Essays on Technological Innovation in the Health Care Industry

Tao Wang

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

School of Management

University College London

November 29, 2018

I, Tao Wang, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

This thesis investigates the relationship between knowledge management and innovation performance in the context of new product development in the biotechnology industry. The dissertation contains three empirical essays in three different research settings. Chapter two focuses on how to improve the outcome-dependence of experiential learning. The findings demonstrate that two innovation strategy choices - the novelty of innovation and the primary ownership - enable prior failure experience to reduce the incidence of future failure more than prior success experience does. Chapter three assesses stage specific learning and knowledge spillover. The findings demonstrate that the productivity of drug research increases only with previous upstream research experience, whereas the productivity of drug development increases only with previous downstream development experience. Chapter four investigates the interaction between network structure and interfirm governance and its effect on knowledge appropation and innovation behaviour. The findings show that interfirm governance contingently determines the outcome of patent applications and patent citations within a dense network. Together these three essays provide three primary contributions. First, the dissertation shows that failure experience has knowledge benefits and investing in failures allows an organisation to build capabilities that improve future performance. In addition, two innovation strategies - innovation novelty and primary ownership - moderate the search behaviour and help to translate the inferences from an organisational experience into knowledge and routines. Second, the dissertation illustrates that both learning by doing and product innovation explain economic growth externality and improve productivity, and knowledge spillover is bounded in the same stage for long-run growth. Third,

when considering the formation of alliances to advance innovation, neither alliance network structure nor interfirm governance guarantee superior performance. Therefore, the dynamic process of a strategic alliance by focusing on the interaction between network structure and interfirm governance need to be considered.

Impact Statement

The total expenditures of the US health care is $3.2 trillion or 17.8% of GDP in 2015, and the expenses for drugs is $1.5 trillion, which is nearly half of the total expenditures of health care. Therefore, the biopharmaceutical industry is a key player in the health care system to develop medications to improve global health, prosperity and economic productivity by reducing the incidence of diseases, treating illnesses and enhancing the quality of life of people. The industry's main contribution is engaging drug research and development (R&D) through technological innovation to meet the complex health care demands of population.

Despite the critical role of this industry and the tremendous improvement in the science and technology underpinning drug R&D, there has been little improvement in the output of this industry. In other words, the number of new drugs approved by the US Food and Drug Administration each year has not changed for the past twenty years. Even worse, after taking the R&D expenses into consideration, the number of new FDA-approved drugs per billion US dollars of R&D spending has halved every 9 years since 1950. In this dissertation, I look at three specific knowledge management processes - outcome-dependent learning, knowledge spillover across stages and knowledge appropriation in alliance networks - that help explain the innovation heterogeneity that can be seen among organisations and act as the potential factors to improve technological innovation.

The findings of this dissertation show that failure experience is more than a disaster for organizations but has knowledge benefits, and investing in failures allows organisations to build capabilities that improve future performance. In addition, two innovation strategies - innovation novelty and primary ownership - improve the build-up capability from failure experience. These two strategic choices help practitioners to develop superior strategies for coping with failures and enable organisations to reap substantial gains. Policy makers need to understand that balancing incremental innovation and radical innovation or allocating more resource to inventors than to licensees could promote learning from previous failures and enhance

innovation efficiency.

Second, the dissertation illustrates that both learning by doing and product innovation explain economic growth externality and improve productivity, and managers should make use of this knowledge management tool to achieve high productivity during innovation. By setting the knowledge boundary and understanding the sources of learning by innovating also enable decision makers to resolve low efficiency of learning by matching knowledge and experience. Furthermore, due to the knowledge boundary condition, the government should use tax laws to encourage investment in intra-stage technological progress instead of promoting cross-stage technological innovation.

Third, this dissertation bolsters the case for the dynamic process of strategic alliance by focusing on the interaction between network structure and alliance governance since these two factors increase innovation independently and in combination. In other words, decision-makers have to look beyond the dyadic level of relationship and consider the network level characteristics in conjunction in order to optimize their alliance strategy and to achieve better innovation performance.

Acknowledgements

I am especially grateful to Simcha Jong, my supervisor, for his ceaseless devotion to his students and for his encouragement, enthusiasm, and support over the years and his faith in my potential. I particularly appreciate the continued support and guidance of Simcha through the extended process of this dissertation. I also thank Bart Vanneste and Martin Kilduff for offering their guidance throughout this process. I am also grateful for the support from other faculty members and fellow PhD students at the UCL School of Management.

My deepest gratitude also goes to my mentors and collaborators when I was visiting Yale University and Harvard University. Professor Olav Sorenson, my thanks to you for sharing data and providing feedback on my projects. Professor Josh Lerner, thank you for exemplifying quality research and teaching excellence, and thank you for the support when I was teaching at Harvard. Professor Ariel Pakes, thank you for broadening my knowledge and perspective in industrial organization and strategy. Professor Richard Freeman, thank you for your expert guidance and mentoring. In addition, I would like to thank Professor Jim Baron, Marissa King, Ian Cockburn, Jeff Furman, Scott Stern and Bill Kerr for their comments and suggestions on my work. I am also grateful to the seminar participants of Harvard IO Lunch and NBER Productivity Lunch for their valuable suggestions.

I am greatly indebted to my parents who gave me the freedom and support to fulfil my professional ambitions. Without their encouragement, I would not have had the strength to complete this programme. Finally, I would like to dedicate this dissertation to my grandparents, who were looking forward to my graduation, but passed away before I had fully completed my dissertation.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In many industries, technological innovation is the most important driver of competitive success. Introducing new products helps firms protect their margins, while investing in process innovation helps firms lower their costs. Besides raising the competitive bar for firms, technological innovation also positively affects society. Innovation enables a wider range of goods and services to be delivered to people worldwide, raising living standards and economic performance.

Gross domestic product (GDP), the total annual output of an economy, reflects the aggregate impact of technological innovation. Figure 1.1 shows the GDP adjusted for inflation for the US from 1970 to 2011. As shown in the figure, the US GDP has risen steadily since 1970. A series of studies conducted at the National Bureau of Economic Research (NBER) shows that the historic rate of economics growth in GDP cannot be explained entirely by growth in labour and capital inputs (e.g. [Romer, 1990]; [Islam, 1995]). Nobel Prize winner Robert Merton Solow argues that this unaccounted residual growth could be explained by technological innovation [Solow, 1994]. In other words, innovation increases the amount of output achievable from a given quantity of labour and capital and improves social welfare.

Knowledge is a resource requiring explicit and specific management policies and practices to be acquired, processed and exploited efficiently. According to the endogenous growth theory in macroeconomics, knowledge production and diffusion are central elements determining economic growth [Aghion and Howitt, 1998]; [Romer, 1990] and the sources of economic expansion [Warsh, 2007]. Strategy

Figure 1.1: Product Innovation Productivity

US Gross Domestic Product (in Real 2017 $US Billion)
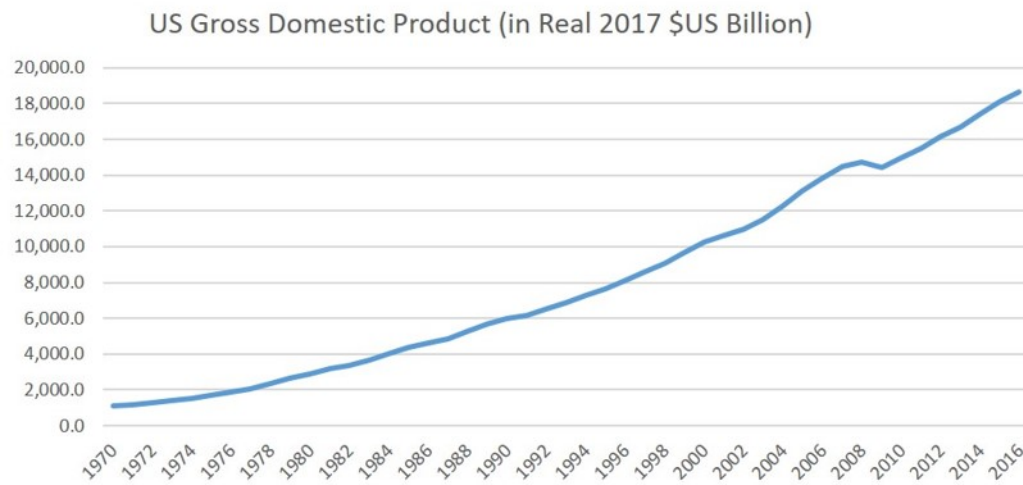
Notes: US gross domestic product from 1970 to 2016. All numbers have been converted into 2017 dollars and adjusted for inflation.

scholars are developing a knowledge-based view of the firm: theorising that firms exist because they provide efficiency advantages in the use, creation, and commercialisation of knowledge relative to markets [Kogut and Zander, 1996] and that aspects of the knowledge creation process influence a firms' scale and scope [Nickson and Zenger, 2004]. Among other factors, knowledge determines firm productivity and its medium and long term competitive advantage [Grant, 1996]; [Roberts, 1999]; [Eisenhardt and Martin, 2004].

