balance the dual and conflicting roles, FDA has been asked to consult a range of interest groups including pharmaceutical companies. Some believe this is an implicit exhortation to cooperate with sponsors for timely approval of drug applications (Zelenay, 2005). Ability to perpetuate the rules of the game by old players might be the reason why new drug discovery technologies and system shocks have not imparted change in the tripartite of the innovation system itself, its markets, and the regulatory system to bring about a disruptive change.

[74] "Tufts Study Finds Big Rise In Cost Of Drug Development | Chemical & Engineering News." Accessed September 18, 2017. http://cen.acs.org/articles/92/web/2014/11/Tufts-Study-Finds-Big-Rise.html.

Table 26 Assessment of Innovation System Functions

Function	Insights from the Survey	Verdict
Knowledge development and diffusion	• 5 universities were deemed most important source of basic research (Harvard, Stanford, MIT, UCSF, JHU) • Reductionism flawed	• Still top basic research producers are universities • Translational research required • R&D strategy of reductionism and target-based drug discovery flawed but preferred by corporate sector
Influence on the direction of search	• Most ***enabling*** legislations: Bayh–Dole Act; FDA; Hatch-Waxman Act; Patent Law; NIH Funding Bills; Orphan Drug Act; Small Business Innovation Research (SBIR) program • Most ***burdensome*** legislations: FDA; Patent Law; European Laws (related to reimbursement of drug costs; EMA; EUMEA); Hatch-Waxman Act; Medicaid/Medicare; Animal Safety Laws; Clinical Trials for FDA Approval; Stem Cell Research Limitations; Slow funding/Underfunding at/for NIH (one comment for slow funding other for underfunding) • Companies pursuing same drug targets	• Same legislations enabling and hindering (right amount of incentive to the right target is needed) • Enabling legislations are incentives; burdensome legislations are mostly regulatory. • Avenues for new research are restricted (stem cell/clinical trial regulations) • Industry level policies could help cut redundant efforts
Entrepreneurial experimentation	• Start-ups in gene-editing and gene-therapy areas are noteworthy based on respondents • Funding grants not going to new investigator/innovator; no recourse • Early-stage funding shortage • Need to reduce mergers	• New technology may be creating new wave of entrants and therapeutics (gene-editing); like the biotech wave of the past. • No change in major payers and rules of the game and entrants play by old rules of the game
Market formation	• Unmet medical needs (as opposed to monetary size of the market) is also important	• Prices might be artificially inflated
Resource mobilization	• Reduced discovery spending and increased development spending	• VC prefers development stage and biotechnology • Public research spending on the decline since 2003 • A gap in early stage discovery either exists or forming
Development of positive externalities	• Doubt over the *sustainability* of the pharmaceutical business model • More Collaborative Research; Academia, Government, and Industry/Public-Private	• Vicious circle of unsustainable dynamics • Weak network failures (dynamic complementarities' failure)
Institutional issues	• Risk-aversion and short-termism	• Hard institutional failure (formal institutions) • Soft institutional failure (informal institutions)

4.13 Conclusion

This study attempted to ascertain some of the barriers and drivers of innovation in drug discovery that were closely connected to the Eroom's Law. Extant literature, existing data, and a survey specifically designed for the study, were used to construct a profile of the innovation system.

Respondents showed nuanced knowledge of R&D spending and drug approval trends. The overall vote for drug approval was stagnant, in line with the Eroom's Law. Regarding R&D spending, the majority view is that the spending is stagnant. While not in line with the Eroom's Law long term trend, R&D spending figures of the recent years, indeed, confirm this trend.

Based on counting total responses, the top three drivers of innovation are "skilled R&D scientists", "R&D investment", and "good R&D management". Likewise, the top three barriers were deemed to be "complex clinical trials", "designing drug substances with a single or narrow therapeutic benefits" tied with "companies pursuing the same drug targets", and "lack of inter-organizational collaboration in R&D".

The story behind the top drivers is intuitive and clear: to innovate in the science-based drug discovery field, you need skilled scientists, good investment and insightful management. However, the barriers, understandably, may represent a more complex dynamics. Clinical trials represent an important bottleneck in advancing drugs to the market and its size and complexity has grown over the decades. There are a number of reasons for this including the 'better than the Beatles" problem of existing good drugs and the urge to go beyond them, cautions regulation, and even marketing effects (e.g. the drug was tested on thousands of patients) (see Scannell et al., 2012). The second ranked

barriers to discovery can be linked to the new paradigm of molecular reductionism entailing similar corporate libraries of compounds and "brute force screening" methods (see Scnnell et al., 2012). Ironically, if the next barrier of "lack of inter-organizational collaboration in R&D" was to be truly addressed, firstly, more knowledge building blocks from diverse (e.g. academic, non-profit, and corporate; foreign and domestic, etc) organizations could be combined for innovative solutions and, secondly, probably redundant activities would be reduced[75]. The "depth of specialized knowledge" was mentioned more in the pool of responses than the "diversity of knowledge"; hence, indicating traces of the foundational view.

Survey results, the literature, and secondary data point to some systemic failures that can be the basis for policy intervention. There is evidence of "lock-in/path dependency failures" by which the system has not managed to change in all aspects to exploit novel technologies to the fullest extent. New biotechnology firms may have had "transition failures" for not developing competencies for independent existence and competition. Instead, they seem to play second fiddle in the trifecta model. This latter effect may also be linked to both hard (i.e. regulatory) and soft (i.e. cultural) institutional shortcomings that have let regulatory capture and perpetuation of the old rules of the game.

The survey results contain some intuitive and confirmatory results such as the continuing importance of universities in basic science production and the role of NIH as a public source of research funds. However, it also contains some more specific and

[75] Depending on how secrecy for competitive advantage would be mitigated

nuanced findings. For example, the lack of funding for translational research, shortage of early-stage funding, high university overhead, and opaque review procedures at granting agencies can be the seeds of further inquiries.

One of the shortcomings of the study might be the small respondent pool and incomplete responses. However, the study does not fall under the conventional survey research purview as there is no clear population of interest that can be used as a sampling frame. In fact it can be argued that the research only involved innovators and no attempt was made to engage a vast array of players and beneficiaries of the innovation system such as practitioners, patients, policymakers, and even citizens not currently on medication (as both future beneficiary of the system and funder of the system by tax money).

CHAPTER FIVE: CONCLUSION AND POLICY IMPLICATIONS

New drug discovery plays a key role in the continued health and welfare of the population. Moreover, a whole industry has been build around drug discovery and development. However, the historical trend of drug discovery output of the industry has been stagnant despite increasing resources allocated to research and development. This has been dubbed the "Eroom's Law", the Moore's law revered, given its paradoxical nature.

Several arguments have been set forth as possible explanations[76]. First, there is the issue of increasing stock of available safe and effective drugs for many ailments. This makes opportunities for innovation scarce. Second, drugs are highly regulated for safety and effectiveness reasons. Over years, the regulatory retched has been gearing up in response to adverse effects or higher expectations of safety and effectiveness. The increasing stock of therapies is also contributing to tighter regulation. A new therapy is held against higher standards when therapies exist while the regulator might be willing to be easy on a therapy for an untreated disease.

Third, patent protection is effective in the pharmaceutical sector. Moreover, there are a number of market exclusivity incentives in place for development of certain drugs. These issues, coupled with a lack of effective drug price regulation, have triggered a race

[76] Mostly drawing on Scannell et al. (2012)

for drug development where the winner reaps big benefits and being a runner-up is not rewarding. The "winner takes all" mentality has led to massive investments in R&D in the hopes of being the winner.

Finally, there has been a move towards mechanistic rationalization of the discovery process as a backlash to the previous era's apparent serendipitous discovery processes. Molecular reductionism has tried to displace the human judgment form the early phases of discovery and has not improved the overall discovery process.

With the aforementioned background, the present study attempted to uncover the drivers and barriers of drug discovery. Given the complexity of the issue, insights form three levels of analyses were developed. First, firm-level drivers of innovative output were explored based on patent data. Second, breakthrough innovations were examined for inklings into their characteristics that could help adjust R&D processes for better performance. Finally, the analysis was broadened to the whole innovation system level by means of an expert opinion survey.

Firms are the main innovative units of the innovation system. Understanding the drivers of innovation in these units can help identify possible causes of the Eroom's Law. A panel dataset was built by starting form patents assigned to the pharmaceutical international patent class (A61K). An observation was that about one quarter of firms contributing to the pool of pharmaceutical inventions come from outside the pharmaceutical industry.

Drug discovery has been a science-based endeavor since inception of the modern day industry. However, the scientific and technological advances of recent decades have

not changed the innovative output of the industry. Motivated by these observations, the use of codified science in inventive out of firms was examined. The results indicate a curvilinear first negative and then positive; relationship between reliance on codified scientific knowledge (enshrined in patent references) and inventive output. This may indicate that firms needs to assimilate scientific knowledge and spread their use over a number of inventions rather than regurgitate the scientific knowledge in a small set of inventions.

Another foray is the knowledge recombination profile of firms. Knowledge diversity positively impacts quality-weighted inventive output while knowledge heterogeneity (i.e. based on distance or differentness from own inventions) negatively influences both inventive output and quality-weighted inventive output. These results are in line with the foundational view of creativity that contends a deep grasp of domain knowledge is important in developing breakthrough innovation.

Certain groups of firms may display systematic differences form their peers. Hence, nuanced subsample analysis revealed that for firms primarily active in the pharmaceutical sector higher knowledge diversity in inventive output does not improve performance. This may indicate that such a capability is pervasive among the firms hence does not differentiate the inventive output.

Moreover, the influence of a driver of innovation may change over time. Before the year 1995, reliance on codified science boosted the inventive output in comparison with post 1995. This can also be in line with science intensity of patents becoming pervasiveness hence having reduced differentiating capacity.

Turning to breakthrough innovation study, compared with the dominant literature, a different approach was used by defining breakthrough in terms of inclusion of a patent in the Orange Book. A first observation is that about 28% of patents listed in the Orange Book are not assigned to the pharmaceutical IPC patent class (A61K). This may indicate the spillover effect of invention in cognate areas, the generic properties of some chemical inventions, or at a mundane level, the fuzzy nature of chemical patent classifications. Moreover, almost half (48.5%) of the Orange Book listed patents have a family member in the Orange book, indicating that a smaller pool of inventions underpin the Orange Book patent set.

Results of the analysis indicate that inventions with fewer applications (i.e. based on generality index) have higher probability of entering the Orange Book. Broader legal scope boosts the probability of being an Orange Book patent. This might be because inventors carve out larger legal protection for more valuable inventions. In terms of knowledge recombination, higher technological knowledge diversity reduces the probability of being an Orange Book patent while technological knowledge distance (i.e. not coming from the patent's technology class) increases the probability of being an Orange Book patent.

A tentative look at the effect of co-assigned patent reveals that being assigned to multiple entities does not have a significant impact on the probability of being an Orange Book patent while being assigned to different organizational types reduces the probability of being an Orange Book patent.

A subsample analyses revealed differences between drug substance patents and Orange Book patents. Broader invention applications (i.e. generality) boosts the probability of being a drug substance patent. This may indicate active chemical substances have broader applications than other inventions. Similarly, broader "envisioned" applicability (i.e. patent scope) of an invention boosts the probability of being a drug substance patent. In terms of knowledge recombination profile, higher technological knowledge diversity (i.e. originality index) increases the probability of being a drug substance patent. However, higher technological knowledge distance reduces the probability of being a drug substance patent.

The combined observations regarding knowledge recombination indicate that when focusing on the Orange Book patents, the tension view on creativity is more relevant while the drug substance patents within the Orange Book show traces of the foundational view.

While insights from firm inventive output and breakthrough innovations are useful, tackling the Eroom's Law needs a boarder systemic perspective and input from experts. Expert opinion was garnered by means of an online survey. Regarding Eroom's Law, the majority of respondents believed new drug approval rate was stagnant (41.5%), a close proportion (39.6%) of them implied there was an uptick in recent years. In terms of R&D spending, the majority believed it is stagnant. This is in line with R&D spending figures of recent years.

Based on counting total responses of respondents, the top three drivers of innovation were "skilled R&D scientists", "R&D investment", and "good R&D

management". Likewise, the top three barriers were deemed to be "complex clinical trials", "designing drug substances with a single or narrow therapeutic benefits" tied with "companies pursuing the same drug targets" for the second place, and "lack of inter-organizational collaboration in R&D" ranked third. The top ranked items in this exercise point to the shortcomings of molecular reductionism paradigm of drug discovery. To begin with, automation and high throughput screening techniques attempt to reduce the human judgment in drug discovery. But respondents have ranked skilled scientists as the top driver of innovation. Likewise, the barriers of "designing drug substances with a single or narrow therapeutic benefits" and "pursuing the same drug targets" are the symptoms of the same molecular reductionist paradigm. Therefore, the R&D model of the industry reduces the role of the top driver of innovation and accentuates (or creates) a number of barriers to innovation.

The barriers can be linked to several systemic level failures such as "lock-in/path dependency failures", "transition failures", and both hard (i.e. regulatory) and soft (i.e. cultural) institutional failures. Moreover, there are traces of "regulatory capture". These systemic failures might be the reason why previous policy interventions have failed to meaningfully change the Eroom's Law. Likewise, they can also indicate new directions for policy intervention. Table 27 depicts the systemic failures and some examples.

Table 28 depicts the insights form the study regarding the impact of knowledge recombination and science use on innovative output. Results of the survey are particularly interesting in that experts ranked basic science, depth of knowledge and diversity of knowledge in the respective order as drivers of innovation. While matching

the broad theoretical view is tentative, it is interesting that expert opinion and firm panel data results match the foundational view. Moreover, the Orange Book patent profiles are more in line with the tension view, which is consistent with the classic breakthrough conception of resting on radical (i.e. different form the norm) ideas.

The fact that policy interventions have not significantly impacted the drug discovery performance has another set of policy formulation implications. To begin with, any initiative aimed at the innovative output of the industry should comprise of a mix of synergistic policies. Moreover, the effectiveness of each policy, and importantly the policy mix, may change over time hence there is a need for policy adjustments. For instance, per one survey respondent comment, there is a gap in translational research funding. If this is indeed the case, an effective policy apparatus should identify and address this gap. Finally, an independent policy apparatus is required to stave off regulatory capture.

A few points on the shortcomings are also in order. The patent indicators may not always quite fit and characterize the complex phenomenon (e.g. knowledge diversity in innovation) under study. Especially, when moving up the conceptualization ladder from the invention to the firm level, this issue is more apparent. However, the final essay brings in the voices of experts to remedy this situation. Perhaps, a bigger proviso is that the study was formulated from the perspective of the innovator. However, the drug discovery innovation has more actors and interest groups such as patients, physicians, policy makers, and any member of the community as a potential beneficiary or

contributor (e.g. by tax) at some point in life. Such a broad take on innovation was beyond the scope of the present study.

Table 27 Systemic Failures and Examples

Systemic failure	Examples
Lock in path dependence	- Restricted funding for early-stage technology development: Reduced NIH funding; VC, Big pharma risk aversion and short-termism; gap in translational research - Restricted entrepreneurial experimentation: Industry/rules of game dominated by established firms (barriers to entry); surge of M&A activity
Hard and Soft Intuitional Failure	- "Pisano Puzzle" (speculative finance in biotech) - Abuse of policy incentives (e.g. pediatric exclusivity for popular among adults - Regulatory capture: Prescription Drug User Fee Act (PDUFA); GSK's alosetron (Lotronex), a drug for irritable bowel syndrome, has been approved, withdrawn, and approved again.
Transition Failures	- Technical issues (e.g. target identification) - Managerial issues (e.g. poor R&D leadership) - Lack of capability development for biotech firms

Table 28 Knowledge and Science Recombination Impact on Innovative Output

	Panel of Firms	Breakthrough Innovations	Survey
Science Intensity	Curvilinear (negative then positive)	No impact	Basic science ranks higher than depth and diversity of knowledge
Knowledge Diversity	Positive for quality of innovations	Negative	Depth of knowledge ranks higher than knowledge diversity and lower than basic science
Knowledge Distance/Heterogeneity	Negative for quantity and quality of innovations	Positive	Knowledge diversity ranks third.
Theoretical Explanation	Foundational view	Tension view	Foundational view

* Survey results are based on total responses for main drivers of innovation put to the test

APPENDIX A – CHAPTER TWO SUPPLEMENTAL INFORMATION

a) The Trifecta Model

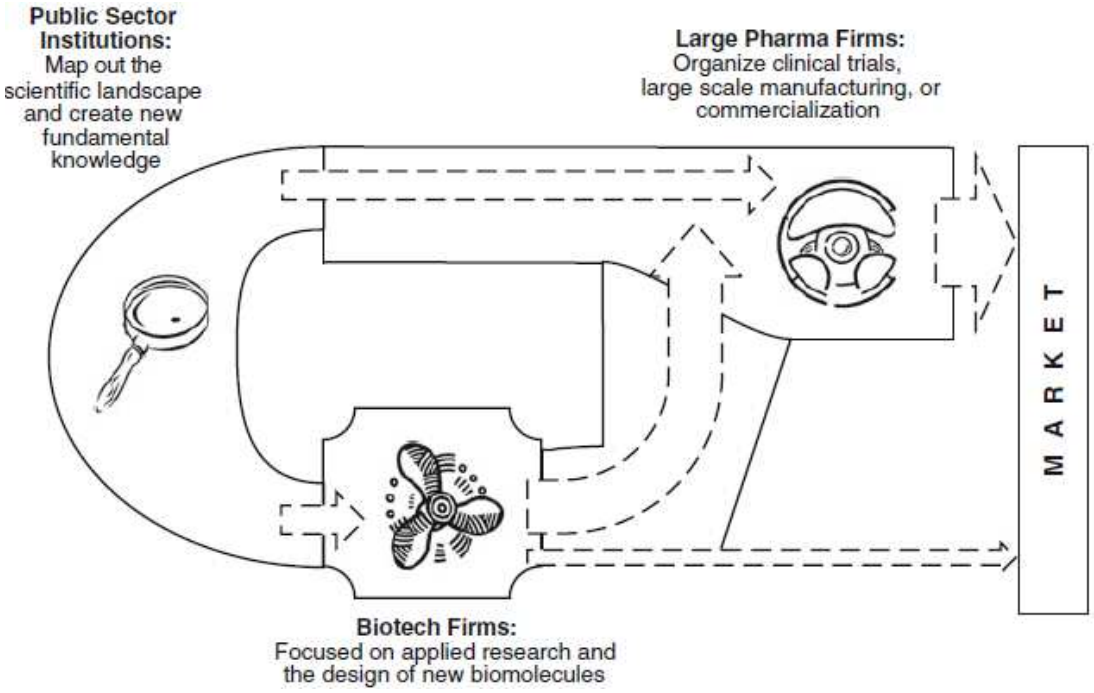

Figure 1a The Trifecta Model of Innovation

Source: Petrova (2014)

b) Data Construction Details

There are several patent datasets available, each with its own merits and drawbacks. The NBER patent data project was developed over a decade, involving multiple researchers, institutions, and financial resources and was meant to be made

accessible to researchers. It includes detailed information on approximately 3 million U.S. patents granted between January 1963 and December 1999. Moreover, it has a reasonably broad match of patents to Compustat (i.e. the financial data of all firms traded in the U.S. stock market) (Hall et al., 2001). Later, the dataset was expanded to include patents up to 2006.[77] The biggest advantages of the NBER dataset is matching patents to firm level data in the Compustat database. According to Golec and Vernon (2008) the Compustat-based data series are more reliable and comprehensive than the alternative data series (i.e. PhRMA and NSF data) for the pharmaceutical industry. However, a major drawback of the NBER dataset is the fact that it is not current and stopped in 2006.

OECD Patent Datasets is another data source that contains patent quality indicators and patent families. It is more recent than the NBER, contains various patent level indicators including the originality and radicalness indices, and is being updated continually. The February 2016 release of the dataset is at hand and will be used in this study. The OECD Patent Datasets (i.e. OECD Patent Quality Indicators database, February 2016" and "OECD Triadic Patent Families database, February 2016") will be used for patent data and indicators. The Compustat database (by the Wharton Research Data Services: WRDS) will be used for firm level financial data. Other datasets have been used to complement certain aspects of the data and will be introduced as they are drawn upon.