Knowledge management is the process of creating, sharing and managing knowledge, and fostering innovation [Alavi and Leidner, 2001]; [Barley et al., 2018]. It is the main driver and determinant of technological innovation. Schumpeterian innovation competences are based on the growth of knowledge stock and generative learning [Schumpeter, 2010]. These competences lead to changes with regard to the practices developed in the organisation, and strengthen its competences and capabilities. Therefore, knowledge management can contribute to the sustainability of competitive advantage, enabling the development of distinctive competences.

Despite the importance of technological innovation and knowledge management, however, very little research has systematically examined the advanced strate-

gies and well-developed processes to manage knowledge to improve technological innovation. Therefore, many firms choose knowledge management practices that are a poor fit with their resources and objectives; manage the new product development process wrongly, causing high failure rates and long development cycles; and organize structures that cannot effectively support the innovation.

This dissertation and my broader research agenda look at the knowledge management factors that play a crucial role in determining the heterogeneity in firms' innovation behaviour and competitive performance. It looks at the role that experience and structures play both in developing capabilities, and in improving innovation productivity. The research context of this dissertation is the US biotechnology industry and the new products focused on here are new drugs.

One of the most obvious sources of knowledge production and innovation is the firms' own research and development (R&D) efforts. The term research and development refers to a range of activities that extend from early exploration of a scientific domain to specific commercial implementations. Studies show that R&D innovation is the lifeblood of the biotechnology industry [Henderson and Cockburn, 1996]; [Azoulay, 2004]. This 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. Biotechnology firms become industry leaders by spending large sums on R&D in order to produce a steady stream of patents and successful products. In other words, biotechnology firms' R&D novelty has a strong positive correlation with their sales growth rate, sales from new products, and profitability. Because patents only protect firms from generic competition for a limited period of time (usually 15 years) [Griliches, 1990], these firms need to continually innovate to ensure survival and growth.

Searching for innovative products is extremely difficult in almost any industry. It is especially challenging in the biotechnology industry, because products come from highly complex fields, such as molecular biology and biochemistry, and involve the most delicate and complicated system - human body. The quest for new

drug development must combine an understanding of the complexity of the human body with a knowledge of life science.

The drug R&D process starts from synthesising and screening new compounds which might interact with specific genes or proteins to combat particular diseases. Once candidate molecules are identified, they are studied in test tubes and in animals to determine their side effects, efficacy, and pharmacological dynamics. Only the promising candidates can proceed to human clinical trial by submitting an investigational new drug (IND) application to the Federal Drug Administration (FDA) first, informing the agency that human studies will start in 30 days unless it objects.

The FDA requires drugs to go through three phases of clinical trials on humans. During phase I, a small number of healthy people receive the drug to test safety and dose range. During the second phase of clinical trials, a large number of subjects who have the disease or conditions that the drug is intended to treat, are tested in placebo-controlled trials. Besides safety and optimal doses, Phase II also investigates the efficacy of the drug. Phase III is tested on an even larger patient population to assess safety, effectiveness, and optimal doses. Normally Phase III is double-blind, which means that neither doctors nor patients are aware of whether the patients are taking the drug or a placebo. When Phase III is completed, the company needs to submit a New Drug Application (NDA) to the FDA before it can launch the new product.

Generally speaking, hundreds of compounds need to be tested to find one promising candidate. During clinical trials, if 20 drugs enter Phase I, only around 13 will successfully complete it. Of those, about nine will finish Phase II, but only 1.5 will survive Phase III. Even after a drug successfully completes Phase III, the FDA might not approve it because they think the data is insufficient for approval. Ultimately, only one of the original 20 may be approved for marketing [Steven et al., 2010].

Despite tremendous improvement in the science and technology underpinning drug discovery, improvement 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

Figure 1.2: Eroom's Law in Drug R&D

a Overall trend in R&D efficiency (inflation-adjusted)

FDA tightens
regulation
post-thalidomide

FDA clears backlog
following PDUFA
regulations plus small
bolus of HIV drugs

First wave of
biotechnology-
derived therapies

b Rate of decline over 10-year periods

c Adjusting for 5-year delay in spending impact

Drugs per billion US$ R&D
spending 5 years previously

Drugs per billion US$
R&D spending

Source: Scannell et al. (2012)

of new drugs approved by the FDA [Kola and Landis, 2004].

The paradoxical phenomenon of declining R&D efficiency has been docu-
mented by Scannell and his collogues (2012). They use the term "Eroom's Law",
which is "Moore's Law" backwards, to refer to the number of new FDA-approved
drugs per billion US dollars of R&D spending has halved every 9 years since 1950
(See Figure 1.2). As Figure 1.2 depicts, the decline rate is almost the same across
different 10-year segments (Panel b). Moreover, the trend seems to be robust in
relation to various assumptions on average delay between R&D spending and drug
approval (Panel c).

In my dissertation, using the biotechnology industry as my research context, I
look at three specific knowledge management processes - outcome-dependent learn-
ing, knowledge spillover across stages and knowledge appropriation in alliance net-

works - that help explain the innovation heterogeneity that can be seen among organisations (See Table 1.1 for a brief overview of each chapter). In Chapter 2, I investigate how to improve the outcome-dependence of experiential learning. While we know organizations need to learn how to innovate and revitalize [Grant, 1996]; [Argote et al., 2003], we know less about how to improve the outcome-dependence of learning [Madsen and Desai, 2010]. This study acknowledges the gap and disentangles the theories of learning from failures and successes. To address project heterogeneity and to alleviate the non-randomness of innovation strategy choices for causal inference, I deploy Competing Risk Analysis [Jenkins, 2005] and Coarsened Exact Matching [Blackwell et al., 2010] methods to analyse more than two thousand drug development projects from 37 US biotechnology companies. I find that two innovation strategy choices - the novelty of innovation and the primary ownership - enhance learning from failure experience. These two strategy choices also enable prior failure experience to reduce the incidence of future failure more than prior success experience does. In addition, by exploring interview data, I seek to provide additional insight into the mechanisms of learning from failures and learning from successes. Taken together, these findings offer both theoretical and practical insights into how innovative firms could improve their learning from successes and failures differently.

In Chapter 3, I examine learning by innovating and the boundary of learning by innovating, by using detailed project-level data in the biotechnology industry. Both product innovation and learning by doing reduce uncertainty and improve productivity [Arrow, 1972]; [Syverson, 2011]; [Aghion and Jaravel, 2015]. Previous research shows that learning by innovation, the combination of these two factors, is a crucial antecedent for innovation success [de Ven and Polley, 1992]; [Aghion and Jaravel, 2015]. However, whether and under what conditions learning by innovating to enhance productivity is not well understood [Nooteboom, 2004]. This study illustrates learning by innovating drives productivity in the biotechnology industry. In addition, this study documents that learning by innovating is stage-specific, which means that one stage of drug innovation cannot fully appropriate

the productivity gains acquired through experience with another stage of drug innovation. In other words, the productivity of drug research increases only with previous upstream experience, whereas the productivity of drug development increases with only previous downstream development experience. This paper proposes that the lack of knowledge spillover between stages is due to limited attention, and proves that decision makers' attention reduce this learning constraint. These results shed light on how learning by innovating shapes the productivity trajectory and the boundary of learning by innovating.

Finally, in Chapter 4, I consider the structuring condition of alliance networks and their effect on knowledge appropriation and innovation behaviour. Previous research indicates that there is a growing recognition of the importance of collaborative R&D networks for successful innovation [Freeman, 1991]; [Hargadon and Sutton, 1997]; [Ahuja and Lampert, 2001]. Collaboration is especially crucial in high-technology sectors, such as the biotechnology industry, where it is unlikely that a single individual or organisation will possess all of the resources and capabilities necessary to develop and implement technological innovation [Powell et al., 1996]. Alliance networks blend the features of market-based exchanges as well as of hierarchies, so both network structure and interfirm governance present interesting challenges and opportunities for firms' innovation behaviour [Powell et al., 1996]; [Oxley, 1997]; [Ahuja, 2000]; [Polidoro et al., 2011]. Instead of treating them separately, this study examines the interaction of network structure and interfirm governance on the " twin tasks" in innovation search, namely access to novel information on the one hand and the build-up of absorptive capacity on the other. Using a combination of propensity-score weighting and difference-in-difference estimation strategies to address endogeneity, I find that interfirm governance, the transaction feature of alliance networks, contingently determines the outcome of patent applications and patent citations within dense networks. More specifically, new partners and non-equity relationships improve the positive influence of dense network when the density is low, but mitigate the cost when the density is high, and enhance the amplitude of the effect of dense networks. These

results indicate the benefits of network closure contingent on interfirm governance, namely partner selection and ownership structure. Therefore firms embedded in a certain type of network structure should clearly recognise such contingency and its implications, and devise optimal alliance strategies to manage it.

	Phenomenon	Theory	Key Construct	Identification
Chapter 2	Improvement in the outcome-dependent experiential learning	Organisational search	Innovation experience	Competing risk analysis
		Innovation Strategy	Innovation strategy: novelty, co-development and primary ownership	Coarsen exact matching
		Experiential learning	Firm performance: future failure/success hazard ratio	
Chapter 3	Innovation productivity based on the configuration of value chain activities	Cobb-Douglas production function	Stage-specific experience	Competing risk analysis
	Organisational boundary	Knowledge spillover construct	Productivity: future failure/success hazard ratio	
		Solow growth model		
Chapter 4	Alliance network features of network structure and inter-firm governance determine firm's innovation behaviour	Innovation search	Governance design: ownership structure and partner selection	Propensity score weighting
		Transaction cost economics theory	Firm innovation performance: patent applications and citations	Semiparametric difference in differences

Table 1.1: Overview of the Following Three Chapters

Below I will first summarise the relevant research questions and key findings for each of the three core chapters. Subsequently, I will summarise the key contributions of these three chapters. Finally, I will introduce the future research direction and possibilities.