The construction of the dataset began form the OECD USPTO patent quality file. The file contains both granted patent and patent application data (total rows of

[77] "Patent Data Project." Accessed April 11, 2016. https://sites.google.com/site/patentdataproject/Home.

observations 9,229,908). Since some granted patents have their application in the data file, only granted patent data was kept for the study. Another advantage of focusing on granted patents is that a granted patent has passed a certain threshold of invention quality (i.e. USPTO examination). Dropping patent applications yields 5,015,706 rows of data.

The data file was explored for non-unique PATSTAT (i.e. the European Patent Office (EPO) Worldwide Patent Statistical Database) patent application identifiers. There we two patents under certain duplicate application-IDs. The oldest patents seem to have been withdrawn; hence, the latest patents were kept (167 observations were dropped in this process).

To focus on pharmaceutical patents, patents in the technology field 16 were retained (yielding 117,442 observations). Technology field 16 contains patents that have been assigned to the international patent class (IPC) A61K that represent pharmaceutical patents. It excludes patents on cosmetics (Squicciarini et al., 2013; Schmoch, 2008).

To boost the chances of having unique inventions in the dataset, the concept of "*Triadic Patent Families*" is used. A patent family is a set of related patents linked by one or more patents called priority filings that have been filed in several countries (Squicciarini et al., 2013). The priority patent of a patent family is the first patent application filed for the invention of each family (Criscuolo, 2006). Triadic families have been filed in the U.S., Europe and Japan. Hence, the patent sample resulting from the aforementioned steps was merged with the OECD Triadic Patent Families (TPF) database (February 2016 release) to identify patent families. This procedure revealed that 80,683

patents were in both OECD Quality and OECD Triadic Patent Families (TPF) datasets (hence, are Triadic) and 36,759 were non-Triadic and excluded from the dataset.

The first and last members (patents) of each family were tagged. 39500 patents had the same first and last patent (i.e. single USPTO granted patent of the family). In determining the first filing of each family, if there were more than one identical filing years (e.g. two patents with filing year 1999), first granted patent by patent number was tagged. For the last filing of the family, the last granted patent was tagged if there were tied patents on the filing year. The first patent tag will be used to construct the final sample after working on assignees as delineated in the following paragraphs.

The patent assignee files in the OECD Patent Dataset were deemed too unwieldy and non-standardized to work with. Hence, the resulting patent list form the abovementioned procedure was searched in the LENS[78] patent search database. The COMETS[79] (Connecting Outcome Measures in Entrepreneurship, Technology, and Science) patent assignee data files were drawn upon for some comparison and checking of assignees. A good point about the COMETS was that it has an assignee type designation (e.g. firm, academic, etc.) that helps in identifying patents assigned to firms. The problem with the lens "*applicant*" field was that it sometimes lists all "*applicants*" including inventors rather than the "*assignee*" of the patent. As a result, it required cleaning and parsing to separate the assignee. The LENS downloaded filet was cleaned by OpenRefine and manually to yield an acceptable assignee list. The manual phase was tedious and time-consuming sometimes involving Googling names to find out about

[78] "The Lens." Accessed April 11, 2017. https://www.lens.org/lens/.
[79] "COMETS." Accessed April 11, 2017. http://www1.kauffman.org/COMETS/.

name changes. An attempt was made to keep assignees as distinct as possible. For instance, if firm A is independent and then merges into firm B, its patents are not subsumed under firm B. If there is an entry for firm A and B in the Compustat, they both are independent entries. The logic is that a merger creates both qualitative and quantitative change in the firm; hence, the innovative dynamics of firm A and firm B many be different from the dynamics of the resulting merged entity.

The resulting assignee file was merged with the patent file constructed earlier from the OECD patent datasets. 56,804 patents were in both files. The remaining 11,366 were only in master data (e.g. non-firm entities including individuals). The next step was to tag unique assignee names (7118) and (arbitrarily) choose assignees with five or greater years of patenting activity (1170). This was done for the practical purpose of easing the next step of linking assignees to the Compustat database as well as focusing on assignees that have some history of patenting hence more likely to show regularity in innovation for statistical analysis.

After identifying assignees with five or more years of patenting, their name was searched in the Compustat. If there was a hit, the unique Compustat firm key (gvkey) was obtained to extract their available financial data. Figure 2a depicts the numbers of years of data (usable data points) after merging financial data with patent data for the total of 631 firms as depicted. Financial data might not be available for all years; hence, the loss of some data points. Joint assigned patents were assigned to both assignees listed. The complete list of firms is available in Appendix A. The final data set, comprising of

29,554 patents, is unbalanced and firm data points are not necessarily chronologically complete (e.g. 1998, 1999, 2010, 2011, 2012).

Another interesting issue to explore about the panel of firms would be their primary industrial activity. Figure 3a depicts the Global Industry Classification (GIS)[80] of the firms. A few seeming anomalies might be worth mentioning here. The energy sector firm is the French-based "ELF AQUITAINE SA" active in the petroleum and gas sector. The Materials companies include familiar names such as DOW CHEMICAL, DU PONT, and ROHM AND HAAS; as well as other chemical companies. The capital goods group consists of companies such as 3M CO, GENERAL ELECTRIC CO, MITSUI & CO LTD and some holding companies such as the Korean SK HOLDINGS CO LTD. The Commercial & Professional Services firms consist of KOKEN LTD (a Japanese company focused on occupational health and safety products) and NELSON RESEARCH & DEV CO (a small consultancy). The "Consumer Durables & Apparel" group includes SEKISUI CHEMICAL CO LTD (focused on high performance plastics for medical and other uses); UNITIKA LTD (Japanese firm focused on advanced materials used in a variety of industries), and LVMH MOET HENNESSY LOUIS V (also known as LVMH, is a French conglomerate). The Food & Staples Retailing company is ALLIANCE BOOTS PLC, which was a multinational pharmacy-led health and beauty group (Wikipedia). The "Food, Beverage & Tobacco" firms include diverse names such as NESTLE SA/AG, JAPAN TOBACCO INC, and MORINAGA MILK INDUSTRY

[80] The GIS is an industry taxonomy developed in 1999 by MSCI and S&P for the global financial community use. It consists of 11 sectors, 24 industry groups, 68 industries and 157 sub-industries (Wikipedia).

CORP. The “Household & Personal Products” group includes companies such as PROCTER & GAMBLE CO, L'OREAL SA, and NUTRITION 21 INC. The Diversified Financials company is Inspired Capital plc (formerly Renovo Group plc), an SME financial solutions company. Finally, the Technology Hardware & Equipment group includes names such as FUJIFILM HLDGS CORP, EASTMAN KODAK CO, and HITACHI LTD.

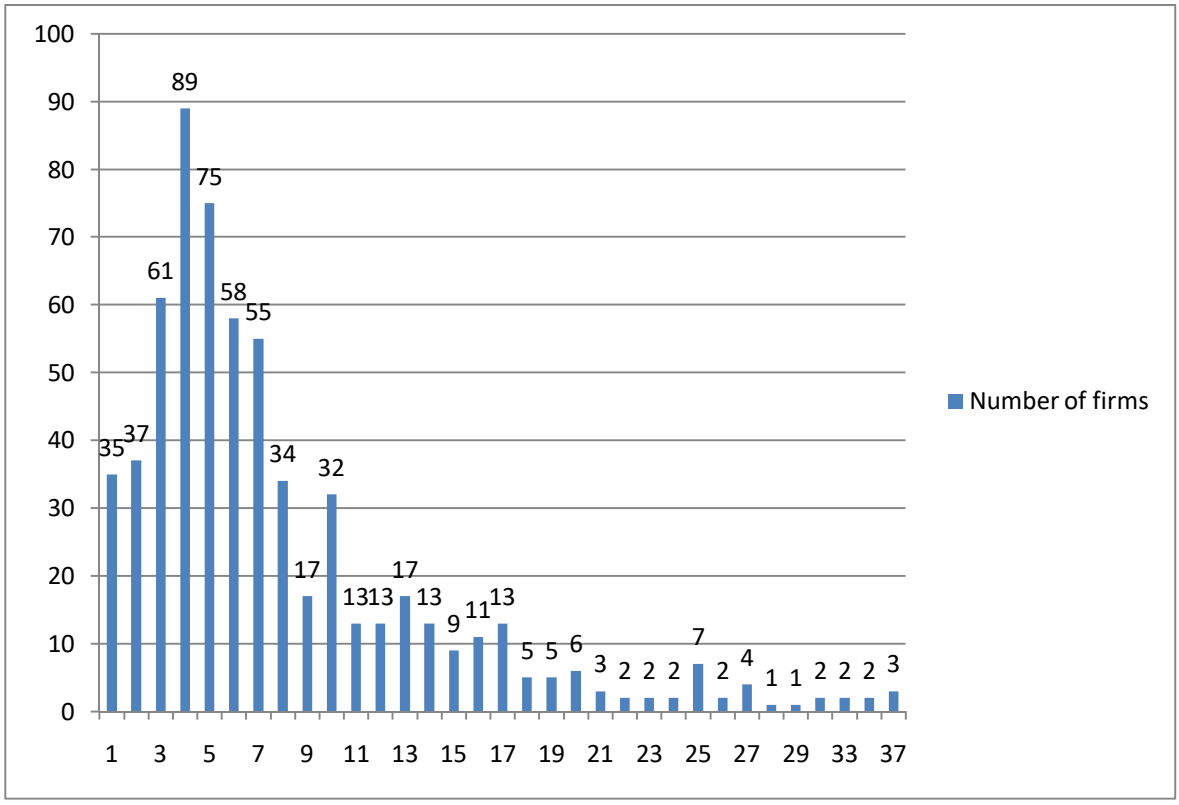

Figure 2a-A Years of Data for Total Firms (total firms=631)

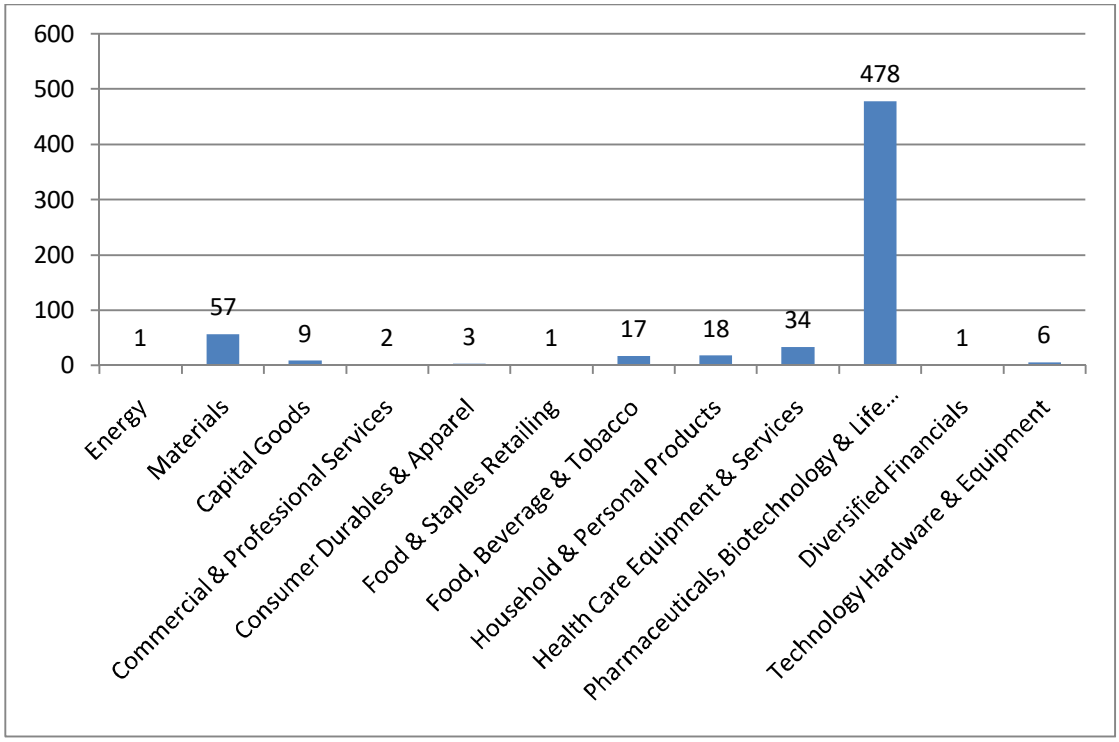

Figure 3a-A Firms by Global Industry Classification (GIC) Groups

The industrial classification of the firms indicates the diversity of firms contributing to the innovation pool related to pharmaceuticals. These firms may not be a member of pharmaceutical industry associations such as PhRMA this may indicate one advantage of the bottom-up sample building (i.e. starting from patent pool rather than a list of companies) used in this study.

The abovementioned steps were taken to ascertain the patent dataset represents valuable and unique (i.e. non-duplicate) inventions. The resulting sample represents one USPTO granted patent per Triadic Patent Family of each firm in the dataset that is listed in the Compustat (both North America and Global lists were searched). In count terms,

the patent count (i.e. firm-year-patent) actually represents the Triadic Patent Family count of the firm per year. Triadic patents are deemed more valuable and one patent per family improves chances of having unique inventions.

c) List of Panel Firms

Firm Name	Years of Data	Beginning Year	End Year	GIC Groups	GIC Industries	GIC Sectors	GIC Sub-Industries	Country HQ
JOHNSON & JOHNSON	37	1978	2014	3520	352020	35	35202010	USA
PFIZER INC	37	1977	2013	3520	352020	35	35202010	USA
MERCK & CO	37	1977	2013	3520	352020	35	35202010	USA
PROCTER & GAMBLE CO	34	1978	2012	3030	303010	30	30301010	USA
3M CO	34	1977	2012	2010	201050	20	20105010	USA
BRISTOL-MYERS SQUIBB CO	33	1981	2013	3520	352020	35	35202010	USA
GLAXOSMITHKLINE PLC	33	1980	2014	3520	352020	35	35202010	GBR
SCHERING-PLOUGH	31	1978	2008	3520	352020	35	35202010	USA
BAXTER INTERNATIONAL INC	31	1980	2014	3510	351010	35	35101010	USA
ABBOTT LABORATORIES	29	1980	2012	3510	351010	35	35101010	USA
WYETH	28	1978	2008	3520	352020	35	35202010	USA
KYOWA HAKKO KIRIN CO LTD	27	1987	2013	3520	352020	35	35202010	JPN
DOW CHEMICAL	27	1980	2012	1510	151010	15	15101020	USA
NESTLE SA/AG	27	1983	2014	3020	302020	30	30202030	CHE
TAKEDA PHARMACEUTICAL CO	27	1987	2013	3520	352020	35	35202010	JPN
SHIONOGI & CO LTD	26	1987	2012	3520	352020	35	35202010	JPN
TEVA PHARMACEUTICALS	26	1986	2012	3520	352020	35	35202010	ISR
ALLERGAN INC	25	1988	2012	3520	352020	35	35202010	USA
SUMITOMO DAINIPPON PHARMA CO	25	1987	2012	3520	352020	35	35202010	JPN
AJINOMOTO CO INC	25	1987	2012	3020	302020	30	30202030	JPN

ELAN CORP PLC	25	1984	2010	3520	352010	35	35201010	IRL
NOVO NORDISK A/S	25	1988	2013	3520	352020	35	35202010	DNK
GENENTECH INC	25	1984	2008	3520	352010	35	35201010	USA
AMGEN INC	25	1986	2013	3520	352010	35	35201010	USA
ROHM AND HAAS CO	24	1981	2008	1510	151010	15	15101050	USA
DU PONT (E I) DE NEMOURS	24	1978	2012	1510	151010	15	15101020	USA
CHUGAI PHARMACEUTICAL CO LTD	23	1987	2011	3520	352020	35	35202010	JPN
AKZO NOBEL NV	23	1984	2010	1510	151010	15	15101050	NLD
TEIJIN LTD	22	1987	2011	1510	151010	15	15101010	JPN
ROCHE HOLDING AG	22	1993	2014	3520	352020	35	35202010	CHE
BAYER AG	21	1993	2013	3520	352020	35	35202010	DEU
TORAY INDUSTRIES INC	21	1987	2012	1510	151010	15	15101010	JPN
PHARMACIA & UPJOHN INC	21	1978	1999	3520	352020	35	35202010	USA
WARNER-LAMBERT CO	20	1980	1999	3520	352020	35	35202010	USA
MEIJI SEIKA KAISHA LTD	20	1987	2008	3020	302020	30	30202030	JPN
VERTEX PHARMACEUTICALS INC	20	1990	2014	3520	352010	35	35201010	USA
IMMUNOMEDICS INC	20	1985	2012	3520	352010	35	35201010	USA
NIPPON KAYAKU CO LTD	20	1987	2010	1510	151010	15	15101050	JPN
ALZA CORP	20	1978	2000	3520	352020	35	35202010	USA
NOVARTIS AG	19	1996	2014	3520	352020	35	35202010	CHE
IONIS PHARMACEUTICALS INC	19	1992	2012	3520	352010	35	35201010	USA
RHONE-POULENC RORER	19	1978	1996	3520	352020	35	35202010	USA
NEKTAR THERAPEUTICS	19	1994	2013	3520	352020	35	35202010	USA
CELGENE CORP	19	1993	2012	3520	352010	35	35201010	USA
ALKERMES PLC	18	1990	2013	3520	352010	35	35201010	IRL
AMYLIN PHARMACEUTICALS INC	18	1990	2010	3520	352010	35	35201010	USA
CEPHALON INC	18	1993	2010	3520	352010	35	35201010	USA
NISSAN CHEMICAL INDUSTRIES	18	1988	2013	1510	151010	15	15101020	JPN
DAIICHI	18	1987	2005	3520	352020	35	35202010	JPN

PHARMACEUTICAL CO								
GILEAD SCIENCES INC	17	1990	2013	3520	352010	35	35201010	USA
NITTO DENKO CORP	17	1987	2013	1510	151010	15	15101050	JPN
IDERA PHARMACEUTICALS INC	17	1994	2011	3520	352010	35	35201010	USA
KOWA SPINNING CO LTD	17	1987	2008	2010	201050	20	20105010	JPN
XOMA CORP	17	1987	2012	3520	352010	35	35201010	USA
UCB SA-NV	17	1989	2012	3520	352020	35	35202010	BEL
ASTRAZENECA PLC	17	1996	2012	3520	352020	35	35202010	GBR
FRESENIUS SE & CO KGAA	17	1989	2011	3510	351020	35	35102015	DEU
SOLVAY SA	17	1990	2008	1510	151010	15	15101020	BEL
PHARMACIA CORP	17	1981	2002	3520	352020	35	35202010	USA
FUJISAWA PHARMACEUTICAL CO	17	1987	2003	3520	352020	35	35202010	JPN
SUNTORY HOLDINGS LTD	17	1994	2011					JPN
SANTEN PHARMACEUTICAL	17	1996	2012	3520	352020	35	35202010	JPN
KAO CORP	16	1988	2011	3030	303020	30	30302010	JPN
REGENERON PHARMACEUTICALS	16	1994	2012	3520	352010	35	35201010	USA
CSL LTD	16	1995	2012	3520	352010	35	35201010	AUS
KURARAY CO LTD	16	1988	2011	1510	151010	15	15101010	JPN
AMERICAN CYANAMID CO	16	1978	1993	3520	352020	35	35202010	USA
SYNTEX CORP	16	1978	1993	3520	352020	35	35202010	PAN
SHIN-ETSU CHEMICAL CO LTD	16	1988	2011	1510	151010	15	15101050	JPN
MALLINCKRODT INC	16	1982	2000	3510	351010	35	35101010	USA
SEPRACOR INC	16	1990	2008	3520	352020	35	35202010	USA
INCYTE CORP	16	1994	2013	3520	352010	35	35201010	USA
YAKULT HONSHA CO LTD	16	1989	2010	3020	302020	30	30202030	JPN
IMMUNOGEN INC	15	1992	2012	3520	352010	35	35201010	USA
TOKUYAMA CORP	15	1994	2010	1510	151010	15	15101050	JPN
NISSHIN SEIFUN GROUP INC	15	1987	2005	3020	302020	30	30202030	JPN
BASF SE	15	1998	2012	1510	151010	15	15101020	DEU
HUMAN GENOME	15	1994	2010	3520	352010	35	35201010	USA