Chapter 2

The Wisdom of Learning: Knowledge Strategies in Technological Innovation

2.1 Introduction

Technological innovation is essential to competitive advantage and to economic development more generally; however, many innovating firms fail to obtain sustainable innovation and significant economic returns [Teece, 1986]. The ability of firms to learn and acquire knowledge has emerged as a key factor influencing performance and survival [Grant, 1996]; [Argote et al., 2003]. Organisational learning, the key building block of innovation, is a process of searching through prior experience [Levinthal, 1997] and translating the inferences from such experience into knowledge and routines that systematically alter subsequent behaviour [Cyert and March, 1963]; [Levitt and March, 1988]. This process is not the retrieval of some entity, but a process of sensemaking to recombine existing materials [Nelson and Winter, 1982]; [Penrose, 2009] or to reconfigure the ways in which knowledge elements are linked [Henderson and Clark, 1990]. Different interpretations lead to different organisational responses, which ultimately shape strategy, norms, forms and protocols for learning [Daft and Weick, 1984]. This implies, in particular, that learning is dependent upon current strategy choices, which gain at-

tention and engender actions for sensemaking to interpret collective experience. In other words, current innovation strategies affect learning outcomes by assisting organisations in making sense of the experience through emphasising discovery through enactment and interpretation [McGrath, 2001].

Indeed, the interplay between past experience and current strategy are central to the creation of meaning and development of knowledge [Colville et al., 2014]. Previous research shows that experience is appreciated through current strategies. For example, the strategy choices help to link between events and models, plot narratives and selectively retain the interpretations considered valuable or preferable [Henderson and Cockburn, 1994]; [Maitlis and Sonenshein, 2010]. Despite these significant advances, however, important questions, such as how the current innovation strategy improves the process of experience interpretation and enhances learning from failures, and whether it also affects learning from successes in the same way, remain unanswered. This is a crucial gap, because learning during innovation is not straightforward due to the use of heuristics and insights in relation to ambiguous contexts, and some evidence suggests that learning is highly specific to the innovation context (e.g. [du Plessis, 2007]; [Madsen and Desai, 2010]). The purpose of the present study is to address this gap by identifying innovation strategies which enhance the sensemaking process in order to improve learning. More specifically, I consider distinct innovation strategies to be delineated by innovation novelty (incremental vs. radical), co-development behaviour (collaboration vs. solo) and primary ownership (inventor vs. co-developer). Within each of these six innovation strategies, I further disaggregate failure and success experience and address their relative learning outcome.

Even though many organisational learning theorists have long held that organisations learn primarily through processes of 'problemistic search', wherein they engage only after experiencing failure [Cyert and March, 1963]; [March and Shapira, 1992]; [Sitkin, 1992]; [Dahlin et al., 2018]; [Posen et al., 2018], any existing evidence that failure is more important than success for organisational learning is entirely anecdotal. Indeed, the way in which learning is contingent on prior outcomes has received

little empirical attention compared to the literature on learning from aggregated experience. Disaggregating success and failure experience helps us understand which experience drives the improvement of organisational performance. Only three studies have directly compared learning from failure with learning from success but they report disparate results: Madsen and Desai (2010) found that orbit launch vehicle companies learn more effectively from failure experience than from success experience, and the magnitude of failures influences the learning process. Muehlfeld and his colleagues (2012) uncovered that newspaper-producing firms could learn from both previous failure and success experiences in relation to acquisition but that this learning process depends on two context factors: the degree of structural variance and the level of stimulation of deliberate learning. Gong and his colleagues (2017) discovered that the experience of successful acquisitions drives up the acquisition premium whereas the experience of failed acquisitions reduces the premium, and the magnitude of these experiences affects the learning outcomes. In addition, none of these studies has investigated learning from innovation in which increasing effectiveness is the result of searching for and exploring alternative routines to practise and the refinement of innovation-related skills. Therefore, whether firms learn from success and failure innovation experiences and whether innovation strategies could change learning behaviour remain unclear.

To explore these issues, I examine the drug innovation process in the US biotechnology industry during the period 1987-2012. This process provides an interesting context in which to study exploratory learning during technological innovation, considering that drug development is the core competence of biotechnology firms and that these firms are usually proactive and willing to take risks during this process. The empirical setting is the US, where biotechnology is well developed and drug development is extensive. However, there is an empirical challenge to answering my research question: information on how and why decision makers choose certain innovation strategies is generally unavailable. Therefore, innovation strategy choices could be the outcome of a process which matches purposely with previous experience, making it difficult to uncover causal effects. To overcome the

endogeneity of the strategic choices, I make use of a matching method in order to make projects comparable.

My results provide support for the hypotheses that two dimensions of innovation strategies - innovation novelty (radical or incremental innovation) and primary ownership (inventor or co-developer) - moderate learning from failure and success. In particular, I find that working on incremental innovation projects improves sensemaking and knowledge recombination and enhances learning from failure more than working on radical innovation projects does. An additional Wald test also shows that working on incremental innovation advances learning from failure experience more than from success experience. Primary ownership has a similar effect on making sense of previous experience. Compared to a licensees, working as an inventor reduces future failure by giving meaning to previous failure experience more than working as a licensee does. It also empowers prior failure experience to reduce the likelihood of future failure more than prior success experience does.

This study contributes to existing learning theories in several ways: firstly, this study illustrates that innovation strategies serve as a moderator of sensemaking and organisational learning [du Plessis, 2007]. More specifically, this paper proposes that two innovation strategies - innovation novelty and primary ownership - moderate sensemaking behaviour and knowledge production. Secondly, this study demonstrates the relationship between failure experience and innovation performance. Previous research shows that investing in a failing technology helps an organisation to build absorptive capacity, which will improve future performance [Eggers, 2014]. This study confirms this conclusion and notes that the two above-mentioned innovation strategies could improve the build-up capability from failure experience. Thirdly, this study addresses the importance of learning from failure [Madsen and Desai, 2010]. This highlights that improved innovation performance of increasing experience is mainly driven by learning from failure instead of learning from success. In addition to the theoretical contribution, this study also advocates project-level data so as to address the microfoundation of learning variation.

2.2 Theories and Hypotheses

2.2.1 Innovation Strategy and Learning

Organisational learning is the result of a problem-solving process in which solutions to problems are discovered via search [Cyert and March, 1963]; [Dosi, 1988]. The search framework was first conceptualized by Kauffman in evolutionary biology [Kauffman, 1993]. In innovation and organisation literature, learning is a process of search over previous experience landscapes and the typography of the experience landscape determines the likelihood of fruitful search and expected success of search [Levinthal and March, 1981]; [Fleming and Sorenson, 2001]. In other words, search is often planned and guided by routines and heuristics of past experience [Nelson and Winter, 1982]; [Dosi, 1988].

Such search reflects experiential wisdom in that they are the outcome of the selection and interpretation of prior behaviour and experience. Organisational interpretation is a process of translating events and developing shared understanding and conceptual schemes among members [Daft and Weick, 1984], and gives meaning to experience in the early stage of the organisational learning process. To understand the complexity associated with organisations, innovation strategies of current projects assist organisations in making sense of the experience by emphasizing discovery through enactment and interpretation [McGrath, 2001]. Specifically, previous literature identifies three basic trade-offs: innovation novelty, co-development and primary ownership, that require strategic decisions during technological innovation (e.g. [Bierly and Chakrabarti, 1996]). These innovation strategies create responses and actions based on the interpretation of task interactions (e.g. [Haunschild and Rhee, 2004]; [Madsen and Desai, 2010]).

First, innovation strategy can affect resource allocation for experiential search. Innovation learning - a process of acquiring, distributing, interpreting and synthesizing existing components - needs resources. Organisations use these resources to perform trials to learn what an error is and to test presumed constraints to discover what is feasible. Certain innovation strategies intensify competition and uncertainty and create new problems, demanding more resources be put into search-related be-

haviours [McGrath, 2001].

Second, innovation strategy also affects equivocality of previous experience. Daft and Weick (1984) define equivocality as 'the extent to which previous experience is unclear and can be interpreted in several ways'. Organisational decision makers will always experience some equivocality in their experience and the uncertainty makes it difficult for decision makers to accurately interpret their past performance outcomes [Levitt and March, 1988]. The current innovation strategy provides a blueprint of novel combination [Schumpeter, 2010] and improves search competence [Levinthal and March, 1981] by combining the results of practice and the refinement of innovation-related skills and providing knowledge.

Third, innovation strategy presents a pattern recognition system. Searching the experience landscape is also a process of knowledge and pattern comparison between new data and pre-existing patterns of activity. Tushman and Anderson (1986) dichotomise innovation as either continuous or discontinuous. Continuous innovation is consistent with current belief and allows decision makers to apply existing mental models to identify the critical information. Discontinuous innovation, on the other hand, is unique and may alter the trajectory of the interpretation process.

2.2.2 Learning from Experiences of Success and Failure

The key insight of learning models is that change is triggered by adaptive search: satisfactory or superior outcomes from success tend to result in local search and organisation slacks, while unsatisfactory outcomes from failure call for non-local search and result in resource shortage [Cyert and March, 1963]; [Levinthal and March, 1981]; [Dahlin et al., 2018]. In other words, failed innovation leads to organisational change and brings in alternative knowledge, while satisfactory and superior performance only refine existing knowledge [March and Shapira, 1992]; [Chuang and Baum, 2003].

According to a behaviour theory of the firm, organisational decision makers respond quite differently to failure from the way they do to success

[Cyert and March, 1963]. Decision makers interpret the experience of success as evidence that the existing knowledge is up to date and the development of new knowledge is unnecessary [Lant, 1992]; [March and Shapira, 1992]. As a result, prior successful innovation induces decision makers to ignore information about the outside world and narrows down search to the neighbourhood of the status quo [Cyert and March, 1963]; [March, 1982]. Neighbourhood search represents a balance between the need to exploit the current wisdom associated with existing action and, at the same time, to engage in some degree of search for alternatives in the immediate landscape of the current behaviours [Gavetti and Levinthal, 2000].

In contrast, failed innovation upsets the status quo, calls existing routines and practices into question, draws attention to potential problems, and stimulates a search for alternatives which are likely to vary greatly from current behaviours. As a result, organisations are more likely to undertake major changes and to initiate the exploration of new practices, strategies, and courses of action to raise performance above the aspiration level. Put differently, failed innovation leads to exploratory search, while successful innovation continues to exploit search [March and Shapira, 1992]; [Chuang and Baum, 2003]; [Dahlin et al., 2018].

There are several innovation strategic choices that decision makers formulate which shape and direct an organisation's search and learning process and, subsequently, determine its performance. Specifically, three basic trade-offs that require strategic decision arise from innovation and knowledge management literature: organisations need to decide to focus on either radical or incremental innovation to make either a long jump search or just a short climb to search for the local optimum [March, 1991]); they need to find the proper balance between internal knowledge and external knowledge [Lant and Mezias, 1992]; they also need to resolve the tension between the depth and breadth of their knowledge by inventing everything in house or licensing key technology from other companies [Holt and Cornelissen, 2014].

2.2.3 Innovation Novelty and Learning

One crucial decision that a company needs to make during new product development is the innovation novelty of its product. The type of innovation novelty strongly influences the factors that shape the innovative performance, notably learning (e.g.[Tushman and Anderson, 1986]; [Garcia and Calantone, 2002]. Incremental innovation uses the same technology, introduces relatively minor changes to the existing product and reinforces the current competence [Tushman and Anderson, 1986]; [Henderson and Clark, 1990]; [Schumpeter, 2010]). Firms examine the immediate neighbourhood to identify alternative forms whose fitness value is superior to their current level of fitness. During this process, firms focus on fine-tuning products by means of incremental improvements which are inspired by other sources of innovation that use existing technology continuity and previous knowledge [Utterback and Abernathy, 1975], and conduct short climbing and neighbourhood search to recombine knowledge. Incremental innovation often needs less investment and effort, and the performance implications appear to be more modest [Marsili and Salter, 2005]. Accordingly, incremental innovation is a routinised behaviour which requires less attention and mindfulness.

Organisational search and learning require both changes and stability, and the level of uncertainty of a current project determines the effectiveness of the conditions of search and learning [Starbuck et al., 1978]. If both current projects and previous knowledge are too complex and dynamic, information overload may occur and learning will not take place since too much change and turbulence make it difficult to map the knowledge path and interpret information [March and Olsen, 1975]. Therefore, firms engaging in exploratory search from failure experience and long jump search to recombine components so as to create radical products typically suffer the costs of experimentation without harvesting many of its benefits [Nooteboom, 2004]. On the other hand, too much stability can also be dysfunctional because there is little inducement to learn or to change. The experience of success proves that the existing knowledge is not obsolete and change is not necessary [Lant and Mezias, 1992]. In this case, both the current innovation strategy

- notably the incremental innovation - and the internal processing of success experience stifle the variation of search. Therefore, decision makers are less likely to choose a risky, less cost-efficient information-generating process. In addition, they have little motivation to pursue a new information-generating process by searching beyond current knowledge bases due to satisfaction with current knowledge. Such responses are likely to lead unwaveringly to a trap, in which organisations only follow routinised behaviour and become stymied in a competence trap.

In contrast, the experience of failure provides variation. Pursuing an incremental innovation strategy only gradually searching through neighbourhood landscapes requires less attention and mindfulness, and might be positively challenged by the previous experience of failure and result in search for increased risk. Feeling challenged, in turn, might stimulate researchers to experiment and explore further and come up with new proposals [Fiol and Lyles, 1985]. In addition, it is the nature of scientists to face challenges and investigate failure, especially when they are not constrained. Failures also trigger their curiosity to find novel ways to interpret data with dialogic practices and analogical reasoning [Adams, 1990].

Furthermore, conducting incremental innovation also alleviates the resource allocation problem. The experience of success promotes the refinement of existing routines without too much variation [Sitkin, 1992], economising on resources [Cyert and March, 1963]. However, decision makers still tend to allocate more resources to successes to avoid risk. On the other hand, failure experience is less explored and search behaviours are insufficient due to lack of resource [Levinthal and March, 1981]. Therefore, imbalanced resource allocation exacerbates the problem: "learning from organisational failures is anything but straight-forward" [Edmondson, 2011].

Unlike undertaking a radical innovation strategy, which is highly uncertain and drains cognitive and other resources, conducting incremental innovation relies heavily on past routine and being able to attend to other stimuli [Shiffrin and Schneider, 1977]. Although learning from innovation failure is not straight-forward, it is necessary, since only by doing this can an organisation un-

cover the causal relationship between experience and outcomes [Sitkin, 1992]. Increasing resources could help to detect and interpret failures, which is the first step to learning from failure. In addition, the spare attention saved from incremental innovation also enables organisations to conduct a non-local problemistic search to discover the underlying mechanism of failures and generate new knowledge in order to find better solutions.

To sum up, working on incremental innovation enables organisations to learn more effectively from the experience of failure, and also to improve learning outcomes by searching through failure experience landscapes more than by searching through success experience ones.

Hypothesis 1a *Working on incremental innovation improves learning outcome from failure more than working on radical innovation does.*

Hypothesis 1b *Working on incremental innovation helps a company to reduce the incidence of future failure by learning from failure more than learning from success.*

2.2.4 Co-development and Learning

The organisation's strategy partially determines its learning capability by shaping the goals and objectives and the breadth of actions available for carrying out learning activities. A co-development strategy expands the boundaries of decision making and duplicates the experience landscapes for searching and learning [Cyert and March, 1963]. The enlarged landscapes, presenting more space to search and increasing the flexibility of the firm, are critical to firms in a dynamic environment [Grant, 1996]. In other words, a collaborative task facilitates both the search for information and its interpretation, and co-developed products are conducive to developing new knowledge [Powell et al., 1996].

In addition, co-development with external members complements a firm's internal capacity [Cohen and Levinthal, 1990]; [Arora and Gambardella, 1994], and this capacity evaluates the progress, while external collaboration provides knowledge and resources [Nelson, 1982]. Extended collaboration enables firms to view

some issues from different perspectives, which is essential in order for them to over-come routines and bias [Bierly and Chakrabarti, 1996].

A co-development strategy also helps organisations to benefit from external knowledge by searching extended landscapes. Although searching internal experience could yield new knowledge, it sometimes leads to suboptimal stable equilibria due to the depleting of recombinant materials. Collaboration can alleviate this dilemma by bringing new routines and novel knowledge for experiential search. In addition, decision makers tend to ignore their own failure and only learn from the failure of others, because they attribute their own failure to external reasons and others' failure to the efforts and actions of the latter [KC et al., 2013]. In this sense, collaboration facilitates vicarious learning, especially from failure experience, and provides new routines and novel knowledge which could reduce the ambiguity of self-failure. Therefore, collaboration could help these companies combine the internal and external knowledge together, enhancing the knowledge repository, and promote an understanding of the underlying mechanism and better interpretation of previous failure experience.

Furthermore, co-development could facilitate learning from failure by attending to weak clues. A key feature of collaboration is a high level of mindfulness, a condition where new product decisions and participants' actions reflect an integrated understanding of the agenda and constraint [Jassawalla and Sashittal, 1998]. Co-development processes are characterised not only by participants who think globally, act locally, and achieve high levels of openness about one another's motives and mindset, but also by participants who understand, accept, and internalise differences that exist and agree to focus on common objectives [Dougherty, 1992]. Accordingly, collaborative partners tend to focus on weak clues which are ignored by self-developers, and uncover the causal relationship of failure experience even though some of the information is not well captured.