SCIENCES INC								
H LUNDBECK A/S	15	1997	2012	3520	352020	35	35202010	DNK
NEUROGEN CORP	15	1990	2007	3520	352010	35	35201010	USA
KYORIN HOLDINGS INC	15	1996	2012	3520	352020	35	35202010	JPN
MITSUBISHI CHEMICAL HLDGS CO	15	1987	2002	1510	151010	15	15101020	JPN
ASTELLAS PHARMA INC	14	2000	2013	3520	352020	35	35202010	JPN
ORION CORP	14	1994	2008	3520	352020	35	35202010	FIN
SKYEPHARMA PLC	14	1996	2012	3520	352020	35	35202010	GBR
BECTON DICKINSON & CO	14	1985	2007	3510	351010	35	35101010	USA
SANOFI	14	2000	2013	3520	352020	35	35202010	FRA
EISAI CO LTD	14	2000	2013	3520	352020	35	35202010	JPN
FMC CORP	14	1982	2010	1510	151010	15	15101030	USA
CHIRON CORP	14	1987	2005	3520	352010	35	35201010	USA
NIPPON SHINYAKU CO LTD	14	1994	2009	3520	352020	35	35202010	JPN
AETERNA ZENTARIS INC	14	1995	2012	3520	352010	35	35201010	USA
BIOGEN INC	14	1999	2013	3520	352010	35	35201010	USA
NEUROSEARCH A/S	14	1996	2009	3520	352010	35	35201010	DNK
ASTEX PHARMACEUTICALS INC	14	1996	2011	3520	352010	35	35201010	USA
EGIS PHARMACEUTICALS	13	1997	2012	3520	352020	35	35202010	HUN
SEIKAGAKU CORP	13	1997	2013	3520	352020	35	35202010	JPN
BAUSCH & LOMB HLDGS -REDH	13	1985	2012	3510	351010	35	35101020	USA
KAKEN PHARMACEUTICAL CO LTD	13	1989	2014	3520	352020	35	35202010	JPN
HISAMITSU PHARMACEUTICAL CO	13	2000	2013	3520	352020	35	35202010	JPN
SMITHKLINE BEECHAM (UK) PLC	13	1986	1999	3520	352020	35	35202010	GBR
MORINAGA MILK INDUSTRY CORP	13	1991	2010	3020	302020	30	30202030	JPN
TAISHO PHARMACEUTICAL HLDGS	13	1998	2011	3520	352020	35	35202010	JPN
ONO PHARMACEUTICAL CO LTD	13	2000	2014	3520	352020	35	35202010	JPN

GENZYME CORP	13	1996	2009	3520	352010	35	35201010	USA
FUJIFILM HLDGS CORP	13	1992	2010	4520	452020	45	45202030	JPN
JAPAN TOBACCO INC	13	1993	2012	3020	302030	30	30203010	JPN
ARBUTUS BIOPHARMA CORP	13	1995	2013	3520	352010	35	35201010	CAN
BIOGEN INC-OLD	13	1982	2002	3520	352010	35	35201010	USA
MITSUI CHEMICALS INC	13	1994	2011	1510	151010	15	15101010	JPN
KUREHA CORP	13	1990	2007	1510	151010	15	15101010	JPN
ENZON PHARMACEUTICALS INC	13	1992	2009	3520	352010	35	35201010	USA
PHARMA MAR SA	12	1996	2010	3520	352010	35	35201010	ESP
SMITHKLINE BECKMAN CORP	12	1977	1988	3520	352020	35	35202010	USA
LILLY INDS INC - CL A	12	1980	1995	1510	151010	15	15101050	USA
MARINA BIOTECH INC -OLD	12	1985	2010	3520	352010	35	35201010	USA
RIGEL PHARMACEUTICALS INC	12	1999	2012	3520	352010	35	35201010	USA
LIPOSOME CO INC	12	1985	1999	3520	352010	35	35201010	USA
SQUIBB CORP	12	1977	1988	3520	352020	35	35202010	USA
EMISPHERE TECHNOLOGIES INC	12	1993	2013	3520	352020	35	35202010	USA
MILLENNIUM PHARMACEUTICALS	12	1995	2007	3520	352010	35	35201010	USA
MOCHIDA PHARMACEUTICAL CO	12	1994	2012	3520	352020	35	35202010	JPN
NEUROCRINE BIOSCIENCES INC	12	1995	2011	3520	352010	35	35201010	USA
L'OREAL SA	12	1999	2010	3030	303020	30	30302010	FRA
NICOX SA	12	1998	2009	3520	352010	35	35201010	FRA
KANEKA CORP	11	1996	2009	1510	151010	15	15101010	JPN
AVENTIS SA	11	1993	2003	3520	352020	35	35202010	FRA
IDEXX LABS INC	11	1998	2011	3510	351010	35	35101010	USA
FIDIA SPA	11	1999	2012	2010	201060	20	20106020	ITA
PHARMACYCLICS INC	11	1995	2012	3520	352010	35	35201010	USA
INSPIRE PHARMACEUTICALS INC	11	1998	2010	3520	352020	35	35202010	USA
KISSEI	11	1998	2011	3520	352020	35	35202010	JPN

PHARMACEUTICAL CO LTD								
NOVELION THERAPEUTICS INC	11	1989	2007	3520	352010	35	35201010	CAN
SK HOLDINGS CO LTD	11	1999	2012	2010	201050	20	20105010	KOR
KANEBO LTD	11	1987	2001	3030	303020	30	30302010	JPN
EXELIXIS INC	11	2002	2012	3520	352010	35	35201010	USA
RECORDATI SPA	11	1989	2001	3520	352020	35	35202010	ITA
ALBANY MOLECULAR RESH INC	11	2000	2012	3520	352030	35	35203010	USA
NPS PHARMACEUTICALS INC	10	1993	2007	3520	352010	35	35201010	USA
ASAHI KAGAKU KOGYO CO LTD	10	2004	2014	1510	151010	15	15101010	JPN
PLIVA DD	10	1997	2007	3520	352020	35	35202010	HRV
SYNTHELABO SA	10	1989	1998	3520	352020	35	35202010	FRA
MITSUBISHI TANABE PHARMA	10	2003	2012	3520	352020	35	35202010	JPN
IMMUNEX CORP	10	1986	2001	3520	352010	35	35201010	USA
ZYMOGENETICS INC	10	1999	2009	3520	352010	35	35201010	USA
BOSTON SCIENTIFIC CORP	10	2001	2012	3510	351010	35	35101010	USA
BIOMARIN PHARMACEUTICAL INC	10	2003	2012	3520	352010	35	35201010	USA
PONIARD PHARMACEUTICALS INC	10	1987	2003	3520	352010	35	35201010	USA
LIGAND PHARMACEUTICAL INC	10	1995	2012	3520	352010	35	35201010	USA
BAVARIAN NORDIC AS	10	1998	2011	3520	352010	35	35201010	DNK
GUERBET SA	10	1989	2008	3510	351010	35	35101020	FRA
SEKISUI CHEMICAL CO LTD	10	1987	2011	2520	252010	25	25201030	JPN
AVADEL PHARMACEUTICALS -ADR	10	1995	2008	3520	352020	35	35202010	IRL
DANONE	10	2000	2011	3020	302020	30	30202030	FRA
IPSEN SA	10	2002	2014	3520	352020	35	35202010	FRA
DOW CORNING CORP	10	1985	1998	1510	151010	15	15101050	USA
SK CHEMICALS	10	1996	2012	1510	151010	15	15101010	KOR
SHIRE PLC	10	1997	2011	3520	352010	35	35201010	IRL

FOREST LABORATORIES - CL A	10	1983	2013	3520	352020	35	35202010	USA
VERNALIS PLC	10	1999	2010	3520	352010	35	35201010	GBR
ACTELION LTD	10	2000	2011	3520	352010	35	35201010	CHE
MYLAN NV	10	1983	2010	3520	352020	35	35202010	GBR
OSI PHARMACEUTICALS INC	10	1999	2009	3520	352010	35	35201010	USA
JOHNSON MATTHEY PLC	10	1989	2010	1510	151010	15	15101050	GBR
KIRIN HOLDINGS CO LTD	10	1988	2005	3020	302010	30	30201010	JPN
DURECT CORP	10	2000	2010	3520	352020	35	35202010	USA
CURIS INC	10	1998	2012	3520	352010	35	35201010	USA
SUZUKEN CO LTD	10	1993	2009	3510	351020	35	35102010	JPN
SURMODICS INC	10	1998	2011	3510	351010	35	35101010	USA
SYNTA PHARMACEUTICALS CORP	10	2004	2013	3520	352010	35	35201010	USA
CV THERAPEUTICS INC	9	1995	2006	3520	352010	35	35201010	USA
ARENA PHARMACEUTICALS INC	9	1999	2012	3520	352010	35	35201010	USA
VESTAR INC	9	1984	1993	3520	352020	35	35202010	USA
HENKEL AG & CO KGAA	9	1991	2009	3030	303010	30	30301010	DEU
ISHIHARA SANGYO KAISHA LTD	9	1987	2009	1510	151010	15	15101020	JPN
GREEN CROSS CORP	9	1987	1996	3520	352020	35	35202010	JPN
WELLCOME PLC	9	1986	1994	3520	352020	35	35202010	GBR
DEPOMED INC	9	1996	2012	3520	352020	35	35202010	USA
GENETICS INSTITUTE INC	9	1987	1995	3520	352010	35	35201010	USA
MEDAREX INC	9	1998	2008	3520	352010	35	35201010	USA
ARRAY BIOPHARMA INC	9	2001	2012	3520	352010	35	35201010	USA
ZERIA PHARMACEUTICAL CO LTD	9	1996	2010	3520	352020	35	35202010	JPN
MEDIVIR AB	9	1996	2011	3520	352010	35	35201010	SWE
SYNAPTIC PHARMACEUTICAL CORP	9	1992	2001	3520	352010	35	35201010	USA
ONYX PHARMACEUTICALS INC	9	1995	2011	3520	352010	35	35201010	USA

KARO PHARMA AB	9	1997	2010	3520	352010	35	35201010	SWE
EASTMAN KODAK CO	9	1985	1994	4520	452020	45	45202030	USA
TSUMURA & CO	8	1989	2005	3520	352020	35	35202010	JPN
YUHAN CORP	8	1995	2011	3520	352020	35	35202010	KOR
DYNAVAX TECHNOLOGIES CORP	8	2001	2011	3520	352010	35	35201010	USA
ARQULE INC	8	1996	2012	3520	352010	35	35201010	USA
PALATIN TECHNOLOGIES INC	8	2000	2012	3520	352010	35	35201010	USA
ALNYLAM PHARMACEUTICALS INC	8	2003	2012	3520	352010	35	35201010	USA
MITSUI & CO LTD	8	1990	2009	2010	201070	20	20107010	JPN
UBE INDUSTRIES LTD	8	1987	2009	1510	151010	15	15101020	JPN
DELSITE INC	8	1986	2002	3030	303020	30	30302010	USA
DYAX CORP	8	2000	2012	3520	352010	35	35201010	USA
DEGUSSA AG	8	1989	2001	1510	151010	15	15101020	DEU
VECTURA GROUP PLC	8	1998	2009	3520	352020	35	35202010	GBR
APRICUS BIOSCIENCES INC	8	1998	2012	3520	352020	35	35202010	USA
ASHLAND GLOBAL HOLDINGS INC	8	1991	2010	1510	151010	15	15101050	USA
BIOCHEM PHARMA INC	8	1992	2000	3520	352010	35	35201010	CAN
ACORDA THERAPEUTICS INC	8	2004	2012	3520	352010	35	35201010	USA
RECKITT BENCKISER GROUP PLC	8	1993	2009	3030	303010	30	30301010	GBR
TELESTA THERAPEUTICS INC	8	1995	2011	3520	352010	35	35201010	CAN
TERUMO CORP	8	1989	2005	3510	351010	35	35101010	JPN
TARO PHARMACEUTICL INDS LTD	8	1998	2011	3520	352020	35	35202010	ISR
BEIERSDORF AG	8	1989	2010	3030	303020	30	30302010	DEU
MOLECULAR BIOSYSTEMS INC	8	1987	1997	3510	351010	35	35101010	USA
REPROS THERAPEUTICS INC	8	1994	2009	3520	352020	35	35202010	USA
SUCAMPO PHARMACEUTICALS INC	8	2006	2013	3520	352020	35	35202010	USA
DONG A SOCIO HOLDINGS CO LTD	8	1999	2009	3520	352020	35	35202010	KOR

SEATTLE GENETICS INC	8	1999	2012	3520	352010	35	35201010	USA
MATEON THERAPEUTICS INC	8	1992	2011	3520	352010	35	35201010	USA
CYGNUS INC	8	1989	1997	3510	351010	35	35101010	USA
TASLY PHARMACEUTICAL GROUP	8	2002	2011	3520	352020	35	35202010	CHN
CETUS CORP	8	1984	1991	3520	352010	35	35201010	USA
ALMIRALL SA	8	2003	2012	3520	352020	35	35202010	ESP
SCHERER (R P)/DE	8	1985	1997	3520	352020	35	35202010	USA
ALLIANCE PHARMACEUTICAL	8	1987	2007	3520	352030	35	35203010	USA
SCICLONE PHARMACEUTICALS INC	8	1993	2010	3520	352020	35	35202010	USA
SALIX PHARMACEUTICALS LTD	7	2004	2012	3520	352020	35	35202010	USA
ABBVIE INC	7	2008	2014	3520	352010	35	35201010	USA
LEXICON PHARMACEUTICALS INC	7	2006	2012	3520	352010	35	35201010	USA
EPIX PHARMACUETICALS INC	7	2000	2008	3520	352010	35	35201010	USA
MIRATI THERAPEUTICS INC	7	2005	2011	3520	352010	35	35201010	USA
NYCOMED ASA	7	1990	1996	3520	352010	35	35201010	NOR
MEDTRONIC PLC	7	1989	2009	3510	351010	35	35101010	IRL
GUILFORD PHARMACEUTICAL INC	7	1995	2002	3520	352020	35	35202010	USA
SMITH & NEPHEW PLC	7	1993	2010	3510	351010	35	35101010	GBR
COLGATE-PALMOLIVE CO	7	1989	2010	3030	303010	30	30301010	USA
MEDICINES CO	7	2005	2013	3520	352020	35	35202010	USA
ARIAD PHARMACEUTICALS INC	7	2000	2011	3520	352010	35	35201010	USA
NIPRO CORP	7	1999	2008	3510	351010	35	35101020	JPN
CORCEPT THERAPEUTICS INC	7	2000	2013	3520	352020	35	35202010	USA
ANORMED INC	7	1997	2005	3520	352020	35	35202010	CAN
L'AIR LIQUIDE SA	7	2000	2011	1510	151010	15	15101040	FRA
ADDEX PHARMACEUTICALS SA	7	2004	2012	3520	352010	35	35201010	CHE

NIPPON SODA CO LTD	7	1989	2011	1510	151010	15	15101020	JPN
CORIXA CORP	7	1995	2004	3520	352010	35	35201010	USA
HOKURIKU SEIYAKU CO LTD	7	1987	2001	3520	352020	35	35202010	JPN
SHOWA DENKO KK	7	1987	2009	1510	151010	15	15101020	JPN
ICOS CORP	7	1994	2002	3520	352010	35	35201010	USA
ADOLOR CORP	7	1997	2010	3520	352020	35	35202010	USA
MEDIGENE AG	7	1999	2012	3520	352010	35	35201010	DEU
MEDICIS PHARMACEUT CP - CL A	7	2000	2010	3520	352020	35	35202010	USA
SUNESIS PHARMACEUTICALS INC	7	2003	2012	3520	352010	35	35201010	USA
GLENMARK PHARMACEUTICALS LTD	7	2005	2011	3520	352020	35	35202010	IND
VIRBAC CORP	7	1994	2004	3520	352020	35	35202010	USA
IDENIX PHARMACEUTICALS INC	7	2002	2012	3520	352010	35	35201010	USA
ANDRX CORP	7	1995	2004	3520	352020	35	35202010	USA
NOF CORP	7	1994	2009	1510	151010	15	15101020	JPN
UNIGENE LABORATORIES INC	7	2001	2010	3520	352010	35	35201010	USA
BIOCRYST PHARMACEUTICALS INC	7	1994	2011	3520	352010	35	35201010	USA
STRYKER CORP	7	1995	2011	3510	351010	35	35101010	USA
SWEDISH ORPHAN BIOVITRUM AB	7	2001	2011	3520	352010	35	35201010	SWE
AVON PRODUCTS	7	2000	2010	3030	303020	30	30302010	GBR
TRANSGENE	7	1996	2004	3520	352010	35	35201010	FRA
ACHILLION PHARMACEUTICALS	7	2005	2012	3520	352010	35	35201010	USA
EVONIK INDUSTRIES AG	7	2005	2012	1510	151010	15	15101050	DEU
ECOLAB INC	7	1991	2011	1510	151010	15	15101050	USA
ANGES MG INC	7	2001	2011	3520	352010	35	35201010	JPN
NIHON NOHYAKU CO LTD	7	1993	2012	1510	151010	15	15101030	JPN
CELTRIX PHARMACEUTICALS	7	1991	1998	3520	352020	35	35202010	USA
ZEALAND PHARMA AS	7	2000	2009	3520	352010	35	35201010	DNK

UNITIKA LTD	7	1990	2007	2520	252030	25	25203030	JPN
PIRAMAL ENTERPRISES LTD	7	2005	2012	3520	352020	35	35202010	IND
WEST PHARMACEUTICAL SVSC INC	7	1994	2003	3510	351010	35	35101020	USA
ONCOLYTICS BIOTECH INC	7	1999	2011	3520	352010	35	35201010	CAN
BIOTEST AG	7	1989	2012	3520	352010	35	35201010	DEU
MANNKIND CORP	7	2005	2011	3520	352010	35	35201010	USA
MITSUBISHI PHARMA CORP	7	1999	2005	3520	352020	35	35202010	JPN
ACADIA PHARMACEUTICALS INC	7	2002	2013	3520	352010	35	35201010	USA
NOVEN PHARMACEUTICALS INC	7	1991	2008	3520	352020	35	35202010	USA
RANBAXY LABORATORIES LTD	7	2002	2011	3520	352020	35	35202010	IND
PENNWALT CORP	7	1981	1987	1510	151010	15	15101020	USA
GENERAL ELECTRIC CO	6	2002	2011	2010	201050	20	20105010	USA
CADILA HEALTHCARE LTD	6	2001	2011	3520	352020	35	35202010	IND
BAYER SCHERING PHARMA AG	6	1999	2007	3520	352020	35	35202010	DEU
SICOR INC	6	1989	1999	3520	352020	35	35202010	USA
CTI BIOPHARMA CORP	6	1995	2006	3520	352010	35	35201010	USA
PROGENICS PHARMACEUTICAL INC	6	1995	2011	3520	352010	35	35201010	USA
INTERMUNE INC	6	2005	2010	3520	352010	35	35201010	USA
DIC CORPORATION	6	1987	2002	1510	151010	15	15101050	JPN
SANGAMO THERAPEUTICS INC	6	1999	2013	3520	352010	35	35201010	USA
SCIOS INC	6	1993	2001	3520	352010	35	35201010	USA
SYNGENTA AG	6	1999	2009	1510	151010	15	15101030	CHE
FUISZ TECHNOLOGIES LTD	6	1992	1998	3520	352020	35	35202010	USA
BIONOMICS LTD	6	2003	2010	3520	352010	35	35201010	AUS
MEMORY PHARMA CORP	6	2002	2007	3520	352010	35	35201010	USA
DR REDDY'S LABORATORIES LTD	6	1999	2007	3520	352020	35	35202010	IND