However, a co-development strategy may not be useful for learning from success due to a knowledge sharing dilemma [Gulati, 1998]. Organisational decision-makers are usually over-confident about their existing knowledge when interpreting

success experience and conclude that further development of knowledge is unnecessary [March and Shapira, 1992]. As a result, they tend to ignore external suggestions and knowledge, and only search their own experience for knowledge recombination. At the same time, co-development strategy tends to redirect attention to the collaboration structure design instead of searching and interpreting success experience [Gulati, 1998]. Even though successes are normally causally unambiguous, simply refining existing assumptions and approaches is not sufficient, especially when the level of technological sophistication is high.

Taking these insights together, I argue that collaboration enables innovative organisations to learn more effectively from failure experience, and to improve learning outcomes by searching through the experience of failure more than through that of success.

Hypothesis 2a *Working with other partners improves learning outcome from failure more than when working alone.*

Hypothesis 2b *Working with other partners helps a company to reduce the incidence of future failure by learning from failure more than learning from success.*

2.2.5 Primary Ownership and Learning

The final element of a firm's innovation strategy is the decision of being the inventor or the licensee who acquires the preliminary products from other firms. As the innovation inventors, organisations shape a responsibility-taking culture. This culture and its norms influence the behavioural and cognitive development that organisations can achieve. In other words, inventors have greater search competence which discovers and refines experience.

Accordingly, inventors are more mindful of and attentive to all clues. However, this mindfulness has no privilege in searching success experience landscapes. Information on successful products is more transparent than that on failed products, and easily accessible to anyone. Both inventors and licensees feel confident about the adequacy of their knowledge since the knowledge generated during success is akin to the existing framework [Madsen and Desai, 2010]. Thus the searching

activities become redundant [Wildavsky, 1988] and the extra attention is wasted. On the other hand, mindfulness can pick up the weak clues of failure and trigger non-local searches for potential problems [Cyert and March, 1963]; [Weick, 1995]; [Rerup, 2009]; [Posen et al., 2018]. It also helps organisations to learn how to prioritise failure and raises the awareness of the insignificant signs of failure for future reference [Baum and Dahlin, 2007]; [Dahlin et al., 2018].

Moreover, innovation inventors also possess a monopoly of knowledge in the field and have the right to reject new ideas from outside [Katz and Allen, 1982]. This monopoly has the potential to disentangle the causal relationship of failure, so as to understand such relationships of previous failures and help to replace existing routines and knowledge with more useful and accurate ones in relation to failure experience [Haunschild and Sullivan, 2002]; [Henderson and Stern, 2004]. Only the inventors have the resources and ability to trigger problemistic searches and look for solutions or alternatives that could address the problems. With diversified knowledge, they can easily detect earlier errors so as to prevent subsequent expensive failure [Edmondson, 2011].

Overall, I posit that organisations cannot learn from success experience because the mindfulness of innovation inventors does not improve 'local search'. However, because inventors have a monopoly of unique knowledge and the ability to find the reason for failure based on weak clues, organisations could learn effectively from failure experience.

Hypothesis 3a *Working as innovation inventors improves learning outcomes from failure more than working as non-inventors.*

Hypothesis 3b *Working as innovation inventors helps reduce the incidence of future failure by learning from failure more than learning from success.*

2.3 Methods

2.3.1 Quantitative Analysis

2.3.1.1 Setting

Biotechnology firms need to constantly identify new opportunities and produce knowledge to survive and prosper. The process of drug development is the procedure of introducing innovation and is the key characteristic that identifies core competence in biotechnology firms. According to the US Food and Drug Administration (FDA), drug development is divided into five steps. The first is discovery and development, where scientists screen thousands of existing compounds, modify chemical structures, and use biological assays to test drug candidates in treating certain disease. The second step is preclinical research, in which scientists engage in both *in vitro* (in a test tube) and *in vivo* (in a living organism) tests of the drug candidates to discover potential toxicity and signs of success against the disease of interest. After this there are three phases of clinical trials to assess the safety and efficacy in humans. After completing all the clinical testing, the company can file a New Drug Application (NDA) to the FDA for approval.

This process of drug R&D is a lengthy process and most experimentations do not reach the approval stage. It is widely believed that the chemicals which do not reach potential launch status are failures [Henderson and Cockburn, 1994]. There are various reasons for failure, such as safety issues of clinical trials, side effects and poor target validation (no causal linkage between drug target and clinical impact). Previous research shows that nearly 90% of newly developed drugs fail in this process [Kola and Landis, 2004]; [Cannon and Edmondson, 2005]. Although this figure is somewhat higher than other innovations, the distribution of failures is highly skewed, as in other innovative industries [Scherer and Ross, 1980].

Even though failures challenge the status quo and trigger distant search, they are the natural running course of innovation, running the risk of going unnoticed or being deliberately ignored. In contrast, successes, confirming the adequacy of current knowledge and leading to local search, are treated as rare events and are considered important to a firm in triggering more extensive search for improving

the firms subsequent innovation outcomes. Therefore, whether innovative firms can learn more from failure than from success is unclear.

2.3.1.2 Data and Sample

The sample for this study consists of 37 top US biotechnology companies. To select the sample, I compiled the top 100 biotechnology companies list ranked by revenue from Compustat and excluded companies, based on three criteria: company location, main business and data availability. Among these top 100 companies, 38 companies have their headquarters outside the US, 12 of them do not develop drugs for human, and 13 of them have incomplete financial and R&D data.

Only US firms are included in this study, based on the following two arguments: first, data from the US is more transparent with high reliability, especially for public firms, making the data collection process smoother. Second, although scientific sharing is common in biomedical research, the exorbitantly high cost and the length of time of the R&D process prevent researchers from sharing information freely. Indeed, sometimes they are unwilling to share knowledge even when they are in strategic alliances [Rothaermel and Deeds, 2004]; [Pisano, 2006]. Some researchers find that geographically distributed units face the challenges of exchanging and acquiring knowledge [Ingram and Baum, 1997]; [Argote, 2012]. Therefore, only including American firms give me a unified and less complicated sample.

Another concern of my sample is the selection problem, which suggests that my sample is too small and will not represent the industry. Admittedly, since the 1970s, over 1600 new biotechnology companies have emerged, but many of them have disappeared due to failure or acquisition. Most of these firms are small start-ups from universities or research institutions, and the majority market share is controlled by only a few giant companies, most of which are located in the US. Even at the lower end of the top 100 biotechnology companies, many of them have fewer than ten drugs being developed. Therefore, although my sample is a small selection of firms in the biotechnology industry, it is representative of this industry and the data best serves the purpose of this paper.

Some people may argue that my sample is a comparatively successful group of companies and the behaviour of less successful companies may be different. Clearly, the selection of companies is the result of a data availability issue since the data of failed companies are less accessible. However, since this study is based on project-level data, firm characteristics have less effect on the analysis. In addition, the clinical trials and marketing activities often require the capital that only larger biotechnology firms can provide, and only larger and successful firms have the ability to continue and complete a project [Guedj and Scharfstein, 2004]. Therefore, even though the selection of relatively successful and large biotechnology companies is biased, it ensures that the data is complete.

To investigate my hypothesis, I compiled a unique dataset by combining data from several sources. I collected drug development data from Pharmaproject, clinical trials data from clinicaltrial.gov website, ventures alliance data from Deloitte Recap and financial data from Thomson One. The dataset consists of all drug development projects carried out by these 37 top US biotechnology firms in the period from 1987 to 2012. Altogether there are 2,240 projects, including 180 successes and 1,732 failures.

2.3.1.2a Dependent Variable

The dependent variable in my model is the hazard rate of an event - either failure or success - happening to a project. Success is defined as the launch of the product, whereas failure is defined as the suspension or discontinue of the product. The dependent variables include both the observed time (either censoring time or event time) and the indicator of the time corresponds to an event. The observed time is the period (in days) between the start of the project and the date on which the project finishes through success or failure. The observed time is right censored to 2012 since for continuing projects I do not have a recorded completion event. Some may argue that defining learning by a fixed outcome as success or failure could lead to an overly narrow representation of organisational learning [Kim and Miner, 2007].

However, it captures well the knowledge from the process of product development that operates to produce survival-enhancing learning in the biotechnology industry.

Mathematically, the hazard rate of failure or success is the instantaneous failure/success rate as *dt* tends to zero:

$$h_t = \lim_{dt \to 0} \frac{R(t) - R(t + dt)}{dt * R(t)}$$

where *R(t)* is the probability of no event (failure or success) before time *t*.

2.3.1.2b Independent Variables

Success and failure experience The independent variable measuring success experience is the cumulative time (in days) an organisation spends on developing launched drugs in the same therapeutic class until time *t*. The independent variable measuring failure experience is the cumulative time (in days) spent by an organisation on prior suspended or discontinued drugs in the same therapeutic class until time t. Drug development is a tedious process which normally takes more than ten years and varies dramatically across products. So merely including the count of a firm's prior success or failure as in the previous literature (e.g. [Madsen and Desai, 2010]; [Muehlfeld et al., 2012]; [Gong et al., 2017]), would not fully capture the variation of experience. Therefore, the cumulative duration of prior R&D failure or success would be more appropriate and justifiable to proxy experience.

Incremental innovation According to the US FDA, a drug that contains an active moiety which has not been approved by the FDA in other drugs is considered to be a new chemical entity (NCE). The NCEs are different from existing molecules and all of them must be reviewed by an advisory committee before being approved by the FDA, so developing an NCE drug is regarded as a process of radical innovation. On the other hand, developing a non-NCE drug is equivalent to a process of incremental innovation, which is the modification of existing chemicals and drugs. This process includes new dosages of existing drugs, combinations of existing chemicals, new indication and formula changes of existing drugs. I have coded non-NCE, which is incremental innovation, as 1, and NCE as 0.

Co-development Strategic alliances are ubiquitous in the biotechnology indus-try to alleviate uncertainty and prior research shows that an alliance improves or-ganisational learning by building absorptive capacity (e.g., [Powell et al., 1996]). However, how collaboration affects learning from success and failure separately is ignored to a great extent. Therefore, I include collaboration as a dummy variable. If the product is co-developed with other organisations, the variable is coded as 1. Otherwise, it is coded as 0.

Primary ownership As discussed above, many drugs in the biotechnology industry are co-developed by several partners and these partners share the val-ues and risks. However, the benefits and risks are not distributed equally among all partners, and inventors take extra responsibility comparing to licensees. First, the inventors of innovation normally have a monopoly of knowledge in the field [Katz and Allen, 1982]. Second, they face more uncertainty than the licensees, since they invest the majority of the capital and time. Third, they could capitalise more financial reward by charging licensing fees or by obtaining a larger portion of revenue. Therefore, primary ownership affects the willingness and ability to learn from experience. In this research, if the product is solely developed or invented by the focal company, the variable is coded as 1. Otherwise, it is coded as 0.

2.3.1.2c Control Variables

Several control variables are also included to account for factors other than organisational experience that might impact the hazard rate of failure or success.

Vicarious learning experience The organisational learning theory suggests that organisations develop knowledge not only from their own experience but also from observations of other companies experiences [Ingram and Baum, 1997]; [Madsen and Desai, 2010]; [Argote, 2012]. To control for vicarious learning, I measure industry experience by including R&D experience of other pharmaceutical companies, universities and research institutes.

In practice, since there are thousands of biotechnology and pharmaceutical

firms developing new drugs, and many universities and research institutes contribute during the process of drug development, many of their R&D activities are not publicly available and it is impossible to calculate the total time spent on R&D by these organisations. Therefore, I include the cumulative number of FDA approved drugs until time t as the proxy for industry experience.

Firm size Annual drug development at time t is included because larger firms tend to have more products under development and it could represent the size of the firm [Haunschild and Rhee, 2004]. Total assets at time t is also included to control for firm size.

R&D expenditure R&D expenditure may influence knowledge creation and retention by affecting the investment in equipment, human capital, and management, so I include annual R&D expense (indexed to 1980 dollars) at time t in my model following the previous literatures [Haunschild and Sullivan, 2002]; [Haunschild and Rhee, 2004].

Drug classes The project characteristics may affect knowledge recall and processing, so three additional drug development characteristics - delivery routes, delivery medium and drug origins - are also included.

Firm age Previous research has examined the effect of ageing on the organisational performance (e.g. [Tushman and Anderson, 1986]; [Henderson and Cockburn, 1994]). Therefore, besides the control variable discussed above, I also include the age of the firm at time t since it is not only related to technology advancement but also to the knowledge endowment of an organisation [Argote, 2012].

All the variables are also illustrated in Table 2.1.

	Variables	Description	Measurement
DV	Hazard rate of failure	The observed time period (in days) for the project to be a failure	The failure rate at time t+dt conditional on survival until time t for a certain project
	Hazard rate of success	The observed time period (in days) for the project to be succeed	The success rate at time t+dt conditional on survival until time t for a certain project
IV	Success experience	Firms learn from their successes (Ingram & Baum, 1997; Madsen & Desai, 2010)	Cumulative time a firm spends developing launched drugs in the same therapeutic class as the DV's therapeutic class until time t
	Failure experience	Firms learn from their failures (Madsen & Desai, 2010)	Cumulative time a firm spends developing drugs which failed eventually in the same therapeutic class as the DV's therapeutic class until time t
	Incremental Innovation	Innovation novelty strongly influences the factors that shape innovation performance (Garcia & Calantone, 2002)	Non-new chemical entity coded as 1, new chemical entity coded as 0
	Co-development	Strategic alliance could alleviate uncertainty and promote learning (Powell et al., 1996)	Co-development coded as 1, self-development coded as 0
	Primary ownership	Inventor has monopoly of knowledge and faces increasing uncertainty (Katz &Allen, 1982)	Solely developed or invented by focal company coded as 1, otherwise 0
Control	Industry experience	Organisations develop knowledge not only from their own experience but also through observation of others' experiences (Ingram & Baum, 1997; Madasen & Desai, 2010; Argote, 2013)	Cumulative number of FDA approved drugs until time t
	Annual drug development	Control for firm size (Haunschild & Sullivan, 2002; Haunschild & Rhee, 2004)	Annual number of drug projects at time t
	Total assets	Control for firm size	Company total assets at time t
	R&D expenditure	Control for firm size and knowledge creation and retention (Haunschild & Sullivan, 2002; Haunschild & Rhee,2004)	R&D expense each year(Indexed to 1980 dollar) at time t
	Firm age	Ageing has an effect on organisational performance(Tushman & Anderson, 1986; Henderson, 1994)	Company age at time t
	Delivery routes	Control for task characteristics that may affect knowledge dissemination. Four types: oral, injectable, inhaled and not applicable	Categorical variables from 1 to 4 for the 4 delivery routes
	Drug origin	Control for task characteristics that may affect knowledge dissemination. Four types: biological, chemical, natural product and not applicable	Categorical variables from 1 to 4 for the 4 types of drug origins
	Delivery medium	Control for task characteristics that may affect knowledge dissemination. Five types: capsule, powder, solution, tablet, patch	Categorical variables from 1 to 5 for the 5 types of delivery mediums

Table 2.1: Variable Construction

2.3.1.3 Model

Organisational learning researchers tend to aggregate individual project-level data to form firm-year panel and calculate the influence of previous experience on yearly failure outcomes. Compared with cross-section analysis, this captures the complexity of firm behaviour by controlling the impact of omitted variables and reduces measurement errors by observing an individual firm several times. In addition, panel data analysis accounts for the learning outcome when studying only routine and repetitive activities. However, with increasing uncertainty and complexity of learning during innovation, this method ignores individual outcomes and micro-foundations of learning.

If individual projects are similarly conditional on certain variables, firm-year panel data enables the possibility of learning from an individual firm's behaviour by observing others' projects [Hsiao et al., 1993]. However, if there are major variations among projects, which is common under the innovation context, acknowledging the heterogeneity instead of treating each project equally and supplementing observations with data on other firms is a more sensible approach.

In addition, aggregated data analysis often involves the representative agent assumption. However, if micro-units are heterogeneous, not only can the time series properties of aggregated data be very different but also the evaluation based on the data may be grossly misleading [Hsiao, 2005]. Furthermore, the prediction of combined outcomes using aggregate data can be less accurate than the prediction based on micro-project data.

To address these problems, I use project-level data and estimate the hazard ratio. In contrast to traditional analysis, survival analysis can predict whether and when an event will occur [Jenkins, 2005] and allows us to assess the conditional probability of an event given that the project is still in progress. Therefore, survival analysis can cope with right-censored data in which an event has not yet occurred and with time-series data, which has different time horizons [Jenkins, 2005].

The process of drug development I study is characterised by both situations: The data are right censored since many projects are still active after 2012. Further,

the time window is different for each project depending on when in the sample period it fails or succeeds. Last but not least, survival analysis allows me to apply a competing risk model to account for the heterogeneity of the event. While a drug development project can be either ongoing or completed, the modes of completion, which includes either failure or success, also need to be considered. Since the occurrence of one state removes other states from the pool of all risk states at that point in time, the different completion outcomes can be treated as competing risks.

Hence, I estimate the following equation:

$$h_t(t|x_{it}) = h_{0k}(t)exp(x_{ik}\beta_k)$$

where $h_k(t|x_{ik})$ is the hazard of project completion, $h_{0k}(t)$ is the baseline hazard (i.e. the hazard when all covariates are equal to zero), and x_{ik} is a matrix of covariates, including independent variables, control variables as well as company and therapeutic class dummy. The competing risks considered in the analysis are failure and success. Assuming the risks are independent, previous research has shown that the log-likelihood for the competing risk model is additively separable into K terms, with each one being a function of the parameters of a single cause-specific hazard [Narendranathan and Stewart, 1991]. Thus, the estimation of a single risk hazard considers finishing durations for other reasons than the one of interest as censored at the point of completion.

When developing a new drug, it is important to consider the therapeutic class to which the drug belongs. The failure rate varies greatly among different therapeutic classes owing to the innate molecular mechanisms [Kola and Landis, 2004]. To account for therapeutic classes, I include 31 indicator variables, corresponding to the therapeutic class guideline given by the FDA. In addition, I also include a set of company dummies to capture the variation in the trend of failure incidence across companies.

Missing data Since some of the dates are missing in my dataset, I estimate those missing data using different assumptions.