LONZA GROUP AG	6	1999	2010	3520	352030	35	35203010	CHE
INFINITY PHARMACEUTICALS INC	6	2007	2012	3520	352010	35	35201010	USA
ASTRA AB	6	1993	1998	3520	352020	35	35202010	SWE
ALTANA AG	6	2000	2005	1510	151010	15	15101050	DEU
UNITED THERAPEUTICS CORP	6	2000	2011	3520	352010	35	35201010	USA
THRESHOLD PHARMACEUTICALS	6	2003	2009	3520	352010	35	35201010	USA
ICAGEN INC	6	2001	2010	3520	352010	35	35201010	USA
3-DIMENSIONAL PHARMACEUTICAL	6	1995	2001	3520	352020	35	35202010	USA
CJ CORP	6	2002	2010	2010	201050	20	20105010	KOR
ROBINS (A.H.) CO	6	1982	1988	3520	352020	35	35202010	USA
ENDO INTERNATIONAL PLC	6	1999	2012	3520	352020	35	35202010	IRL
XENOPORT INC	6	2004	2010	3520	352020	35	35202010	USA
CHEMOCENTRYX INC	6	2005	2010	3520	352010	35	35201010	USA
KOSAN BIOSCIENCES INC	6	2001	2006	3520	352010	35	35201010	USA
NOVOGEN LTD	6	1999	2006	3520	352020	35	35202010	AUS
LUPIN LTD	6	2004	2011	3520	352020	35	35202010	IND
RESPIRERX PHARMACEUTICALS	6	1994	2011	3520	352020	35	35202010	USA
CURAGEN CORP	6	1997	2002	3520	352010	35	35201010	USA
ANTISENSE THERAPEUTICS LTD	6	2004	2010	3520	352020	35	35202010	AUS
CELL GENESYS INC	6	1994	2002	3520	352010	35	35201010	USA
COLEY PHARMACEUTICAL GROUP	6	2001	2006	3520	352010	35	35201010	USA
PRANA BIOTECHNOLOGY LTD	6	2000	2008	3520	352010	35	35201010	AUS
SCOTIA HOLDINGS PLC	6	1994	1999	3520	352020	35	35202010	GBR
ROHTO PHARMACEUTICAL CO LTD	6	1989	2012	3520	352020	35	35202010	JPN
ONCOTHERAPY SCIENCE INC	6	2006	2012	3520	352010	35	35201010	JPN
TARGACEPT INC	6	2004	2011	3520	352010	35	35201010	USA

MEDA AB	6	2000	2009	3520	352020	35	35202010	SWE
PERSTORP AB	6	1989	1998	1510	151010	15	15101020	SWE
ANADYS PHARMACEUTICALS INC	6	2002	2010	3520	352010	35	35201010	USA
CYTOKINETICS INC	6	2003	2009	3520	352010	35	35201010	USA
VIVUS INC	6	1997	2009	3520	352020	35	35202010	USA
ZOETIS INC	6	2007	2013	3520	352020	35	35202010	USA
JUNIPER PHARMACEUTICALS INC	6	1993	2003	3520	352020	35	35202010	USA
AMICUS THERAPEUTICS INC	6	2004	2012	3520	352010	35	35201010	USA
CELL PATHWAYS INC	6	1994	2000	3520	352020	35	35202010	USA
ASKA PHARM CO LTD	6	2003	2009	3520	352020	35	35202010	JPN
CELLTECH GROUP PLC	6	1998	2003	3520	352010	35	35201010	GBR
CASI PHARMACEUTICALS INC	6	1995	2008	3520	352010	35	35201010	USA
YM BIOSCIENCES INC	6	2003	2011	3520	352010	35	35201010	CAN
GRACE (W R) & CO	6	1978	1997	1510	151010	15	15101050	USA
SUMITOMO CHEMICAL CO LTD	6	1987	2011	1510	151010	15	15101020	JPN
CALPIS CO LTD	6	1996	2006	3020	302020	30	30202030	JPN
INSITE VISION INC	6	1993	2000	3520	352020	35	35202010	USA
CYCLACEL PHARMACEUTICALS	5	2006	2012	3520	352010	35	35201010	USA
DENKI KOGYO CO LTD	5	1989	2009	4520	452010	45	45201020	JPN
KEY PHARMACEUTICALS INC	5	1980	1985	3520	352020	35	35202010	USA
AVIGEN INC	5	1997	2007	3520	352010	35	35201010	USA
SUVEN LIFE SCIENCES LTD	5	2005	2010	3520	352020	35	35202010	IND
TULARIK INC	5	1997	2002	3520	352010	35	35201010	USA
ALEXION PHARMACEUTICALS INC	5	2003	2011	3520	352010	35	35201010	USA
CENES PHARMACEUTICALS PLC	5	1998	2007	3520	352020	35	35202010	GBR
CSPC PHARMACEUTICAL GROUP	5	2004	2011	3520	352020	35	35202010	HKG

ACTIVE BIOTECH AB	5	1999	2005	3520	352010	35	35201010	SWE
BTG PLC	5	1993	1997	3520	352020	35	35202010	GBR
AVANIR PHARMACEUTICALS INC	5	1997	2007	3520	352020	35	35202010	USA
ZIMMER BIOMET HOLDINGS INC	5	2000	2011	3510	351010	35	35101010	USA
NIPPON CHEMIPHAR CO LTD	5	1999	2008	3520	352020	35	35202010	JPN
HEMOSOL CORP	5	1994	2002	3520	352010	35	35201010	CAN
BTG INC	5	1994	2000	4520	452020	45	45202010	USA
GENTA INC	5	1996	2008	3520	352010	35	35201010	USA
MEDEVA PLC	5	1994	1998	3520	352020	35	35202010	GBR
WYETH LTD	5	1997	2001	3520	352020	35	35202010	IND
PROMETIC LIFE SCIENCES INC	5	2004	2011	3520	352010	35	35201010	CAN
MILES LABORATORIES INC	5	1981	1986	3520	352020	35	35202010	USA
PRONOVA BIOPHARMA ASA	5	2005	2010	3520	352020	35	35202010	NOR
GENAERA CORP	5	1991	1999	3520	352010	35	35201010	USA
GLYCOMED INC	5	1990	1994	3520	352020	35	35202010	USA
MITSUI TOATSU CHEMICALS INC	5	1988	1994	1510	151010	15	15101010	JPN
CONNETICS CORP	5	1995	2004	3520	352020	35	35202010	USA
ORCHID PHARMA LTD	5	2003	2010	3520	352020	35	35202010	IND
COR THERAPEUTICS INC	5	1994	2000	3520	352010	35	35201010	USA
GERON CORP	5	1997	2004	3520	352010	35	35201010	USA
OTSUKA HOLDINGS CO LTD	5	2008	2013	3520	352020	35	35202010	JPN
VICAL INC	5	1996	2012	3520	352010	35	35201010	USA
ENZO BIOCHEM INC	5	1998	2009	3520	352030	35	35203010	USA
TELIK INC -OLD	5	2002	2011	3520	352010	35	35201010	USA
FISONS PLC	5	1989	1994	3510	351020	35	35102010	GBR
MAXIM PHARMACEUTICALS INC	5	1995	2002	3520	352010	35	35201010	USA
OXFORD BIOMEDICA LTD	5	1996	2008	3520	352010	35	35201010	GBR
AGOURON PHARMACEUTICALS INC	5	1992	1997	3520	352010	35	35201010	USA

STERLING DRUG INC	5	1981	1986	3520	352020	35	35202010	USA
BLOCK DRUG -CL A	5	1987	1996	3510	351010	35	35101020	USA
INNATE PHARMA SA	5	2004	2012	3520	352010	35	35201010	FRA
MERCIAN CORP	5	1998	2005	3020	302010	30	30201020	JPN
OMEROS CORP	5	2006	2011	3520	352020	35	35202010	USA
ALK-ABELLO A/S	5	2004	2010	3520	352020	35	35202010	DNK
ID BIOMEDICAL CORP	5	1998	2004	3520	352010	35	35201010	CAN
NORSK HYDRO ASA	5	1989	1997	1510	151040	15	15104010	NOR
SEARLE (G.D.) & CO	5	1979	1984	3520	352020	35	35202010	USA
KIKKOMAN CORP	5	1996	2010	3020	302020	30	30202030	JPN
ALIZYME PLC	5	1998	2004	3520	352010	35	35201010	GBR
SUNSTAR INC	5	1990	2006	3030	303020	30	30302010	JPN
LION CORP	5	1988	2013	3030	303010	30	30301010	JPN
NITROMED INC	5	1998	2003	3520	352020	35	35202010	USA
PROBI AB	5	2000	2008	3520	352010	35	35201010	SWE
VASOGEN INC	5	1998	2003	3520	352010	35	35201010	CAN
MERCK SERONO SA	5	2000	2005	3520	352010	35	35201010	DEU
CORVAS INTERNATIONAL INC	5	1991	1998	3520	352010	35	35201010	USA
MORISHITA JINTAN CO LTD	5	2000	2012	3030	303020	30	30302010	JPN
METABASIS THERAPEUTICS INC	5	1999	2008	3520	352010	35	35201010	USA
SIGA TECHNOLOGIES INC	5	2004	2010	3520	352010	35	35201010	USA
ARCHER-DANIELS-MIDLAND CO	5	1996	2008	3020	302020	30	30202010	USA
MEDICINOVA INC	5	2005	2012	3520	352010	35	35201010	USA
MEITO SANGYO CO LTD	5	1995	2008	3020	302020	30	30202030	JPN
ATRIX LABORATORIES INC	5	1991	1999	3520	352020	35	35202010	USA
LVMH MOET HENNESSY LOUIS V	5	1992	2007	2520	252030	25	25203010	FRA
MITSUBISHI RAYON CO LTD	5	1987	1992	1510	151010	15	15101010	JPN
KING PHARMACEUTICALS INC	5	1998	2008	3520	352020	35	35202010	USA
HOECHST AG	5	1994	1998	3520	352020	35	35202010	DEU

ALCHEMIA LTD	5	2003	2011	3520	352010	35	35201010	AUS
SUGEN INC	5	1994	1998	3520	352010	35	35201010	USA
ARCH CHEMICALS INC	5	2002	2010	1510	151010	15	15101050	USA
LABOPHARM INC	5	2000	2006	3520	352020	35	35202010	CAN
SAREPTA THERAPEUTICS INC	5	1997	2002	3520	352010	35	35201010	USA
GENVEC INC	5	1998	2006	3520	352010	35	35201010	USA
AMOREPACIFIC CORP	5	2006	2011	3030	303020	30	30302010	KOR
YUNGSHIN GLOBAL HOLDING CORP	5	2002	2006	3520	352020	35	35202010	TWN
SHAMAN PHARMACEUTICALS INC	4	1991	1996	3520	352010	35	35201010	USA
FH FAULDING & CO LTD	4	1990	1999	3520	352020	35	35202010	AUS
EMERGENT GROUP INC	4	2005	2010	3510	351020	35	35102010	USA
EASTMAN CHEMICAL CO	4	1995	2008	1510	151010	15	15101020	USA
NEUREN PHARMACEUTICALS LTD	4	2002	2008	3520	352020	35	35202010	AUS
ASAHI KASEI CORP	4	1987	1991	1510	151010	15	15101010	JPN
DIATIDE INC	4	1994	1998	3520	352010	35	35201010	USA
GALENICA AG	4	1996	2008	3520	352020	35	35202010	CHE
ARADIGM CORP	4	1997	2002	3520	352020	35	35202010	USA
BONE CARE INTERNATIONAL INC	4	1998	2002	3520	352020	35	35202010	USA
RIBI IMMUNOCHEM RESEARCH INC	4	1991	1998	3520	352010	35	35201010	USA
WOCKHARDT LTD	4	2001	2012	3520	352020	35	35202010	IND
MITSUBISHI GAS CHEMICAL CO	4	1990	2007	1510	151010	15	15101020	JPN
NEUTEC PHARMA PLC	4	1998	2005	3520	352010	35	35201010	GBR
REPLIGEN CORP	4	1992	2000	3520	352010	35	35201010	USA
ANACOR PHARMACEUTICALS INC	4	2007	2011	3520	352010	35	35201010	USA
RAPTOR PHARMACEUTICAL CORP	4	2006	2009	3520	352010	35	35201010	USA
PHARMING GROUP NV	4	1998	2006	3520	352010	35	35201010	NLD
ALPHA BETA TECHNOLOGY INC	4	1988	1994	3520	352020	35	35202010	USA

NOVAVAX INC	4	1995	2008	3520	352010	35	35201010	USA
SHISEIDO CO LTD	4	1989	1992	3030	303020	30	30302010	JPN
CANON INC	4	1992	2011	4520	452020	45	45202030	JPN
CELLEGY PHARMACEUTICALS -OLD	4	1998	2002	3520	352020	35	35202010	USA
ZILA INC	4	1990	2002	3520	352020	35	35202010	USA
BIOCON LTD	4	2003	2008	3520	352010	35	35201010	IND
INTERCELL AG	4	2003	2012	3520	352010	35	35201010	AUT
OXIS INTERNATIONAL INC	4	1996	2000	3520	352010	35	35201010	USA
PHARMACOPEIA INC	4	2002	2007	3520	352010	35	35201010	USA
CIPLA LTD	4	2007	2011	3520	352020	35	35202010	IND
K V PHARMACEUTICAL -CL A	4	1994	2002	3520	352020	35	35202010	USA
GENTIUM SPA - ADR	4	2005	2010	3520	352010	35	35201010	ITA
HITACHI LTD	4	1991	2009	4520	452030	45	45203010	JPN
ATHEROGENICS INC	4	1998	2003	3520	352020	35	35202010	USA
HERON THERAPEUTICS INC	4	1989	1998	3520	352010	35	35201010	USA
OREXIGEN THERAPEUTICS INC	4	2005	2010	3520	352020	35	35202010	USA
GELTEX PHARMACEUTICALS INC	4	1993	1998	3520	352010	35	35201010	USA
BIOMET INC	4	2003	2009	3510	351010	35	35101010	USA
SANGSTAT MEDICAL CORP	4	1995	2000	3520	352010	35	35201010	USA
ELF AQUITAINE SA	4	1991	1994	1010	101020	10	10102010	FRA
PHARMOS CORP	4	1993	2002	3520	352020	35	35202010	USA
KALOBIOS PHARMACEUTICALS INC	4	2009	2012	3520	352010	35	35201010	USA
CHELSEA THERAPEUTICS INTL	4	2005	2010	3520	352010	35	35201010	USA
OREXO AB	4	2005	2013	3520	352020	35	35202010	SWE
TITAN PHARMACEUTICALS INC	4	1996	2012	3520	352020	35	35202010	USA
STARPHARMA HLDGS LTD	4	2001	2007	3520	352020	35	35202010	AUS
SSP CO LTD	4	1998	2003	3520	352020	35	35202010	JPN

CREATIVE BIOMOLECULES INC	4	1991	1995	3520	352010	35	35201010	USA
CRUCELL NV	4	2004	2010	3520	352010	35	35201010	NLD
ENCYSIVE PHARMACEUTICALS INC	4	1994	2000	3520	352010	35	35201010	USA
CENCO INC	4	1977	1980					USA
ABRAXIS BIOSCIENCE INC	4	2006	2009	3520	352010	35	35201010	USA
MICRO THERAPEUTICS INC	4	1996	2002	3510	351010	35	35101010	USA
PHOTOCURE ASA	4	1999	2008	3520	352020	35	35202010	NOR
POWDERJECT PHARMACEUTICALS	4	1998	2002	3520	352010	35	35201010	GBR
OPKO HEALTH INC	4	2008	2012	3520	352010	35	35201010	USA
MEDIVATION INC	4	2008	2012	3520	352010	35	35201010	USA
POINT THERAPEUTICS INC	4	1998	2001	3520	352010	35	35201010	USA
VANDA PHARMACEUTICALS INC	4	2007	2013	3520	352010	35	35201010	USA
GENFIT	4	2004	2010	3520	352030	35	35203010	FRA
CUBIST PHARMACEUTICALS INC	4	1999	2013	3520	352010	35	35201010	USA
HERCULES INC	4	1983	2001	1510	151010	15	15101020	USA
ARALEZ PHARMACEUTICALS INC	4	2000	2010	3520	352020	35	35202010	CAN
HAUSER INC	4	1993	2001	1510	151010	15	15101050	USA
ARROWHEAD PHARMACEUTICALS	4	2011	2014	3520	352010	35	35201010	USA
BIO-RAD LABORATORIES INC	4	1999	2006	3520	352030	35	35203010	USA
STEMCELLS INC	4	1994	1997	3520	352010	35	35201010	USA
NORTHFIELD LABORATORIES INC	4	1999	2005	3520	352010	35	35201010	USA
COMPUGEN LTD	4	2006	2009	3520	352030	35	35203010	ISR
IMCLONE SYSTEMS INC	4	1999	2007	3520	352010	35	35201010	USA
MAXYGEN INC	4	2000	2006	3520	352010	35	35201010	USA
LEK DD	4	1996	2002	3520	352020	35	35202010	SVN
NOVA PHARMACEUTICAL	4	1986	1991	3520	352020	35	35202010	USA

CORP								
INHIBITEX INC	4	2001	2010	3520	352010	35	35201010	USA
BIO REFERENCE LABS	4	1999	2004	3510	351020	35	35102015	USA
CYTOGEN CORP	4	1988	1998	3520	352010	35	35201010	USA
DENDREON CORP	4	2000	2007	3520	352010	35	35201010	USA
NUTRITION 21 INC	4	1995	2005	3030	303020	30	30302010	USA
IRONWOOD PHARMACEUTICALS INC	4	2007	2011	3520	352010	35	35201010	USA
COVIDIEN PLC	4	2006	2011	3510	351010	35	35101010	IRL
APHTON CORP	4	1993	2001	3520	352010	35	35201010	USA
PHARMASSET INC	4	2002	2010	3520	352010	35	35201010	USA
CONJUCHEM BIOTECH INC	4	1999	2003	3520	352010	35	35201010	USA
PRAECIS PHARMACEUTICALS INC	4	1996	2001	3520	352010	35	35201010	USA
CASCADIAN THERAPEUTICS INC	4	2002	2012	3520	352010	35	35201010	USA
CHONG KUN DANG HLDGS CORP	4	1998	2010	3520	352020	35	35202010	KOR
TAKARA HOLDINGS INC	4	1991	2001	3020	302010	30	30201020	JPN
SILENCE THERAPEUTICS PLC	4	2003	2010	3520	352010	35	35201010	GBR
ARYX THERAPEUTICS INC	4	2002	2008	3520	352020	35	35202010	USA
DIADEXUS INC	4	2001	2010	3520	352010	35	35201010	USA
BIODEL INC	3	2005	2009	3520	352020	35	35202010	USA
DAEWOONG PHARM CO LTD	3	2007	2013	3520	352020	35	35202010	KOR
BIOSPHERE MEDICAL INC	3	1999	2006	3520	352010	35	35201010	USA
GENESIS RESEARCH & DEVELOPMT	3	1999	2002	3520	352010	35	35201010	NZL
TOYO SUISAN KAISHA LTD	3	1999	2012	3020	302020	30	30202030	JPN
NEOPHARM INC-OLD	3	1998	2009	3520	352010	35	35201010	USA
RADIUS HEALTH INC	3	2007	2011	3520	352010	35	35201010	USA
TERCICA INC	3	2004	2007	3520	352010	35	35201010	USA
PROXIMAGEN GROUP PLC	3	2008	2010	3520	352010	35	35201010	GBR
ENDOCYTE INC	3	2006	2011	3520	352020	35	35202010	USA
ACUSPHERE INC	3	1998	2001	3520	352020	35	35202010	USA

INDEVUS PHARMACEUTICALS INC	3	1989	1998	3520	352010	35	35201010	USA
HESKA CORP	3	1995	2000	3520	352020	35	35202010	USA
LARGE SCALE BIOLOGY CORP	3	1997	1999	3520	352010	35	35201010	USA
VALEANT PHARMACEUTICALS -OLD	3	1999	2008	3520	352020	35	35202010	USA
PTC THERAPEUTICS INC	3	2005	2010	3520	352010	35	35201010	USA
HARBOR DIVERSIFIED INC	3	2006	2009	3520	352010	35	35201010	USA
INSPIRED CAPITAL PLC	3	2005	2007	4020	402010	40	40201040	GBR
MOLECULAR INSIGHT PHARMACTLS	3	2003	2009	3520	352010	35	35201010	USA
BIOGAIA AB	3	2007	2009	3520	352010	35	35201010	SWE
GLOBEIMMUNE INC	3	2009	2012	3520	352010	35	35201010	USA
COCENSYS INC	3	1995	1997	3520	352020	35	35202010	USA
MERRIMACK PHARMACEUTICALS	3	2008	2012	3520	352010	35	35201010	USA
INNOGENETICS SA	3	2000	2002	3520	352010	35	35201010	BEL
BIOMERIEUX	3	2003	2007	3510	351010	35	35101010	FRA
LAVIPHARM SA	3	1999	2002	3520	352020	35	35202010	GRC
HALOZYME THERAPEUTICS INC	3	2006	2010	3520	352010	35	35201010	USA
SONUS PHARMACEUTICALS INC	3	1994	2004	3520	352010	35	35201010	USA
HELIX BIOMEDIX INC	3	2002	2008	3520	352010	35	35201010	USA
ARDEA BIOSCIENCES INC	3	2009	2011	3520	352010	35	35201010	USA
PANACEA BIOTEC LTD	3	2002	2009	3520	352010	35	35201010	IND
MATRIX PHARMACEUTICAL INC	3	1994	1997	3520	352010	35	35201010	USA
TOYAMA CHEMICAL CO LTD	3	2004	2007	3520	352020	35	35202010	JPN
FAES FARMA SA	3	2000	2008	3520	352020	35	35202010	ESP
BIODELIVERY SCIENCES INTL	3	2007	2012	3520	352020	35	35202010	USA
XENOVA GROUP PLC	3	1995	2001	3520	352010	35	35201010	GBR
LIFECELL CORP	3	1993	2000	3520	352010	35	35201010	USA

WILEX AG	3	2004	2006	3520	352010	35	35201010	DEU
BANYU PHARMACEUTICAL CO LTD	3	1987	2002	3520	352020	35	35202010	JPN
TORRENT PHARMACEUTICALS LTD	3	2002	2008	3520	352020	35	35202010	IND
SYMRISE AG	3	2006	2012	1510	151010	15	15101050	DEU
MEDICURE INC	3	2000	2006	3520	352010	35	35201010	CAN
ASCENT PEDIATRICS INC	3	1994	1996	3520	352020	35	35202010	USA
FENNEC PHARMACEUTICALS INC	3	2005	2011	3520	352010	35	35201010	USA
FIVE PRIME THERAPEUTICS INC	3	2010	2012	3520	352010	35	35201010	USA
VICURON PHARMACEUTICALS INC	3	2001	2003	3520	352010	35	35201010	USA
AMERSHAM PLC	3	2000	2002	3510	351010	35	35101010	GBR
CENTOCOR INC	3	1988	1995	3520	352010	35	35201010	USA
THERAVANCE BIOPHARMA INC	3	2010	2012	3520	352020	35	35202010	CYM
GENMAB AS	3	2007	2009	3520	352010	35	35201010	DNK
BIOCOMPATIBLES INTL PLC	3	2001	2007	3510	351010	35	35101010	GBR
NYMOX PHARMACEUTICAL CORP	3	1998	2007	3520	352010	35	35201010	BHS
FREUND CORP	3	1996	1999	2010	201060	20	20106020	JPN
CARDIOME PHARMA CORP	3	1999	2005	3520	352020	35	35202010	CAN
LG LIFE SCIENCES LTD	3	2002	2011	3520	352020	35	35202010	KOR
SPECTRUM PHARMACEUTICALS INC	3	2002	2011	3520	352010	35	35201010	USA
GTX INC	3	2007	2011	3520	352010	35	35201010	USA
ACRUX LTD	3	2004	2013	3520	352020	35	35202010	AUS
CYTRX CORP	3	1993	2007	3520	352010	35	35201010	USA
HANWHA CHEMICAL CORP	5	1998	2012	1510	151010	15	15101010	KOR
ACAMBIS PLC	3	1999	2002	3520	352010	35	35201010	GBR
LA JOLLA PHARMACEUTICAL CO	3	1999	2001	3520	352010	35	35201010	USA
REXAHN PHARMACEUTICALS INC	2	2007	2010	3520	352010	35	35201010	USA

CONCERT PHARMACEUTICLS INC	2	2011	2012	3520	352010	35	35201010	USA
LIPOCINE INC	2	2011	2012	3520	352020	35	35202010	USA
AGENNIX AG	2	2004	2006	3520	352010	35	35201010	DEU
PORTOLA PHARMACEUTICALS INC	2	2010	2011	3520	352010	35	35201010	USA
AMBIT BIOSCIENCES CORP	2	2011	2012	3520	352010	35	35201010	USA
RICHARDSON-VICKS INC	2	1983	1984	3030	303020	30	30302010	USA
SUPERNUS PHARMACEUTICALS INC	2	2007	2011	3520	352020	35	35202010	USA
NELSON RESEARCH & DEV CO	2	1980	1985	2020	202010	20	20201030	USA
CIMA LABS INC	2	1998	1999	3520	352020	35	35202010	USA
GW PHARMACEUTICALS PLC	2	2011	2012	3520	352020	35	35202010	GBR
BIOGLAN PHARMA PLC	2	1997	1999	3520	352020	35	35202010	GBR
TANOX INC	2	1997	2000	3520	352010	35	35201010	USA
ENANTA PHARMACEUTICALS INC	2	2010	2011	3520	352010	35	35201010	USA
MONSANTO CO	2	1998	1999	1510	151010	15	15101030	USA
GRIFOLS SA	2	2009	2010	3520	352010	35	35201010	ESP
WAKAMOTO PHARMACEUTICAL CO	2	2003	2006	3520	352020	35	35202010	JPN
AVIRAGEN THERAPEUTICS INC	2	2009	2011	3520	352010	35	35201010	USA
PROBIODRUG AG	2	2011	2012	3520	352010	35	35201010	DEU
AMBRX INC-REDH	2	2012	2013	3520	352010	35	35201010	USA
CONNAUGHT BIOSCIENCES INC	2	1984	1987	3520	352010	35	35201010	CAN
AB SCIENCE	2	2011	2012	3520	352020	35	35202010	FRA
AXYS PHARMACEUTICALS INC	2	1996	2000	3520	352020	35	35202010	USA
ZENYTH THERAPEUTICS	2	1997	2003	3520	352010	35	35201010	AUS
JCR PHARMACEUTICALS CO LTD	2	2008	2010	3520	352020	35	35202010	JPN
ZALICUS INC	2	2004	2005	3520	352010	35	35201010	USA
MACROGENICS INC	2	2012	2013	3520	352010	35	35201010	USA

CYTOTOOLS AG	2	2005	2009	3520	352010	35	35201010	DEU
SEQUUS PHARMACEUTICALS INC	2	1993	1997	3520	352020	35	35202010	USA
MAP PHARMACEUTICALS INC	2	2005	2008	3520	352020	35	35202010	USA
BIOPURE CORP	2	1999	2001	3520	352010	35	35201010	USA
GENELABS TECHNOLOGIES INC	2	2000	2001	3520	352010	35	35201010	USA
NIKKEN CHEMICALS CO LTD	2	1999	2000	3520	352020	35	35202010	JPN
CERUS CORP	2	1998	1999	3510	351010	35	35101020	USA
SCINOPHARM TAIWAN LTD	2	2009	2010	3520	352020	35	35202010	TWN
CUMBERLAND PHARMACEUTICALS	2	2009	2010	3520	352020	35	35202010	USA
SUN PHARMA ADVANCED RESEARCH	2	2008	2010	3520	352020	35	35202010	IND
CHOONGWAE HOLDINGS CO LTD	1	2009	2009	3520	352020	35	35202010	KOR
QUARK PHARMACEUTICALS -REDH	1	2005	2005	3520	352010	35	35201010	USA
CAMPINA AG	1	1994	1994					DEU
CYDEX PHARMACEUTICALS-REDH	1	2006	2006	3520	352020	35	35202010	USA
XENETIC BIOSCIENCES PLC	1	2013	2013	3520	352010	35	35201010	GBR
ALLERGY THERAPEUTICS PLC	1	2005	2005	3520	352020	35	35202010	GBR
SKW TROSTBERG AG	1	1999	1999	1510	151010	15	15101020	DEU
SUMMIT THERAPEUTICS PLC	1	2012	2012	3520	352010	35	35201010	GBR
ICN PHARMACEUTICALS -OLD	1	1984	1984	3520	352020	35	35202010	USA
ASAHI HOLDINGS INC	1	2012	2012	1510	151040	15	15104040	JPN
ESPERION THERAPEUTICS INC	1	2012	2012	3520	352010	35	35201010	USA
XENCOR INC	1	2011	2011	3520	352010	35	35201010	USA
NUTRAMAX	1	1994	1994	3030	303020	30	30302010	USA

PRODUCTS INC								
DENTSPLY INTERNATIONAL INC	1	1981	1981					USA
NEURODERM LTD	1	2014	2014	3520	352020	35	35202010	ISR
LIGHT SCIENCES ONCOLOGY-REDH	1	2005	2005	3520	352010	35	35201010	USA
KOKEN LTD	1	1987	1987	2020	202010	20	20201060	JPN
GIST-BROCADES (KONINKLIJ) NV	1	1997	1997	3510	351020	35	35102010	NLD
EURAND NV	1	2008	2008	3520	352020	35	35202010	NLD
XENON PHARMACEUTICALS INC	1	2012	2012	3520	352010	35	35201010	CAN
LIXTE BIOTECHNOLOGY HOLDINGS	1	2011	2011	3520	352010	35	35201010	USA
ONCONOVA THERAPEUTICS INC	1	2011	2011	3520	352010	35	35201010	USA
SAMYANG CORP	1	2011	2011	3020	302020	30	30202030	KOR
OMRIX BIOPHARMACEUTICALS	1	2007	2007	3520	352010	35	35201010	USA
OSCOTEC INC	1	2011	2011	3520	352020	35	35202010	KOR
BURCON NUTRASCIENCE CORP	1	2008	2008	1510	151010	15	15101050	CAN
DEPUY INC	1	1997	1997	3510	351010	35	35101010	USA
MCNEIL CORP	1	1977	1977	2010	201060	20	20106020	USA
SIRTRIS PHARMACEUTICALS INC	1	2006	2006	3520	352020	35	35202010	USA
ALLIANCE BOOTS PLC	1	2001	2001	3010	301010	30	30101010	GBR
HEMAGEN DIAGNOSTICS INC	1	1994	1994	3510	351010	35	35101010	USA
AUSPEX PHARMACEUTICALS INC	1	2012	2012	3520	352010	35	35201010	USA
OSIRIS THERAPEUTICS INC	1	2006	2006	3520	352010	35	35201010	USA
BIONUMERIK PHARMA -REDH	1	2002	2002	3520	352010	35	35201010	USA
HANMI PHARM CO LTD	1	2011	2011	3520	352020	35	35202010	KOR

d) List of Firms with at least 20 Years of Data Points and Assigned to the "Pharmaceuticals, Biotechnology & Life sciences" Industry Group

Firm Name	Years of Data	Beginning Year	End Year	GIC Groups	GIC Industries	GIC Sectors	GIC Sub-Industries	Country HQ
MERCK & CO	37	1977	2013	3520	352020	35	35202010	USA
JOHNSON & JOHNSON	37	1978	2014	3520	352020	35	35202010	USA
PFIZER INC	37	1977	2013	3520	352020	35	35202010	USA
BRISTOL-MYERS SQUIBB CO	33	1981	2013	3520	352020	35	35202010	USA
GLAXOSMITHKLINE PLC	33	1980	2014	3520	352020	35	35202010	GBR
SCHERING-PLOUGH	31	1978	2008	3520	352020	35	35202010	USA
WYETH	28	1978	2008	3520	352020	35	35202010	USA
TAKEDA PHARMACEUTICAL CO	27	1987	2013	3520	352020	35	35202010	JPN
KYOWA HAKKO KIRIN CO LTD	27	1987	2013	3520	352020	35	35202010	JPN
TEVA PHARMACEUTICALS	26	1986	2012	3520	352020	35	35202010	ISR
SHIONOGI & CO LTD	26	1987	2012	3520	352020	35	35202010	JPN
AMGEN INC	25	1986	2013	3520	352010	35	35201010	USA
SUMITOMO DAINIPPON PHARMA CO	25	1987	2012	3520	352020	35	35202010	JPN
ELAN CORP PLC	25	1984	2010	3520	352010	35	35201010	IRL
ALLERGAN INC	25	1988	2012	3520	352020	35	35202010	USA
GENENTECH INC	25	1984	2008	3520	352010	35	35201010	USA
NOVO NORDISK A/S	25	1988	2013	3520	352020	35	35202010	DNK
CHUGAI PHARMACEUTICAL CO LTD	23	1987	2011	3520	352020	35	35202010	JPN
ROCHE HOLDING AG	22	1993	2014	3520	352020	35	35202010	CHE
BAYER AG	21	1993	2013	3520	352020	35	35202010	DEU

PHARMACIA & UPJOHN INC	21	1978	1999	3520	352020	35	35202010	USA
IMMUNOMEDICS INC	20	1985	2012	3520	352010	35	35201010	USA
ALZA CORP	20	1978	2000	3520	352020	35	35202010	USA
VERTEX PHARMACEUTICALS INC	20	1990	2014	3520	352010	35	35201010	USA
WARNER-LAMBERT CO	20	1980	1999	3520	352020	35	35202010	USA

e) Selected Stata Codes

Main Regression:

Count dependent variable:

xtnbreg firm_year_patent xrd_intensity NPL_intensity_year_firm NPL_intensity_year_firm_quad ln_annual_sales ln_emp originality_firm_year radicalness_firm_year i.filing, fe

Weighted dependent variable:

xtnbreg weignted_firm_year_fwd_cits7 xrd_intensity NPL_intensity_year_firm NPL_intensity_year_firm_quad ln_annual_sales ln_emp originality_firm_year radicalness_firm_year i.filing, fe

Dependent variable:

Count dependent variable:

by filing gvkey patent_no, sort: gen numberofpat1=_n

gen numberofpat2=1 if numberofpat1 ==1

by filing gvkey: egen firm_year_patent=count(numberofpat2)

label variable firm_year_patent "Number of patents per year per firm"

drop numberofpat2 numberofpat1

Weighted dependent variable:

*** first seven year citation

sort filing gvkey

by filing gvkey: egen firm_year_fwd_cits7=sum(fwd_cits7)

*** dependent variabale (wieghted) based on 1+citation (Trajtenberg, 1990)

gen weignted_firm_year_fwd_cits7=firm_year_patent+firm_year_fwd_cits7

label variable weignted_firm_year_fwd_cits7 "Weighted patent per year by 7-year patent citations"

NPL intensity per year per firm

** total priorart per year per firm

sort filing gvkey patent_id

by filing gvkey: egen prior_art_per_year_firm=total (prior_art)

label variable prior_art_per_year_firm "Number of Priorart per year per firm"

** NPL intensity per year per firm

gen

NPL_intensity_year_firm=(NPLs_year_firm/prior_art_per_year_firm)/firm_year_patent

label variable NPL_intensity_year_firm "NPL intensity per year per firm"

Originality

*** Average Originality per year per firm

sort filing gvkey patent_id

by filing gvkey: egen originality_firm_year=mean (originality)

label variable originality_firm_year "Average originality per year per firm"

Radicalness

*** Average radicalness per year per firm

```
sort filing gvkey patent_id

by filing gvkey: egen radicalness_firm_year=mean (radicalness)

label variable radicalness_firm_year "Average radicalness per year per firm"
```

f) Summary Statistics

Variable		Mean	Std. Dev.	Min	Max	Observations
$Patents_{it}$	O	4.592308	9.29794	1	137	N =4940
	B		4.537825	1	47.54054	N=631
	W		6.010186	-39.9482	94.05177	T-bar=7.82884
$Patents_{it}$ Weighted by 7-year citations	O	39.32733	85.92369	1	1460	N=4940
	B		38.0221	1	325.973	N=631
	W		64.52015	-283.646	1248.354	T-bar=7.82884
$Sales_{it}$	O	$276,770.30	$3,557,326.00	-$3.94	$120,000,000.00	N=4866
	B		$2,721,706.00	$0.00	$67,200,000.00	N=627
	W		$1,507,908.00	-$48,400,000.00	$52,800,000.00	T-bar=7.76077
$Annual\ Sales_t$	O	$54,700,000.00	$47,700,000.00	$10,237.56	$155,000,000.00	N=4940
	B		$36,000,000.00	$10,237.56	$155,000,000.00	N=631
	W		$38,100,000.00	-$46,400,000.00	$173,000,000.00	T-bar=7.82884
R&D	O	$5,530.57	$20,801.85	$0.02	$453,046.00	N=4527
	B		$16,303.99	$0.10	$188,818.40	N=611
	W		$10,010.18	-$104,136.00	$309,493.60	T-bar=7.40917
$R\&D\ Intesity_{it}$	O	16.67462	441.4469	-90.3846	25684.4	N=4367
	B		361.2019	-16.2359	8666.188	N=593
	W		327.6426	-7442.21	17034.89	T-bar=7.36425
Employment	O	18.61606	37.55868	0.002	361.796	N=4043
	B		29.33049	0.007333	343.9965	n=567
	W		9.478934	-81.3815	117.2185	T-bar=7.13051
$NPL\ Intensity_{it}$	O	0.251009	0.250072	0	1	N=4934
	B		0.175546	0	0.946429	n=631

	W		0.192682	-0.40051	1.144114	T-bar=7.81933
Originality	O	0.837312	0.138998	0	0.981557	N=4833
	B		0.077937	0.394531	0.962883	N=631
	W		0.122307	0.010771	1.231843	T-bar=7.65927
$Original_{it}$	O	0.838273	0.113728	0	0.981557	N=4885
	B		0.074331	0.394531	0.962883	n=631
	W		0.095405	0.034581	1.232804	T-bar=7.74168
Radicalness	O	0.27727	0.225914	0	1	N=4833
	B		0.129694	0	0.701256	n=631
	W		0.196969	-0.3487	1.129558	T-bar=7.65927
$Radical_{it}$	O	0.270516	0.187558	0	1	N=4885
	B		0.123775	0	0.732303	n=631
	W		0.153606	-0.26529	1.121602	T-bar=7.74168

* O=Overall; B=Between; W=Within; Dollar amounts are in millions; C; Employment is in thousands

APPENDIX B – CHAPTER THREE SUPPLEMENTAL INFORMATION

a) Technology Fields of Orange Book Patents

Figure 1b depicts the technology fields of the Orange Book patents. A noticeable issue is that only about 72% of patents are assigned to the "Pharmaceuticals" technology field. About 16 % of the patents have been assigned to the "Organic fine chemistry" field, 7% to the "Medical technology", and 2.43% to the "Biotechnology" fields. While these fields are related to pharmaceutical and drug products, the incidence of some technology fields is not intuitive; hence, taking a closer look is useful. An example of "Electrical machinery, apparatus, energy" patent is US 8269128 titled "Vacuum switch tube" which was submitted for a product with "AEROSOL, FOAM" application.[81] The "Audio-visual technology" patent is US 8021344 titled "Medicament delivery device configured to produce an audible output" which is related to drug application. The "Telecommunications" patent US 8226610 is assigned to class A61M5 "*Devices for bringing media into the body in a subcutaneous, intra-vascular or intramuscular way; Accessories thereof, e.g. filling or cleaning devices, arm rests*" and is titled "Medical injector with compliance tracking and monitoring". An example for "Computer technology" is US 8978966 titled "Dose counters for inhalers, inhalers and methods of assembly thereof", and an example for the "IT methods for management" is US 8731963 "Sensitive drug distribution system and method". The "Optics" patent is called

[81] The exact reason or incentives for listing patents in the Orange Book is not tackled in this study.