In the main model, I assume that all missing data follow the middle point rule, which is (see Appendix Table A.1):

1. if only the day of the current stage is missing, and the current stage is not in the same year and month as either the previous or next stage, assume the date is the 15th;

2. if only the day of the current stage is missing, and the current stage is in the same year and same month as the previous stage, assume the date is the middle point between the previous stage date and the last day of the month;

3. if only the day is missing, and the current stage is in the same year and month as the next stage, assume the date is the middle point between the next stage date and the first day of that month;

4. if both the month and the day are missing, and the current stage is not in the same year and month as either the previous or next stage, assume the month is June and date is the 30th;

5. if both the month and day are missing, and the current stage is in the same year as the previous stage, assume the date is the middle point between the previous stage date and the last day of the year;

6. if both the month and day are missing, and the current stage is in the same year as the next stage, assume the date is the middle point between the next stage date and the first day of that month; and

7. if the year, month, and day are missing, average the existing previous stage date and post-stage date to estimate them.

   In the robustness check section, different assumptions such as minimum, maximum and random will be considered.

Coarsened Exact Matching (CEM) The idea that different strategic choices may foster different types of learning highlights a source of potential bias that I cannot mitigate entirely: potential assortative matching between organisations and strategies. Here, the concern is that learning outcomes are influenced by the expectations of the innovative potential of each strategy. Although I control for as

many observable differences in the project-level characteristics as possible in my regressions, some aspects of these matches may be unobservable, introducing the potential for selection bias. From the pre-CEM section of table 2.2 we can see that different types of strategies have different characteristics. Similar drug classes share identical mechanisms to promote learning. This begs the questions of whether organisations are matched with projects and whether this may, in part, drive the observed results. To address this possibility, I perform CEM to make the projects in the pairwise strategic factors comparable in my analysis [Blackwell et al., 2010].

The matching process is as follows. The first step is to choose a relatively small set of covariates on which I would like to guarantee a balance between projects when searching for experience. In this study, I focus on three drug class types (delivery routes, delivery medium, drug origins; see Table 2.1 for more detailed information) and therapeutic classes. These four variables I use to balance the projects represent observable learning potentials that would be correlated with strategic choices. As the 'pre-CEM' column shows, the incremental innovation and radical innovation sub-samples are significantly different (at the 5% level) across the board for the full set of covariates as are the sub-samples for collaborative projects and non-collaborative projects, and inventors and non-inventors. These differences are reduced, however, 'post-CEM' , where none of the differences is significant at a level higher than 5%, suggesting a balance in the two pairwise sets of samples. The second step is to create a large number of strata to cover the support for the joint distribution of the covariates selected in the previous step. Thirdly, each observation is allocated to a unique stratum. Any stratum that has no project from either of the pairwise strategic factors is then dropped from the data.

Table 2.2: Project Characteristics Before and After Coarsened Exact Matching Procedure

	Pre-CEM		Post-CEM	
	Incremental Innovation	Radical Innovation	Incremental Innovation	Radical Innovation
Delivery Routes	2.56	3.21**	3.09	3.09
	(0.70)	(0.65)	(0.92)	(0.70)
Delivery Medium	3.66	2.21**	3.22	3.27
	(1.71)	(1.01)	(0.86)	(1.21)
Drug Origins	1.48	2.01**	2.05	2.02
	(0.67)	(0.11)	(0.22)	(0.15)
Therapeutic Class	13.8	15.1**	14.26	14.67
	(6.76)	(6.84)	(6.42)	(6.60)

	Collaboration	Non-collaboration	Collaboration	Non-collaboration
Delivery Routes	3.35	2.93**	2.92	2.82
	(0.84)	(0.66)	(0.72)	(0.89)
Delivery Medium	3.14	3.73**	3.19	3.12
	(1.07)	(0.98)	(0.93)	(0.76)
Drug Origins	1.89	1.24**	1.75	1.70
	(0.60)	(0.50)	(0.52)	(0.56)
Therapeutic Class	13.6	15.3**	14.95	14.26
	(6.99)	(6.61)	(6.76)	(6.38)

	Inventor	Non-Inventor	Inventor	Non-Inventor
Delivery Routes	3.03	2.53**	2.92	2.90
	(0.72)	(0.95)	(0.71)	(0.89)
Delivery Medium	3.01	3.84**	3.29	3.54
	(0.94)	(1.51)	(0.94)	(1.49)
Drug Origins	1.25	1.96**	1.75	1.72
	(0.53)	(0.62)	(0.52)	(0.59)
Therapeutic Class	12.9	15.4**	14.56	14.59
	(6.73)	(7.24)	(6.47)	(7.12)

Notes: The mean and standard deviation (in parentheses) are reported. The CEM procedure invloves matching on the delivery routes (Oral, injectable, inhale and not applicable), delivery medium (capsule, powder, solution, tablet, patch, deliver medium and not applicable), drug origin (biological product, chemical product, natural product and not applicable) and therapeutic classes.

**Indicates difference is significant at the 5% or higher level

The procedure is coarse because I do not precisely match the projects on covariate values, but rather, I coarsen the support of the joint distribution of the

covariates into a finite number of strata, and keep the strata if and only if projects from both pairwise strategic factors can be found in the same stratum. An important advantage of CEM is that researchers can guarantee the degree of covariate balance *ex ante*. However, the downside is that the more fine-grained the partitioning of the covariates, the larger the number of unmatched observations. In general, there is a trade-off between the quality of the matches and external validity.

2.3.2 Qualitative Analysis

Learning from innovation failures and innovation successes is quite different from learning from routine works as described in previous studies (e.g. [Adams, 1990]; [Baum and Dahlin, 2007]; [Madsen and Desai, 2010]). Failure is predominant in innovation, but rare in routine works. Therefore, the mindset is different between these two contexts. In addition, learning from innovation is a random trial and error since firms cannot automatically shift from existing routines to new ones [Nooteboom, 2004]. To make sense of the quantitative results and to determine why learning from innovation failures and successes are more difficult, I also conduct several semi-structured interviews with senior scientists. The qualitative data are collected from the research centres of US biotechnology companies in China. The main reason for collecting interview data from China instead of the US is data availability: most of the scientists in the private sectors are not willing to talk about their drug development projects in any sense and the company regulations also forbid employees to release any information concerning their R&D to protect their intellectual property in the field. The author could obtain in-depth information in China because of previous personal relationships.

Semi-structured interviews are conducted with 12 interviewees from 5 firms; they all are senior scientists with at least 10 years' industry experience of drug development. Each interview consists of three sections: the first section covers the interviewee's background, education, and work history. The second consists of a detailed narrative of the relationship among the projects the company has been working on without the technical elements. This section of the interview focuses

on the specific problems the firm faces during the R&D process as well as actions taken by the person with respect to these issues. The goal is to understand how the scientist sees the connections among projects and what the knowledge-sharing activities inside the firm are. In this section, I also explore their opinions on the sky-high attrition rate and whether they can learning by doing. The third section explores specific details and decisions that arise during the interview. Each interview lasts between 30 minutes to 1 hour.

To ensure data validity and minimise informant bias, the interviews are structured to gather specific information and conducted with non-directive questioning [Huber, 1985], and the participants are asked to focus on facts rather than speculation. For example, informants are asked about a specific project's failure and its relationship with other projects, and leading questions (e.g., was the failure blameworthy?) are avoided.

2.4 Results and Mechanisms

2.4.1 Results

Table 2.3 presents the descriptive statistics and correlations between the variables. As can be seen from the table, some very high correlations exist among certain variables. For example, the correlations between total assets and R&D expense (r =0.95) as well as between these two and annual drug development (r= 0.67 and 0.69 respectively). To avoid multicollinearity, I run the analyses by adding the variables sequentially and checked the fitness of each regression. These analyses show that multicollinearity is not affecting my main conclusions.

Table 2.3: Descriptive Statistics

Variables	Mean	SD	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13
1.Failure Experience	10.77	1.50	0	12.69	1.00												
2.Success Experience	8.14	3.76	0	11.37	0.49	1.00											
3.R&D Expense	11.14	1.63	6.51	14.13	0.63	0.52	1.00										
4.Company Age	22.38	15.41	0	81	0.31	0.40	0.46	1.00									
5.Annual Development	38.28	29.72	1	112	0.73	0.48	0.67	0.19	1.00								
6.Total Assets	12.85	2.23	6.77	16.79	0.64	0.59	0.95	0.55	0.69	1.00							
7.NDA Approved	11558.67	507.26	9374	12163	0.46	0.10	0.34	0.24	0.06	0.30	1.00						
8.Incremental Innovation	0.48	0.50	0	1	-0.03	0.03	0.03	0.09	-0.03	0.04	-0.14	1.00					
9.Co-development	0.48	0.50	0	1	0.01	0.05	-0.06	-0.10	-0.03	-0.06	0.06	0.10	1.00				
10.Innovation Inventor	0.85	0.36	0	1	0.16	-0.13	-0.06	-0.16	0.03	-0.09	-0.02	-0.03	-0.32	1.00			
11.Drug Origins	1.75	0.54	1	4	-0.04	0.00	-0.14	-0.11	-0.07	-0.14	0.06	-0.49	-0.01	0.02	1.00		
12.Delivery Medium	3.25	1.06	1	7	-0.04	0.05	0.00	0.02	-0.07	0.01	0.09	-0.02	0.12	-0.11	0.05	1.00	
13.Delivery Routes	2.90	0.75	1	4	0.00	-0.07	-0.06	-0.12	0.00	-0.07	0.03	-0.44	-0.04	0.04	0.24	0.03	1.00

Notes: The samples consist of 2,240 drug development projects from 37 biotechnology firms. See the Data and Sample section for details on the sample construction and Table 2.1 on variables definitions and construction.