"Projection screen" (US 4057323) and was quoted with application for an antidepressant (Bupropion) in the 8th edition of Orange Book (1988). An example of "handling" technology, is US 8122917 "Apparatus and method for dispensing foam". A number of patents are seemingly far removed from what is expected to be in the Orange Book for instance, the "Machine tools" patent is titled "Reversible micromachining locator" (US 5944329); the "Engines, pumps, turbines" patents are titled "Apparatus for controlling rotational speed of prime mover of construction machine" (US 4955344) and "Component arrangement for outboard motor" (US 6062927); and "other special machines" patent is "Arrow mounted self-retracting sight" (US 4105209).

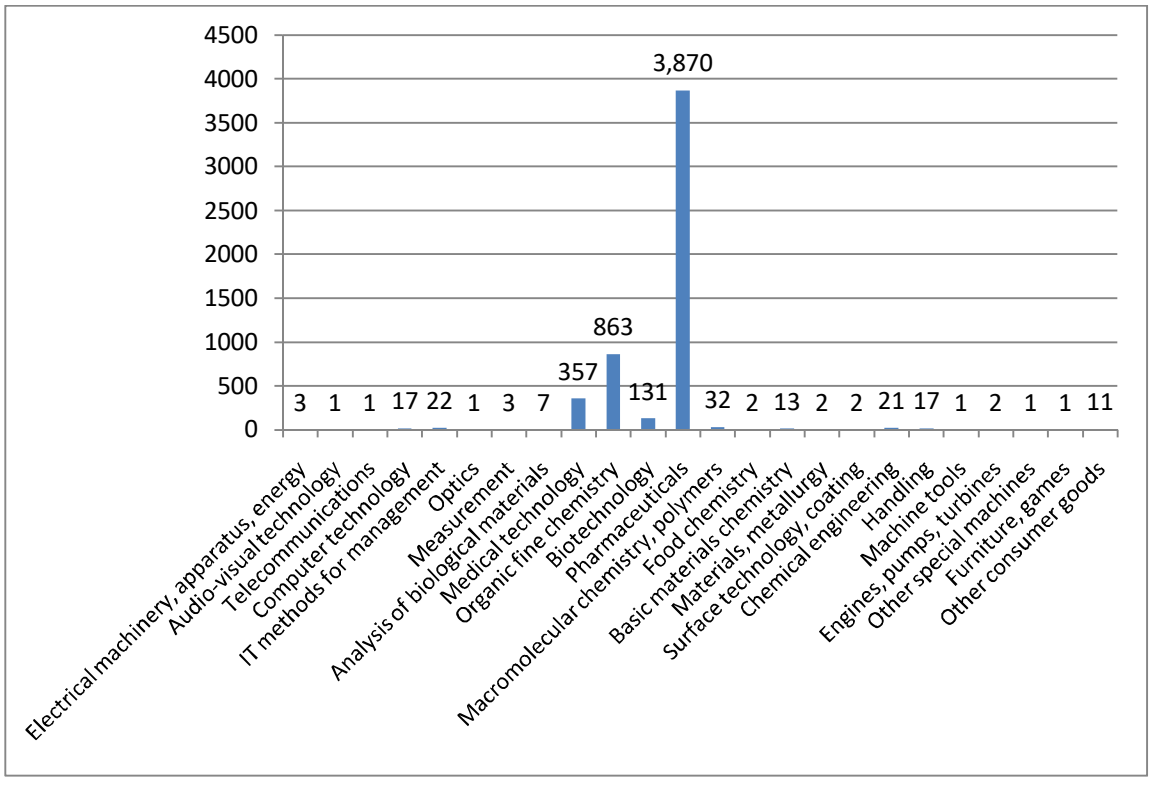

Figure 1b Orange Book Patents by Technology Field (Total= 5381)

The “Furniture, games” patent is more intuitively titled “Balance and coordination exercise device” (US 4828251). Finally, examples of the “Other consumer goods” field are “Inhaler device” (US 8474447), and “Nicotine dispenser with polymeric reservoir of nicotine” (US 4800903).

APPENDIX C – CHAPTER FOUR SUPPLEMENTAL INFORMATION

Survey Instrument

GEORGE MASON UNIVERSITY | Schar School of Policy and Government

Pharmaceutical Innovation System Survey

INFORMED CONSENT FORM

RESEARCH PROCEDURES

This research is being conducted for a PhD dissertation on the drivers and barriers of innovation in drug discovery. If you agree to participate, you will be asked to answer a series of questions on factors impacting innovation. The survey will take about 20-30 minutes. No personal or proprietary information is asked for.

RISKS

There are no foreseeable risks for participating in this research.

BENEFITS

There are no benefits to you as a participant other than to further research in drug discovery innovation.

CONFIDENTIALITY

The data in this study will be confidential. The survey is anonymous and names or individual identifiers will neither be collected nor linked to responses. While it is understood that no computer transmission can be perfectly secure, reasonable efforts will be made to protect the confidentiality of your transmission.

PARTICIPATION

Your participation is voluntary, and you may withdraw from the study at any time and for any reason. If you decide not to participate or if you withdraw from the study, there is no penalty or loss of benefits to which you are otherwise entitled. There are no costs to you or any other party.

CONTACT

This research is being conducted by Alfred Sarkissian, PhD candidate in public policy at the Schar School of Policy and Government, George Mason University. He may be reached at [redacted] for questions or to report a research-related problem. The dissertation chair is Prof. David M. Hart and can be reached at (703) 993-2279. You may contact the George Mason University Institutional Review Board office at 703-993-4121 if you have questions or comments regarding your rights as a participant in the research. This research has been reviewed according to George Mason University procedures governing your participation in this research.

CONSENT

I have read this form, all of my questions have been answered by the research staff, and I agree to participate in this study by pressing the proceed button on this page.

SurveyMonkey Privacy Policy: https://www.surveymonkey.com/mp/policy/privacy-policy/

* 1. Do you agree to the above terms? By clicking Yes, you consent that you are willing to answer the questions in this survey

- ◯ Yes
- ◯ No

GEORGE MASON UNIVERSITY | Schar School of Policy and Government

Pharmaceutical Innovation System Survey

Background

2. What is your main field of specialization? (Choose up to three in order of expertise if applicable or add yours)

	Specialization
1	
2	
3	

Other, please specify in order of expertise if possible.

3. What is your current position? (Optional; select all that applies)

- [ ] Researcher (Corporate Sector)
- [ ] Researcher (Non-profit/Hospital)
- [ ] R&D manager (Corporate Sector)
- [ ] R&D manager (Non-profit/Hospital)
- [ ] Academic
- [ ] Entrepreneur
- [ ] Consultant

Other, please specify.

GEORGE MASON UNIVERSITY | Schar School of Policy and Government

Pharmaceutical Innovation System Survey

Drivers of Innovation

Prior research has identified a number of factors conducive to innovation in the pharmaceutical sector:

a. R&D investment
b. Basic science (e.g. scientific publications)
c. Skilled R&D scientists
d. Good R&D management
e. Collaborative R&D with other outside entities (e.g. universities)
f. The diversity of knowledge available to the inventors (e.g. enzymology, toxicology, etc.)
g. The depth of specialized knowledge available to the inventors
h. Market size of the drug

4. What drives new drug substance discovery in your field of specialization? (Please pick five most influential factors in order of importance. You can add your own options.)

	Driver
1	
2	
3	
4	
5	

Other, Please specify and rank (e.g. 1. x 2. y 3. z).

4

GEORGE MASON UNIVERSITY | Schar School of Policy and Government

Pharmaceutical Innovation System Survey

Innovation System Structure

This section tries to build a list of the most influential players and institutions in your field based on your judgment.

5. List three most important public R&D funding entities (e.g. NIH)

1
2
3

6. List three most important private R&D funding entities (i.e. funding outside R&D such as a venture capital company or big companies)

1
2
3

7. List three most important producers of basic research (e.g. a specific university)

1
2
3

8. List three most enabling legislations or regulations

1
2
3

9. List three most burdensome legislations or regulations (i.e. hindering innovation)

1
2
3

GEORGE MASON UNIVERSITY | Schar School of Policy and Government

Pharmaceutical Innovation System Survey

Innovation System Structure

This section tries to build a list of the most influential players and institutions in your field based on your judgment.

5. List three most important public R&D funding entities (e.g. NIH)

1
2
3

6. List three most important private R&D funding entities (i.e. funding outside R&D such as a venture capital company or big companies)

1
2
3

7. List three most important producers of basic research (e.g. a specific university)

1
2
3

8. List three most enabling legislations or regulations

1
2
3

9. List three most burdensome legislations or regulations (i.e. hindering innovation)

1
2
3

GEORGE MASON UNIVERSITY | Schar School of Policy and Government

Pharmaceutical Innovation System Survey

New Company Formation

New companies are important elements of the innovation system because, for instance, they commercialize new technologies or create new markets.
We define new companies as independent, for-profit entities that were established no more than 10 years ago.

10. List top three new companies in your area of expertise and reason for inclusion in this list (e.g. for new technology development; for new market creation; etc.)

1

2

3

GEORGE MASON UNIVERSITY | Schar School of Policy and Government

Pharmaceutical Innovation System Survey

R&D Efficiency

11. What is your overall judgment about corporate R&D spending and new drug approval trends?

* Please also mention your time frame
* R&D spending is meant to be inflation adjusted
* New drug approvals include small molecules as well as new biologics

	Trend	Time Frame
R&D spending		
New drug approvals		

Comments:

GEORGE MASON UNIVERSITY | Schar School of Policy and Government

Pharmaceutical Innovation System Survey

Barriers to Innovation

Existing studies have identified a number of factors hindering new drug substance discovery:

a. Availability of good drugs for many diseases
b. Over-cautious regulation for safety
c. Inflated R&D wages
d. Designing drug substances with a single or narrow therapeutic benefits
e. Complex clinical trials
f. Reduced quality of published science
g. Patented or proprietary research tools
h. Lack of inter-organizational collaboration in R&D
i. Companies pursuing the same drug targets

12. What do you think is a hindrance to new drug substance discovery? (Please pick five most influential factors in order of importance. You can add your own options.)

	Barrier
1	
2	
3	
4	
5	

Other, Please specify and rank (e.g. 1. x 2. y 3. z).

9

GEORGE MASON UNIVERSITY | Schar School of Policy and Government

Pharmaceutical Innovation System Survey

Your Recommendations

13. Please share any other comments relevant to boosting innovation and lifting barriers to innovation.

GEORGE MASON UNIVERSITY | Schar School of Policy and Government

Pharmaceutical Innovation System Survey

Generalization

14. Do you think your answers apply to other therapeutic areas as well?

- ◯ Yes. To a large extent.
- ◯ Yes. To some extent.
- ◯ No.

Comments:

11

Survey Results

a) Sources of public and private R&D funds.

Table 1c Most Important Public R&D Funding Entities (59 responses)

Entity	Frequency
NIH	58
NSF	22
DOD	16
Foundations/Individual Trusts	7
Gates Foundation	7
Welcome Trust	5
Biomedical Advanced Research and Development Authority (BARDA)	3
DARPA	3
Innovative Medicines Initiative (IMI)	3
National Cancer Institute's (NCI)	3
SBIR/STTR	3
State Economic Development/Entities	3
Cancer Research UK	2
Disease-specific Foundations/Organizations	2
Howard Hughes Medical Institute	2
Medical Research Council (MRC)	2
Alzheimer's Foundation	1
American Diabetes Association (ADA)	1
Department of Veterans Affairs	1
FDA	1
fNIH	1
Funding Agencies	1
Grant Organizations e.g. Gates Foundation	1
Juvenile Diabetes Research Foundation (JDRF)	1
NSA	1
Patient Advocacy Organizations	1
Philanthropy	1
Public Universities	1
Small Business Administration	1
University Seed Funds	1
Venture Philanthropy Partners	1
WHO partnerships - IAVI, GAVI, MMV	1

* A few items listed as example were separated and entered as an independent entry e.g. “foundation such as Gates” was separated into Foundations and Gates Foundation

Table 2c Most Important Private R&D Funding Entities (54 responses)

Entity	Frequency
Venture Capital	25
Big Pharma	24
Gates Foundation	10
Angel Investors (i.e. individuals investing own funds)	9
Foundations	9
JVs/Collaborations/Big Pharma Partnerships/Private Collaborations	4
Biotech	4
Johnson & Johnson (J&J)	4
Philanthropies	4
Atlas Ventures	3
High Net Worth/Individual investors	3
Howard Hughes Medical Institute (HHMI)	3
Michael J. Fox Foundation	3
Early Stage Venture/Early stage venture capital companies	2
Flagship Pioneering	2
Juvenile Diabetes Research Foundation (JDRF)	2
Medium-sized Pharmaceutical/biotech companies	2
Private Equity	2
Third Rock Ventures	2
Venrock	2
Welcome Trust	2
Alphabet (Google)	1
Bristol-Myers Squibb	1
Cantor Fitzgerald	1
Celgene	1
CurePSP (An organization offering services for neurodegenerative diseases to patients, researchers, etc.)	1
Eli Lilly	1
Kleiner Perkins Caufield & Byers	1
Medicines for Malaria Venture (MMV)	1
NEA	1
Non-profit Research Institutions	1
Novartis	1
Novo Nordisk	1
OrbiMed Advisors	1
Patient Advocacy Foundations	1
Pfizer	1
Pharma Funds	1
Pharma Licensees	1

Robert Wood Johnson Foundation (RWJF)	1
SV Life Science Advisors	1
Takeda	1
Tau Consortium	1

b) Most Enabling Legislations or Regulations

One question asked respondents to list three most enabling legislations. 33 responders had some input for this question and in total 31 legislations or Acts were identified from the exercise (table 3c). A responder posed the rhetorical question of if there were any enabling legislations. The top ranked is the Bayh–Dole Act. FDA core legislation, Hatch-Waxman Act, and the "Patent Law" are tied for the second rank. The emphasis on "Patent Law", i.e. intellectual property protection, by responders corroborates the focus of previous essays on patent data and the importance of patents in the pharmaceutical industry. Other aforementioned legislations were identified in the literature and this exercise further validates their importance.

A few surprising mentions are present: two apparently foreign referenced legislations (i.e. European Medicines Agency (EMA) and Health Canada), the international ICH council, and "Foreign Visas". ICH was created in 1990 with the mission of greater harmonization of worldwide drug registration[82]. These responses may indicate the global nature of the pharmaceutical industry, immediate foreign base of a few respondents inadvertently included in the pool of respondents[83]. A few phrases were ambiguous (e.g. "IPO treaty" or "Health Care Law") and could not be reconciled with

[82] http://www.ich.org
[83] As mentioned earlier US-based respondents were the intended targets.

any legislation. The phrase of many “generics ruling by Supreme Court” may be a fruitful lead but is beyond the scope of this study.

Table 3c Three Most Enabling Legislations / Regulations

Act/Regulation	Frequency
Bayh–Dole Act	8
FDA	6
Hatch-Waxman Act	6
Patent Law	6
NIH Funding Bills	5
Orphan Drug Act	5
Small Business Innovation Research (SBIR) program	5
21st Century Cures Act	4
Biologics Price Competition and Innovation Act (Biosimilars Act)	2
European Medicines Agency (EMA)	2
Prescription Drug User Fee Act of 1992 (PDUFA)	2
Accelerated Approval Program	1
American Innovation Act	1
Antibiotic Development to Advance Patient Treatment (ADAPT) Act	1
Breakthrough Therapy Designations	1
Foreign Visas	1
GAIN Act	1
Guidance on the Codevelopment of Two or More New Investigational Drugs for Use in Combination	1
Health Canada	1
“Health Care Law”	1
International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH)	1
“IPO treaty”	1
Jumpstart Our Business Startups (JOBS) Act	1
Myriad generics ruling by supreme court	1
NIH clinical trials publication	1
NIH training grants	1
NSF funding	1
Pharmaceuticals and Medical Devices Agency (PMDA)	1
R&D tax credit	1
State funds for early stage research, e.g. Ben Franklin funds	1
Tax Free IND	1

* Those phrases in quotes could not be linked to a known regulation or Act; ICH was only mentioned by acronym and there is a slight chance that it is misconstrued here.

c) Most Burdensome Legislations or Regulations

One question asked respondents to identify three most burdensome legislations or regulations. 30 responders offered some input for this question. Some issues are very general and some more specific (table 4c). The foreign; (i.e. EU, Canada and "Price Controls Abroad") regulations raised are also of special note. As mentioned before, this could indicate a few respondents having primary experience abroad and/or the interconnected nature of the innovation systems. The latter point is visible by the presence of multinational pharmaceutical firms in the innovation system.

The top ranked burdensome regulation raised is broad FDA rules and regulations with two respondent referring to it as "21 CFR" which stands for the "Code of Federal Regulations" Title 21 governing the FDA, the Drug Enforcement Administration (DEA), and the Office of National Drug Control Policy (ONDCP).[84] While the respondents were not asked for details of why the regulations are burdensome, these may be symptomatic of discontent with a slow bureaucracy. "Patent Law" related issues are second highest ranked in the list. Moreover, the issue of copyright and "confidentiality in competitive environment" are other issues that are related to intellectual property. This is not surprising as intellectual property protection is a double-edged sword with positive and negative effects on innovation. One respondent aptly raised the "IP perverse incentives". These references indicate the potential abuse of intellectual property issues in the pharmaceutical sector.

[84] https://en.wikipedia.org/wiki/Title_21_of_the_Code_of_Federal_Regulations

The Hatch-Waxman and Bayh–Dole Acts have been mentioned as both an enabling and a burdensome legislation. Given the complex incentive mechanism for exclusivity and the abuse potential, this is an expected observation. One responder mentioned "Limited NCE Exclusivity Period" indicating that striking the right balance between exclusivity and subsequent innovation by other parties is important.

Clinical trial related regulations also figure prominently especially if we pool mentions of "Animal Safety Laws" and "Clinical Trials for FDA Approval". Stem cell research regulations were mentioned twice. Given the apparent popularity of gene-editing technologies (i.e. from list of top firms mentioned by responders), stem cell research restriction would naturally be raised here.

There are traces of financing and corporate law issues in some responses (i.e. Corporate Tax Structure, Finance Law, and Sarbanes–Oxley Act). Section 806 of the Sarbanes-Oxley Act is aimed at protecting corporate whistleblowers and has been used in cases related to pharmaceutical firms (e.g. Bio-Rad Laboratories Inc. for violations of the Foreign Corrupt Practices Act, and Progenics Pharmaceuticals Inc. for inaccurate representations about the results of a clinical trial).[85] Exploring the contents of these legislation or the reason why these were deemed burdensome can be avenues for future research.

[85] https://en.wikipedia.org/wiki/Sarbanes%E2%80%93Oxley_Act

Table 4c Three Most Burdensome Legislations / Regulations

Barrier	Frequency
FDA	9
Patent Law	5
European Laws (related to reimbursement of drug costs; EMA; EUMEA)	3
Hatch-Waxman Act	3
Medicaid/Medicare	3
Animal Safety Laws	2
Clinical Trials for FDA Approval	2
Stem Cell Research Limitations	2
Slow funding/Underfunding at/for NIH (one comment for slow funding other for underfunding)	2
American Innovation Act	1
Anything Sans Safety and Efficacy (responder's original comment was that safety is the only real priority and safety and efficacy are the foundation of regulation; hence, implying that other regulation is burdensome)	1
Banning Medicare drug price negotiation	1
Bayh–Dole Act	1
Price Competition and Innovation Act (Biosimilars Act)	1
Confidentiality in competitive environment	1
Conflict of interest regulations	1
Copyright (i.e. copyright by journal vs. open access)	1
Corporate Tax Structure	1
Finance Law	1
Flawed Center for Scientific Reviews and Lack of Recourse	1
Health Canada	1
Health Insurance Portability and Accountability Act (HIPPA)	1
Limited NCE Exclusivity Period	1
Limits on Collaboration	1
NIH Rules	1
Physician Payments Sunshine Act (PPSA)	1
Price Controls Abroad	1
Regulations that discourage the adoption of new manufacturing technologies for approved drugs	1
Requirements for large patient safety data bases	1
Sarbanes–Oxley (SOX) Act (2002 Act meant to protect investors against fraudulent corporate accounting)	1
Stanford Vs Roche	1
Underfunding of SBIR program by Congress	1

New Firms Identified By Survey Respondents

One question in the survey asked respondent to "*List top three new companies*" in their area of expertise and the reason for inclusion in this list "*(e.g. for new technology development; for new market creation; etc.)*". A "*new company*" was defined as "*independent, for-profit*" entity that was "*established no more than 10 years ago*". Overall 87 entities were listed. Not all names came with the reason for inclusion; moreover, not all names met the instructions (one research center and one university). Table 5c depicts the firms cited more than once. Most are noted for new technology possibly because new technologies and products are more noticeable than creating new markets. Four out of the eight firms do not meet the age definition in the question, indicating some respondents did not heed the instructions; however, we may argue that probably respondents considered these firms noteworthy in some respect. Among the remaining four, the top cited is CRISPR[86] Therapeutics, established in 2014 and is a gene-editing company working on transformative gene-based medicines based on own proprietary CRISPR/Cas9 gene-editing platform[87]. CRISPR-Cas9 is a technology that enables genome editing by "*removing, adding or altering sections of the DNA sequence*". It is supposed to be "faster, cheaper and more accurate" than previous techniques and has

[86] Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)
[87] "CRISPR." Accessed September 10, 2017. http://www.crisprtx.com/about-us/overview.php.

a broad range of potential applications.[88] Editas Medicine is also focused on gene-editing especially exploring repairing broken genes and mutations in DNA.[89]

Kite Pharma is focused on curing cancer by developing engineered cell therapies expressing either a chimeric antigen receptor (CAR) or a T cell receptor (TCR). They hope to bring a paradigm shift in cancer treatment by their dual platform.[90] CAR T cells recognize proteins expressed on the "*surface"* of the cancer cell whereas TCRs can "recognize tumor-specific proteins on the *inside* of cells".[91] From these few cases, it seems gene-editing work is popular with the respondents.

Finally, "Third Rock Ventures" was mentioned twice in by the respondents but it is a venture capital firm rather than a traditional drug discovery firm.

[88] "What Is CRISPR-Cas9? | Facts | Yourgenome.org." Accessed September 10, 2017. https://www.yourgenome.org/facts/what-is-crispr-cas9.
[89] "Company Overview | Editas Medicine." Accessed September 10, 2017. http://www.editasmedicine.com/company-overview.
[90] "Kite Pharma." Accessed September 10, 2017. http://www.kitepharma.com/.
[91] "T Cell Receptor Technology (TCRs) | Juno Therapeutics." Accessed September 10, 2017. https://www.junotherapeutics.com/the-science/tcr-technology/.

Table 5c Top Three New Companies Identified in the Survey More Than Once

Name	Reason	Frequency	Established
CRISPR Therapeutics	New technology (gene editing)	5	2014
Genentech	New technology (drug pipeline; innovation) (Long term fundamental research) (Excellence in science and technology)	4	1976
Celgene		3	1986
Editas Medicine	New technology (leaders in genome editing)(gene editing)	4	2013
Gilead		2	1987
GSK (Glaxo)	New technology (Strimvelis, Cell & gene therapy)	2	2000
Kite Pharma	New technology (expanding T-cell therapy)	2	2009
Third Rock Ventures	New companies in new areas; Innovative funding models	2	2007

* Statements in parentheses represent what respondents mentioned

REFERENCES

Acemoglu, Daron, and Joshua Linn. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry." *The Quarterly journal of economics* 119, no. 3 (2004): 1049-1090.

Acs, Zoltan J., and David B. Audretsch. "Innovation, Market Structure, and Firm Size." *The Review of Economics and Statistics* 69, no. 4 (November 1987): 567. doi:10.2307/1935950.

Adams, C. P., and V. V. Brantner. "Estimating The Cost Of New Drug Development: Is It Really $802 Million?" *Health Affairs* 25, no. 2 (March 1, 2006): 420–28. doi:10.1377/hlthaff.25.2.420.

Ahuja, Gautam, and Curba Morris Lampert. "Entrepreneurship in the Large Corporation: A Longitudinal Study of How Established Firms Create Breakthrough Inventions." *Strategic Management Journal* 22, no. 6–7 (2001): 521–43. doi:10.1002/smj.176.

Ahuja, Gautam, Curba Morris Lampert, and Vivek Tandon. "1 Moving Beyond Schumpeter: Management Research on the Determinants of Technological Innovation." *The Academy of Management Annals* 2, no. 1 (January 2008): 1–98. doi:10.1080/19416520802211446.

Andres, Charles. "Overcoming Restriction Requirements On Pharma Patents." LAW360. August 2015. Expert Analysis Article. Accessed October 4, 2017. https://www.wsgr.com/publications/PDFSearch/law360-0815.pdf

Archontopoulos, Eugenio, Dominique Guellec, Niels Stevnsborg, Bruno van Pottelsberghe de la Potterie, and Nicolas van Zeebroeck. "When Small Is Beautiful: Measuring the Evolution and Consequences of the Voluminosity of Patent Applications at the EPO." *Information Economics and Policy* 19, no. 2 (June 2007): 103–32. doi:10.1016/j.infoecopol.2007.01.001.

Ardito, Lorenzo, Antonio Messeni Petruzzelli, and Umberto Panniello. "Unveiling the Breakthrough Potential of Established Technologies: An Empirical Investigation in

the Aerospace Industry." *Technology Analysis & Strategic Management* 28, no. 8 (September 13, 2016): 916–34. doi:10.1080/09537325.2016.1180356.

Arts, Sam, Francesco Paolo Appio, and Bart Van Looy. "Inventions Shaping Technological Trajectories: Do Existing Patent Indicators Provide a Comprehensive Picture?" *Scientometrics* 97, no. 2 (November 2013): 397–419. doi:10.1007/s11192-013-1045-1.

Austin, David H. "Research and Development in the Pharmaceutical Industry." Congress of the United States, Congressional Budget Office, 2006. Accessed April 6, 2017. http://www.cbo.gov/sites/default/files/cbofiles/ftpdocs/76xx/doc7615/10-02-drugr-d.pdf.

Azoulay, Pierre, Joshua S. Graff Zivin, Danielle Li, and Bhaven N. Sampat. "Public R&D Investments and Private-Sector Patenting: Evidence from NIH Funding Rules." National Bureau of Economic Research, 2015. http://www.nber.org/papers/w20889.

Baba, Yasunori, and John P. Walsh. "Embeddedness, Social Epistemology and Breakthrough Innovation: The Case of the Development of Statins." *Research Policy* 39, no. 4 (May 2010): 511–22. doi:10.1016/j.respol.2010.01.016.

Ballentine, Carol. "Sulfanilamide Disaster." FDA Consumer magazine Issue June 1981. Accessed June 28, 2017. https://www.fda.gov/aboutfda/whatwedo/history/productregulation/sulfanilamidedisaster/default.htm.

Barkoff, Aaron ."Editorial: The Orange Book Should Include Process Patents or Be Eliminated | Patently-O." May 4, 2006. Accessed October 4, 2017. https://patentlyo.com/patent/2006/05/editorial_the_o.html.

Bennani, Youssef L. "Drug Discovery in the next Decade: Innovation Needed ASAP." *Drug Discovery Today* 16, no. 17–18 (2011): 779–92. doi:10.1016/j.drudis.2011.06.004.

Bergek, Anna, Staffan Jacobsson, Bo Carlsson, Sven Lindmark, and Annika Rickne. "Analyzing the Functional Dynamics of Technological Innovation Systems: A Scheme of Analysis." *Research Policy* 37, no. 3 (April 2008): 407–29. doi:10.1016/j.respol.2007.12.003.

Bleda, Mercedes, and Pablo del Río. "The Market Failure and the Systemic Failure Rationales in Technological Innovation Systems." *Research Policy* 42, no. 5 (June 2013): 1039–52. doi:10.1016/j.respol.2013.02.008.

Boadi, Kwame. “Erosion of funding for the National Institutes of Health threatens US leadership in biomedical research.” Center for American Progress 25 (2014). https://cdn.americanprogress.org/wp-content/uploads/2014/03/NIHinvestment.pdf.

Branstetter, Lee, and Yoshiaki Ogura. “Is Academic Science Driving a Surge in Industrial Innovation? Evidence from Patent Citations.” National Bureau of Economic Research, 2005. http://www.nber.org/papers/w11561.

Breschi, Stefano, Franco Malerba, and Luigi Orsenigo. “Technological Regimes and Schumpeterian Patterns of Innovation.” *The Economic Journal* 110, no. 463 (2000): 388–410.

Burk, Dan L., and Mark A. Lemley. “Policy Levers in Patent Law.” *Virginia Law Review* 89, no. 7 (November 2003): 1575. doi:10.2307/3202360.

Callaert, Julie, Bart Van Looy, Arnold Verbeek, Koenraad Debackere, and Bart Thijs. “Traces of Prior Art: An Analysis of Non-Patent References Found in Patent Documents.” *Scientometrics* 69, no. 1 (2006): 3–20.

Cardinal, Laura B., and Donald E. Hatfield. “Internal Knowledge Generation: The Research Laboratory and Innovative Productivity in the Pharmaceutical Industry.” *Journal of Engineering and Technology Management* 17, no. 3 (2000): 247–271.

Carlsson, Bo. “Internationalization of Innovation Systems: A Survey of the Literature.” *Research Policy* 35, no. 1 (February 2006): 56–67. doi:10.1016/j.respol.2005.08.003.

Carlsson, Bo, Staffan Jacobsson, Magnus Holmén, and Annika Rickne. “Innovation Systems: Analytical and Methodological Issues.” *Research Policy* 31, no. 2 (2002): 233–245.

Cockburn, Iain, and Rebecca Henderson.“The Economics of Drug Discovery.” Chapter in Ralph Landau, Basil Achilladelis and Alexander Scriabine (eds.) Pharmaceutical Innovation, Chemical Heritage Press, Philadelphia PA, 1999, pp. 308-331.

Criscuolo, Paola. “The’home Advantage’effect and Patent Families. A Comparison of OECD Triadic Patents, the USPTO and the EPO.” *Scientometrics* 66, no. 1 (2006): 23–41.

Declerck, Paul. “Biologicals and Biosimilars: A Review of the Science and Its Implications.” *GaBI-Generics and Biosimilars Initiative Journal* 1, no. 1 (2012): 13–16.

Ding Min, Jehoshua Eliashberg , and Stefan Stremersch. “The Pharmaceutical Industry: Specificity, Challenges, and What You Can Learn from this Book” In *Innovation and Marketing in the Pharmaceutical Industry,* edited by Min Ding, Jehoshua Eliashberg, and Stefan Stremersch, pp. 1-18. Springer New York, 2014a.

Ding Min, Jehoshua Eliashberg, and Stefan Stremersch. “Portfolio Management in New Drug Development” In *Innovation and Marketing in the Pharmaceutical Industry,* edited by Min Ding, Jehoshua Eliashberg, and Stefan Stremersch, pp. 83-118. Springer New York, 2014b.

DiMasi, Joseph A., and Cherie Paquette. “The Economics of Follow-on Drug Research and Development.” *Pharmacoeconomics* 22, no. 2 (2004): 1–14.

Doloreux, David, and Saeed Parto. “Regional Innovation Systems: Current Discourse and Unresolved Issues.” *Technology in Society* 27, no. 2 (April 2005): 133–53. doi:10.1016/j.techsoc.2005.01.002.

Dunlap-Hinkler, Denise, Masaaki Kotabe, and Ram Mudambi. “A Story of Breakthrough versus Incremental Innovation: Corporate Entrepreneurship in the Global Pharmaceutical Industry.” *Strategic Entrepreneurship Journal* 4, no. 2 (June 11, 2010): 106–27. doi:10.1002/sej.86.

Eckelmann, Oliver. “The Business Case for Sustainability - Pharmaceutical Industry Sector Report”. Working Paper Series, Forum for Corporate Sustainability Management International Institute for Management Development (IMD) (2003)

Edquist, Charles. “Reflections on the Systems of Innovation Approach.” *Science and Public Policy* 31, no. 6 (2004): 485–489.

Edquist, Charles. 2005. ‘Systems of innovation – perspectives and challenges’, in J. Fagerberg, D.C. Mowery and R.R. Nelson (eds), *The Oxford Handbook of Innovation*, Oxford: Oxford University Press, pp. 181–208.

Edquist, Charles, and Leif Hommen, eds. 2008. “Comparing national systems of innovation in Asia and Europe: theory and comparative framework.” Edquist, C., and Hommen L. Editors. In *Small Country Innovation Systems: Globalization, Change and Policy in Asia and Europe*. Cheltenham, UK ; Northampton, MA: Edward Elgar.

EFPIA. “The Pharmaceutical Industry in Figures”. Brussels, Belgium: European Federation of Pharmaceutical Industries and Associations (2017). Accessed August

8, 2017. https://www.efpia.eu/media/219735/efpia-pharmafigures2017_statisticbroch_v04-final.pdf

Egli Florian, Nick Johnstone, and Menon, Carlo. "Identifying and Inducing Breakthrough Inventions." OECD Science, Technology and Industry Working Papers, June 24, 2015. doi:10.1787/5js03zd40n37-en.

Fabrizio, Kira R. "Absorptive Capacity and the Search for Innovation." *Research Policy* 38, no. 2 (March 2009): 255–67. doi:10.1016/j.respol.2008.10.023.

Feldman, M. P., and J. W. Yoon. "An Empirical Test for General Purpose Technology: An Examination of the Cohen-Boyer rDNA Technology." *Industrial and Corporate Change* 21, no. 2 (April 1, 2012): 249–75. doi:10.1093/icc/dtr040.

Fernandez, Dennis S., and James T. Huie. "Balancing US patent and FDA approval processes: strategically optimizing market exclusivity." *Drug discovery today* 9, no. 12 (2004): 509-512.

Fleming, Lee. "Recombinant Uncertainty in Technological Search." *Management Science* 47, no. 1 (2001): 117–132.

Fontana, Roberto, Alessandro Nuvolari, Hiroshi Shimizu, and Andrea Vezzulli. "Schumpeterian Patterns of Innovation and the Sources of Breakthrough Inventions: Evidence from a Data-Set of R&D Awards." *Journal of Evolutionary Economics* 22, no. 4 (September 2012): 785–810. https://doi.org/10.1007/s00191-012-0287-z.

Freeman, Chris. "Continental, National and Sub-National Innovation Systems—complementarity and Economic Growth." *Research Policy* 31, no. 2 (2002): 191–211.

Geysen, H. Mario, Frank Schoenen, David Wagner, and Richard Wagner. "A Guide to Drug Discovery: Combinatorial Compound Libraries for Drug Discovery: An Ongoing Challenge." *Nature Reviews Drug Discovery* 2, no. 3 (March 2003): 222–30. doi:10.1038/nrd1035.

Gittelman, Michelle, and Bruce Kogut. "Does Good Science Lead to Valuable Knowledge? Biotechnology Firms and the Evolutionary Logic of Citation Patterns." *Management Science* 49, no. 4 (April 2003): 366–82. doi:10.1287/mnsc.49.4.366.14420.

Golec, Joseph, and John Vernon. "Measuring US Pharmaceutical Industry R&D Spending." *PharmacoEconomics* 26, no. 12 (2008): 1005–1017.

Grabowski, Henry, and Y. Richard Wang. "Do Faster Food and Drug Administration Drug Reviews Adversely Affect Patient Safety? An Analysis of the 1992 Prescription Drug User Fee Act." *The Journal of Law and Economics* 51, no. 2 (2008): 377–406.

Graves, Samuel B., and Nan S. Langowitz. "Innovative Productivity and Returns to Scale in the Pharmaceutical Industry." *Strategic Management Journal* 14, no. 8 (1993): 593–605.

Grove, A. "Rethinking Clinical Trials." *Science* 333, no. 6050 (September 23, 2011): 1679–1679. doi:10.1126/science.1212118.

Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg. "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools." National Bureau of Economic Research, 2001. Accessed April 6, 2017. http://www.nber.org/papers/w8498.

Hall, Bronwyn H., and Rosemarie Ham Ziedonis. "The Patent Paradox Revisited: An Empirical Study of Patenting in the US Semiconductor Industry, 1979-1995." *RAND Journal of Economics*, 2001, 101–128.

Harhoff, Dietmar, and Stefan Wagner. "The Duration of Patent Examination at the European Patent Office." *Management Science* 55, no. 12 (2009): 1969–1984.

Hekkert, M.P., R.A.A. Suurs, S.O. Negro, S. Kuhlmann, and R.E.H.M. Smits. "Functions of Innovation Systems: A New Approach for Analysing Technological Change." *Technological Forecasting and Social Change* 74, no. 4 (May 2007): 413–32. doi:10.1016/j.techfore.2006.03.002.

Hellsmark, Hans, and Staffan Jacobsson. "Opportunities for and Limits to Academics as System builders—The Case of Realizing the Potential of Gasified Biomass in Austria." *Energy Policy* 37, no. 12 (December 2009): 5597–5611. doi:10.1016/j.enpol.2009.08.023.

Henderson, Rebecca, Luigi Orsenigo, and Gary P. Pisano. "The pharmaceutical industry and the revolution in molecular biology: interactions among scientific, institutional, and organizational change." *Sources of industrial leadership: studies of seven industries* (1999): 267-311.

Higgins, Matthew J., and Daniel Rodriguez. "The Outsourcing of R&D through Acquisitions in the Pharmaceutical Industry." *Journal of Financial Economics* 80, no. 2 (May 2006): 351–83. doi:10.1016/j.jfineco.2005.04.004.

Jensen, Elizabeth J. “Research Expenditures and the Discovery of New Drugs.” *The Journal of Industrial Economics* 36, no. 1 (September 1987): 83. doi:10.2307/2098598.

Johnson, Marvin M. “One Point of View: Finding Creativity in a Technical Organization.” *Research-Technology Management* 39, no. 5 (1996): 9-11.

Johnstone, Nick, Carlo Menon, and Florian Egli. “Identifying and Inducing Breakthrough Inventions.” OECD Science, Technology and Industry Working Papers, June 24, 2015. http://www.oecd-ilibrary.org/science-and-technology/identifying-and-inducing-breakthrough-inventions_5js03zd40n37-en.

Kaitin, K I. “Deconstructing the Drug Development Process: The New Face of Innovation.” *Clinical Pharmacology & Therapeutics* 87, no. 3 (March 2010): 356–61. doi:10.1038/clpt.2009.293.

Kaitin, K I, and J A DiMasi. “Pharmaceutical Innovation in the 21st Century: New Drug Approvals in the First Decade, 2000–2009.” *Clinical Pharmacology & Therapeutics* 89, no. 2 (February 2011): 183–88. doi:10.1038/clpt.2010.286.

Kaplan, Sarah, and Keyvan Vakili. “The Double-Edged Sword of Recombination in Breakthrough Innovation: The Double-Edged Sword of Recombination.” *Strategic Management Journal* 36, no. 10 (October 2015): 1435–57. doi:10.1002/smj.2294.

Keijl, S., V.A. Gilsing, J. Knoben, and G. Duysters. “The Two Faces of Inventions: The Relationship between Recombination and Impact in Pharmaceutical Biotechnology.” *Research Policy* 45, no. 5 (June 2016): 1061–74. doi:10.1016/j.respol.2016.02.008.

Kelley, Donna J., Abdul Ali, and Shaker A. Zahra. “Where Do Breakthroughs Come From? Characteristics of High-Potential Inventions: Characteristics of High-Potential Inventions.” *Journal of Product Innovation Management* 30, no. 6 (November 2013): 1212–26. doi:10.1111/jpim.12055.

Kesselheim, Aaron S. “An Empirical Review of Major Legislation Affecting Drug Development: Past Experiences, Effects, and Unintended Consequences.” *Milbank Quarterly* 89, no. 3 (2011): 450–502.

Khanna, R., I. Guler, and A. Nerkar. “Fail Often, Fail Big, and Fail Fast? Learning from Small Failures and R&D Performance in the Pharmaceutical Industry.” *Academy of Management Journal* 59, no. 2 (April 1, 2016): 436–59. doi:10.5465/amj.2013.1109.

Kim, Bongsun, Eonsoo Kim, Douglas J. Miller, and Joseph T. Mahoney. “The Impact of the Timing of Patents on Innovation Performance.” *Research Policy* 45, no. 4 (May 2016): 914–28. doi:10.1016/j.respol.2016.01.017.

Kim, J. H., and A. R. Scialli. “Thalidomide: The Tragedy of Birth Defects and the Effective Treatment of Disease.” *Toxicological Sciences* 122, no. 1 (July 1, 2011): 1–6. doi:10.1093/toxsci/kfr088.

Klein Woolthuis, Rosalinde, Maureen Lankhuizen, and Victor Gilsing. “A System Failure Framework for Innovation Policy Design.” *Technovation* 25, no. 6 (June 2005): 609–19. doi:10.1016/j.technovation.2003.11.002.

Lanjouw, Jean O., and Mark Schankerman. “Characteristics of Patent Litigation: A Window on Competition.” *The RAND Journal of Economics* 32, no. 1 (2001): 129. doi:10.2307/2696401.

Lauer, Michael. “How Many Researchers? | NIH Extramural Nexus.” May 31, 2016. Accessed August 15, 2017. https://nexus.od.nih.gov/all/2016/05/31/how-many-researchers/.

Laursen, Keld, and Ammon Salter. “Open for Innovation: The Role of Openness in Explaining Innovation Performance among U.K. Manufacturing Firms.” *Strategic Management Journal* 27, no. 2 (February 2006): 131–50. doi:10.1002/smj.507.

Lazonick, William, and Öner Tulum. “US Biopharmaceutical Finance and the Sustainability of the Biotech Business Model.” *Research Policy* 40, no. 9 (November 2011): 1170–87. doi:10.1016/j.respol.2011.05.021.

Lerner, Joshua. “The Importance of Patent Scope: An Empirical Analysis.” *The RAND Journal of Economics* 25, no. 2 (1994): 319. doi:10.2307/2555833.

Lichtenberg, Frank, and Suchin Virabhak. “Using Patents Data to Map Technical Change in Health-Related Areas,” 2002. http://www.oecd-ilibrary.org/science-and-technology/using-patents-data-to-map-technical-change-in-health-related-areas_522485718871.

Liebenau, Jonathan. *Innovation Trends: Prioritising Emerging Technologies Shaping the UK to 2017*. UK Department of Trade and Industry London, 2007. http://www.lse.ac.uk/businessAndConsultancy/LSEEnterprise/pdf/file39041.pdf.

Lo, Andrew W., and Sourya V. Naraharisetti. “New Financing Methods in the Biopharma Industry: A Case Study of Royalty Pharma, Inc.” *Inc.(December 15, 2013)*, 2013. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2330089.

Lundvall, Bengt-Åke; Jan Vang; Christina Chaminade, and Joseph K. Joseph. 2009. Innovation system research and developing countries. Lundvall B-Å, Chaminade C, Joseph JK, Vang J, editors. In Handbook of innovation systems and developing countries: Building domestic capabilities in a global setting. Cheltenham: Edward Elgar Publishing, Incorporated. pp. 1-32.

Lynne G. Zucker and Michael R. Darby, Connecting Outcome Measures in Entrepreneurship Technology and Science (COMETS) database, public use version COMETS 2.0 at http://www.kauffman.org/comets/, published to the web February 26, 2014.[92]

Malerba, Franco. "Sectoral Systems of Innovation and Production." *Research Policy* 31, no. 2 (2002): 247–264.

Malerba, Franco, and Luigi Orsenigo. "The Evolution of the Pharmaceutical Industry." *Business History* 57, no. 5 (July 4, 2015): 664–87. doi:10.1080/00076791.2014.975119.

Markiewicz, K. "Firm Capabilities and Absorptive Capacity: Implications for Exploitation of Public Science and the Pace of Knowledge Exploitation." mimeo, UC Berkeley. 9, 2006.

McKelvey, Maureen, and Luigi Orsenigo. "Pharmaceuticals as a Sectoral Innovation System." *ESSY Project (European Sectoral Systems of Innovation), November*, 2001. Accessed April 6, 2017. http://www.druid.dk/uploads/tx_picturedb/dw2002-447.

McKelvey, Maureen, Luigi Orsenigo, and Fabio Pammolli. 2004. "Pharmaceuticals Analyzed through the Lens of a Sectoral Innovation System." In *Sectoral systems of innovation: concepts, issues and analyses of six major sectors in Europe*, edited by Malerba, Franco Cambridge University Press, 2004.

Meldrum, Cliff, Maria A. Doyle, and Richard W. Tothill. "Next-Generation Sequencing for Cancer Diagnostics: A Practical Perspective." *Clin Biochem Rev* 32, no. 4 (2011): 177–195.

Mittra, James. "Life Science Innovation and the Restructuring of the Pharmaceutical Industry: Merger, Acquisition and Strategic Alliance Behaviour of Large Firms."

[92] The reference for COMETS dataset. The COMETS requires the following acknowledgement: Certain data included herein are derived from the COMETS database release 2.0. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Ewing Marion Kauffman Foundation, the National Science Foundation, or other funders of COMETS.

Technology Analysis & Strategic Management 19, no. 3 (May 2007): 279–301. doi:10.1080/09537320701281540.

Moynihan, Ray. "Alosetron: a case study in regulatory capture, or a victory for patients' rights?." *BMJ: British Medical Journal* 325, no. 7364 (2002): 592.

Munos, Bernard. "Lessons from 60 Years of Pharmaceutical Innovation." *Nature Reviews Drug Discovery* 8, no. 12 (December 2009): 959–68. doi:10.1038/nrd2961.

Narin, Francis, Kimberly S. Hamilton, and Dominic Olivastro. "The increasing linkage between US technology and public science". *Research policy*, *26, no.* 3 (1997), pp.317-330.

Nesta, Lionel, and Pier Paolo Saviotti. "Coherence of the Knowledge Base and the Firm's Innovative Performance: Evidence from the Us Pharmaceutical Industry*." *The Journal of Industrial Economics* 53, no. 1 (2005): 123–142.

Nelson, Richard R., and Katherine Nelson. "Technology, Institutions, and Innovation Systems." *Research Policy* 31, no. 2 (2002): 265–272.

Niosi, Jorge. "Complexity and Path Dependence in Biotechnology Innovation Systems." *Industrial and Corporate Change* 20, no. 6 (December 1, 2011): 1795–1826. doi:10.1093/icc/dtr065.

Nonaka, Ikujiro, and Noboru Konno. "The Concept of "*ba*": Building a Foundation for Knowledge Creation." *California Management Review* 40, no. 3 (1998): 41.

OECD Triadic Patent Families database, September 2017

OECD Patent Quality Indicators database, September 2017

Organisation for Economic Co-operation and Development, ed. *OECD Patent Statistics Manual*. Paris: OECD, 2009.

Pammolli, Fabio, Laura Magazzini, and Massimo Riccaboni. "The Productivity Crisis in Pharmaceutical R&D." *Nature Reviews Drug Discovery* 10, no. 6 (June 2011): 428–38. doi:10.1038/nrd3405.

Peneder, Michael, and Martin Wörter. "Competition, R&D and Innovation: Testing the Inverted-U in a Simultaneous System." *Journal of Evolutionary Economics*, 2013, 1–35.

Penner-Hahn, Joan, and J. Myles Shaver. "Does International Research and Development Increase Patent Output? An Analysis of Japanese Pharmaceutical Firms." *Strategic Management Journal* 26, no. 2 (February 2005): 121–40. doi:10.1002/smj.436.

Petrova, Elina. "Innovation in the Pharmaceutical Industry: The Process of Drug Discovery and Development." In *Innovation and Marketing in the Pharmaceutical Industry,* edited by Min Ding, Jehoshua Eliashberg, and Stefan Stremersch, pp. 19-81. Springer New York, 2014.

Pew Research Center. "What Low Response Rates Mean for Telephone Surveys" May, 2017. Accessed September 5, 2017. http://assets.pewresearch.org/wp-content/uploads/sites/12/2017/05/12154630/RDD-Non-response-Full-Report.pdf

Pharmaceutical Research and Manufacturers of America (PhRMA). 2016. Biopharmaceutical Research Industry Profile. Washington, DC: PhRMA; 2016.

Phene, Anupama, Karin Fladmoe-Lindquist, and Laurence Marsh. "Breakthrough Innovations in the U.S. Biotechnology Industry: The Effects of Technological Space and Geographic Origin." *Strategic Management Journal* 27, no. 4 (April 2006): 369–88. doi:10.1002/smj.522.

Popp, David, Ted Juhl, and Daniel KN Johnson. "Time in Purgatory: Examining the Grant Lag for US Patent Applications." *Topics in Economic Analysis & Policy* 4, no. 1 (2004). https://www.degruyter.com/view/j/bejeap.2004.4.issue-1/bejeap.2004.4.1.1329/bejeap.2004.4.1.1329.xml.

Powell, Walter W. "Learning from collaboration: Knowledge and networks in the biotechnology and pharmaceutical industries." *California management review* 40, no. 3 (1998): 228-240.

Reuters, Thomson. "2015 CMR International Pharmaceutical R&D Factbook - Executive Summary." (2015). Accessed April 6, 2017. http://cmr.thomsonreuters.com/pdf/Executive_Summary_Final.pdf

Rosenkopf, Lori, and Atul Nerkar. "Beyond Local Search: Boundary-Spanning, Exploration, and Impact in the Optical Disk Industry." *Strategic Management Journal* 22, no. 4 (April 2001): 287–306. doi:10.1002/smj.160.

Rzakhanov, Zaur. "Regulatory Policy, Value of Knowledge Assets and Innovation Strategy: The Case of the Orphan Drug Act." *Research Policy* 37, no. 4 (May 2008): 673–89. doi:10.1016/j.respol.2007.12.009.

Sampat, Bhaven N., David C. Mowery, and Arvids A. Ziedonis. "Changes in University Patent Quality after the Bayh–Dole 18 Act: A Re-Examination," *International Journal of Industrial Organization*, 21(9), pp.1371-1390.

Sauermann, Henry, and Michael Roach. "Increasing Web Survey Response Rates in Innovation Research: An Experimental Study of Static and Dynamic Contact Design Features." *Research Policy* 42, no. 1 (February 2013): 273–86. doi:10.1016/j.respol.2012.05.003.

Scannell, Jack W., Alex Blanckley, Helen Boldon, and Brian Warrington. "Diagnosing the Decline in Pharmaceutical R&D Efficiency." *Nature Reviews Drug Discovery* 11, no. 3 (2012): 191–200.

Schmoch, Ulrich. "Concept of a Technology Classification for Country Comparisons." *Final Report to the World Intellectial Property Office (WIPO), Karslruhe: Fraunhofer ISI*, 2008. Accessed April 11, 2017. http://www.wipo.int/edocs/mdocs/classifications/en/ipc_ce_41/ipc_ce_41_5-annex1.pdf.

Schultz, Laura I., and Frederick L. Joutz. "Methods for Identifying Emerging General Purpose Technologies: A Case Study of Nanotechnologies." *Scientometrics* 85, no. 1 (October 2010): 155–70. doi:10.1007/s11192-010-0244-2.

Schüpbach, Jan M. "Foreign direct investment in agriculture: the impact of out grower schemes and large-scale farm employment on economic well-being in Zambia". PhD Thesis, No. 22287 Eth Zurich (2014).

Schumpeter, Joseph Alois. *Business Cycles*. A Theoretical, Historical and Statistical Analysis of the Capitalist Process. Abridged, with an introduction, by Rendigs Fels. McGraw-Hill New York, 1939. http://lewebpedagogique.com/jineu66/files/2010/10/tes_spec_1_1939_business_cycles.pdf.

Shackelford, Brandon, and Francisco Moris. "A Snapshot of Business R&D Employment in the United States." October 2016. National Science Foundation, National Center for Science and Engineering Statistics. Accessed September 6, 2017. https://pdfs.semanticscholar.org/dd0b/135a06590b9e5730337e5f4f00ac32563ed4.pdf.

Sharif, Naubahar. "Emergence and Development of the National Innovation Systems Concept." *Research Policy* 35, no. 5 (June 2006): 745–66. doi:10.1016/j.respol.2006.04.001.

Singh, Jasjit, and Lee Fleming. “Lone Inventors as Sources of Breakthroughs: Myth or Reality?” *Management Science* 56, no. 1 (January 2010): 41–56. https://doi.org/10.1287/mnsc.1090.1072.

Sisko Patana, Anne, Matti Pihlajamaa, Kirsi Polvinen, Tamara Carleton, and Laura Kanto. “Inducement and Blocking Mechanisms in the Finnish Life Sciences Innovation System.” Edited by Tugrul Daim. *Foresight* 15, no. 6 (November 11, 2013): 428–45. doi:10.1108/FS-10-2012-0081.

Stakleff, N. Nicole. “Drug Life: The Chemistry of Patent and Regulatory Exclusivity for Pharmceuticals, A.” *Fla. Coastal L. Rev.* 16 (2014): 27.

So, A.D., N. Gupta, S.K. Brahmachari, I. Chopra, B. Munos, C. Nathan, K. Outterson, et al. “Towards New Business Models for R&D for Novel Antibiotics.” *Drug Resistance Updates* 14, no. 2 (April 2011): 88–94. doi:10.1016/j.drup.2011.01.006.

Somaya, Deepak, Ian O. Williamson, and Xiaomeng Zhang. “Combining Patent Law Expertise with R&D for Patenting Performance.” *Organization Science* 18, no. 6 (December 2007): 922–37. doi:10.1287/orsc.1070.0292.

Sternitzke, C., 2007. Reducing uncertainty in the patent application procedure – insights from malicious prior art in European patent applications. The R&D Management Conference, Risk and Uncertainty in R&D Management, Bremen, Germany, 4-6 July 2007, ISBN 0-9549916-9-9. Accessed April 10, 2017. http://www.fofi.uni-bremen.de/files/moehrle/alumini/6)Sternitzke_et_al_2007_Reducing_uncertianty_in_the_patent_application_procedure.pdf

Squicciarini, Mariagrazia, Chiara Criscuolo, and Hélène Dernis. “Measuring Patent Quality: Indicators of Technological and Economic Value” OECD Science, Technology and Industry Working Papers, No. 2013/03. Accessed April 6, 2017. http://www.oecd-ilibrary.org/science-and-technology/measuring-patent-quality_5k4522wkw1r8-en.

Surden, Harry. “Efficient Uncertainty in Patent Interpretation.” *Wash. & Lee L. Rev.* 68 (2011): 1737.

Suurs, Roald A.A., and Marko P. Hekkert. “Cumulative Causation in the Formation of a Technological Innovation System: The Case of Biofuels in the Netherlands.” *Technological Forecasting and Social Change* 76, no. 8 (October 2009): 1003–20. doi:10.1016/j.techfore.2009.03.002.

Suurs, Roald A.A., Marko P. Hekkert, Sander Kieboom, and Ruud E.H.M. Smits. “Understanding the Formative Stage of Technological Innovation System

Development: The Case of Natural Gas as an Automotive Fuel." *Energy Policy* 38, no. 1 (January 2010): 419–31. doi:10.1016/j.enpol.2009.09.032.

Tait, Joyce. "Systemic Interactions in Life Science Innovation1." *Technology Analysis & Strategic Management* 19, no. 3 (May 2007): 257–77. doi:10.1080/09537320701281524.

Toole, Andrew A. "The Impact of Public Basic Research on Industrial Innovation: Evidence from the Pharmaceutical Industry." *Research Policy* 41, no. 1 (February 2012): 1–12. doi:10.1016/j.respol.2011.06.004.

Trajtenberg, Manuel. "A Penny for Your Quotes: Patent Citations and the Value of Innovations." *The RAND Journal of Economics* 21, no. 1 (1990): 172. doi:10.2307/2555502.

Tsai, Fu-Sheng. "Knowing What We Know Differently: Knowledge Heterogeneity and Dynamically Ambidextrous Innovation." *Journal of Organizational Change Management* 29, no. 7 (November 14, 2016): 1162–88. doi:10.1108/JOCM-01-2016-0021.

Van Looy, Bart, Tom Magerman, and Koenraad Debackere. "Developing Technology in the Vicinity of Science: An Examination of the Relationship between Science Intensity (of Patents) and Technological Productivity within the Field of Biotechnology." *Scientometrics* 70, no. 2 (February 2007): 441–58. doi:10.1007/s11192-007-0211-8.

Van Looy, Bart, Edwin Zimmermann, Reinhilde Veugelers, Arnold Verbeek, Johanna Mello, and Koenraad Debackere. "Do Science-Technology Interactions Pay off When Developing Technology?" *Scientometrics* 57, no. 3 (2003): 355–367.

Verbeek, Arnold, Koenraad Debackere, Marc Luwel, Petra Andries, Edwin Zimmermann, and Filip Deleus. "Linking Science to Technology: Using Bibliographic References in Patents to Build Linkage Schemes." *Scientometrics* 54, no. 3 (2002): 399–420.

Verhoeven, Dennis, Jurriën Bakker, and Reinhilde Veugelers. "Measuring Technological Novelty with Patent-Based Indicators." *Research Policy* 45, no. 3 (April 2016): 707–23. doi:10.1016/j.respol.2015.11.010.

Ward, Michael R., and David Dranove. "The vertical chain of research and development in the pharmaceutical industry." *Economic Inquiry* 33, no. 1 (1995): 70-87.

Weisberg, Robert W. "Creativity and Knowledge: A Challenge to Theories." *Handbook of creativity* 226-250 (1999).

Wellman-Labadie, Olivier, and Youwen Zhou. "The US Orphan Drug Act: Rare Disease Research Stimulator or Commercial Opportunity?" *Health Policy* 95, no. 2–3 (May 2010): 216–28. doi:10.1016/j.healthpol.2009.12.001.

Williams, David R. *The funding of biopharmaceutical research and development*. Elsevier, 2013

Wirth, Steffen, and Jochen Markard. "Context Matters: How Existing Sectors and Competing Technologies Affect the Prospects of the Swiss Bio-SNG Innovation System." *Technological Forecasting and Social Change* 78, no. 4 (May 2011): 635–49. doi:10.1016/j.techfore.2011.01.001.

Wuyts, Stefan. "New Challenges in Alliance Portfolio Management" In *Innovation and Marketing in the Pharmaceutical Industry,* edited by Min Ding, Jehoshua Eliashberg, and Stefan Stremersch, pp. 83-118. Springer New York, 2014.

Yao, Lixia, James A. Evans, and Andrey Rzhetsky. "Novel Opportunities for Computational Biology and Sociology in Drug Discovery." *Trends in Biotechnology* 27, no. 9 (September 2009): 531–40. doi:10.1016/j.tibtech.2009.06.003.

Youtie, Jan, Maurizio Iacopetta, and Stuart Graham. "Assessing the Nature of Nanotechnology: Can We Uncover an Emerging General Purpose Technology?" *The Journal of Technology Transfer* 33, no. 3 (June 2008): 315–29. doi:10.1007/s10961-007-9030-6.

Zelenay Jr, James L. "The Prescription Drug User Fee Act: Is a Faster Food and Drug Administration Always a Better Food and Drug Administration." *Food & Drug LJ* 60 (2005): 261.

BIOGRAPHY

Alfred Sarkissian holds a bachelor's and a master's degree in industrial management. His research interests revolve around the determinants, processes, and outcomes of innovation.