In addition, the correlation between success experience and failure experience is also high. One reason for the fairly high correlation could be that occasions of both success and failure increase as the organisation gains overall experience. To determine whether success and failure experiences contribute information to the models independent of general experience, I conduct preliminary tests to estimate the impact of an organisation's general experience. Although total experience has a significant effect on performance improvement, models separating general experience into success experience and failure experience yielded a significantly better fit. This finding suggests that success experience and failure experience contribute independent information to the models despite their high correlations. Therefore, only failure experience and success experience are included in the models and the general experience is omitted.

Table 2.4 reports the partial likelihood method of estimation for the fixed-effects hazard model analysis of learning from failure and success. The second and third columns do not confine the sample to observations matched via CEM because I initially want to describe the learning pattern for all my samples. I examine both the timing and consequences of future failures as my dependent variables. I take the log value of experience, R&D expense and total assets and include three fixed effects, namely company fixed effects, year fixed effects, and therapeutic class fixed effects. We can see that innovation novelty and primary ownership moderate the relationship between experience and future performance.

Specifically, researching incremental innovation projects helps organisations learn from failure experience to reduce the failure risk. In addition, innovation inventors take more responsibility and have a higher tendency to learn positively from failure experience. Therefore, hypotheses 1a and 3a are supported but not 2a. An additional Wald test also shows that all three factors help prior failure experience enabling learning more effectively than prior success experience, even though the moderating effects of co-development on learning from failure and learning from success separately are not statistically significant. Therefore, hypotheses 1b, 2b and 3b are supported.

Table 2.4: Boundary Condition of Learning from Success and Failure

	CRA		CEM	
Variables	Risk of Failure	Risk of Success	Risk of Failure	Risk of Success
Failure Experience	1.46**	0.83**	1.12	1.04
	(0.17)	(0.15)	(0.20)	(0.18)
Success Experience	0.91**	1.04**	0.95	1.02
	(0.04)	(0.02)	(0.06)	(0.08)
Failure Experience * Incremental	0.87**	1.17**	0.82**	1.25***
	(0.05)	(0.07)	(0.08)	(0.07)
Success Experience * Incremental	1.01	0.98	1.02	0.97
	(0.02)	(0.03)	(0.04)	(0.03)
Failure Experience * Collaboration	0.92	1.06	0.98	1.05
	(0.05)	(0.06)	(0.07)	(0.06)
Success Experience * Collaboration	1.00	1.02	0.99	1.01
	(0.02)	(0.03)	(0.02)	(0.03)
Failure Experience * Inventor	0.78**	1.27**	0.78*	1.25**
	(0.07)	(0.08)	(0.11)	(0.10)
Success Experience * Inventor	1.09***	0.87***	1.09**	0.91**
	(0.03)	(0.04)	(0.04)	(0.04)
Company Fixed Effect	Y	Y	Y	Y
Therapeutic Fixed Effect	Y	Y	Y	Y
Year Fixed Effect	Y	Y	Y	Y
Loglikelihood	-2584.89	-2583.85	-1516.33	-1527.57
Chi-squared	543	544	313	307
N	2240	2240	1309	1309

Notes: Estimates stem from the partial likelihood method of estimation. Dependent variable is the hazard rate of future project failure or success. All models incorporate year fixed effect and firm fixed effect, as well as therapeutic class fixed effect. Robust standard errors in parentheses. Two specifications are used and they differ depending on whether CEM is performed before the regression. The sample size drops to only a half after CEM.

$***p < 0.001; **p < 0.01; *p < 0.05$

This estimation has not taken the possible self-selection of different learning modes based on unobservable into consideration. I therefore employ CEM to better understand the relationship between innovation strategy and learning outcomes. The last two columns of the table report the CEM estimates, balancing on drug class types (delivery routes, delivery medium, drug origins) and therapeutic classes to define the pairwise task samples. I find that my key results hold. This specification contains my full set of controls, company fixed effects, year fixed effects, and therapeutic class fixed effects. In terms of the magnitudes of the effects, if the current project has incremental innovation, one year of failure experience will decrease future failure incidence by 3.6% and increase the future success incidence by 4.3%. If the focal company is the inventor of the drug, one year of failure experience will decrease future failure likelihood by 4.8% and increase the future success incidence by 5.3%. We can also see that neither the coefficient of failure nor the coefficient of success experience is significant, indicating that both of them

are less likely to reduce future failure incidence or induce success incidence. These results are surprising and contradictory to those of previous studies alleging that organisations can learn from both failures and successes. Additional test and data in the next section will explain this in detail.

2.4.2 Disentangling the Mechanisms

The above results fail to uncover evidence of significant learning by firms from the observation of their own failures and successes. Coefficients estimating the effect of success experience and failure experience on future performance are indistinguishable from one. Although several properties of innovation strategy improve learning from failures, the fact that organisations do not demonstrably learning from the experience of failure or success is confusing and inconsistent with previous literature.

In addition, in this study I find that co-development does not enhance learning from failure. This is not a rejection of previous study on alliance and learning. The present study only takes intra-organisational learning into consideration, whereas most studies of alliance and learning focus on inter-organisational learning (e.g., [Powell et al., 1996]). So it is highly likely that co-development contributes to inter-organisational learning, but is unable to moderate intra-organisational learning appropriately. Limited resources, both cognitive and financial, may limit the ability of organisations to strike a balance between inside knowledge and outside knowledge. To understand this 'unlearn' situation and explore the effect of co-development, I conduct further qualitative analysis.

It is possible that learning from innovation failures and successes are drastically different from learning from routine works. Organisations cannot automatically abduct from an existing working modus operandi to a new one that in the future will turn out to be better but which now is not known [Nooteboom, 2004]). They do not know all the options and so random trial and error might be the only choice. To shed light on this issue and explain the above quantitative result, I conduct twelve semi-structured interviews with the senior scientists from five biotechnology firms.

A first potential explanation that arises from the interviews is that it takes a long time for the disease mechanisms which are the direct information from previous drug development to be applied in new drug development. One chief scientist from one of the largest biotechnology companies states:

> *Although pharmaceutical development is aimed at producing new drugs, the most valuable product of translation efforts is information about disease and drug mechanisms. This information is valuable because it informs drug development and it guides clinical practice. However, unlike other fields, the translational process in pharmaceutical development could not be accomplished overnight and there is a huge lag.*

Similar points and detailed examples are stated repeatedly by other scientists. One scientist recalls that several oncogenes which cause tumour growth have been discovered from clinical research that was conducted more than thirty years ago; however, our understanding of the mechanism of these genes is still limited and the discovery of drugs which target these oncogenes is still a long way off. My analysis only covers 25 years of data and this short time span potentially poses a challenge to find a real learning effect in the baseline model. This also promotes the importance of research in finding context factors which accelerate the learning process as has happened in this study.

Another possible reason that emerges from the interviews is that the methods for generating information may not be effective. Over the past three decades, the development of biotechnology has transformed drug development from low-throughput *in vivo* testing and medicinal chemistry optimization to aim-specific genetic manipulation and targeting in order to improve success rates. Although the latter reduces the cost and enhances the speed dramatically, it also produces more false positives since the assays are less rigorous. Before the 1990s, the standard approach for small-molecule drug discovery involved synthesizing and screening a relatively smaller number of compounds (normally fewer than ten at one time). Repeating the assessment cycle could establish a structure activity relationship and advance the structures of lead compounds through the chemical space. This approach prevents trial compounds from being confined to minor local optima and

avoids spurious clinical promise due to bias or random variation.

To reduce cost and to speed up the process, many drug development activities, especially early stage investigations, are conducted on small sample sizes and surrogate endpoints. However, this also comes at a cost, since small and less rigorous studies tend to produce more false positives, with misleading results. Several scientists complain in their interviews that in order to explore a vast, multidimensional landscape of agents, doses, disease indications and treatment schedules on a tight budget and within a limited time frame, they have to look for other intermediate indicators or use the minimal sample size. They think these short cuts may work for a while, but are detrimental to the field. One of the researchers became really upset when discussing this issue, telling me:

> *These managers only care about the success rate and profit, and they do not have any sympathy for the suffering victims. Science should not and cannot be measured by money, especially medical science. Money is only a number and it becomes useless when you can no longer enjoy your life. They are the real causes of unethical behaviours in medical research.*

Several neuroscientists from my interviews also mentioned that many spurious clinical promises are caused by bias or random variation in their field. They alleged that some information generated from previously launched products cannot be verified and this is hard to vindicate in the later drug development.

Most of the scientists also mentioned that a considerable amount of information and knowledge generated from drug development is not well-captured. For example, several researchers noted that the reporting and publication scheme is inadequate in pharmaceutical research. In one case, a senior scientist who used to work at a top US university criticized the fact that only limited pathophysiological data and theories are shared among peers even in the same company. If he wishes to publish the clinical results, he has to withhold some part of the methods or data to avoid leaking too much information. In this sense, knowledge is only embedded in a few particular people. If they leave the company, at least part of the knowledge will be lost.

Furthermore, negative or inconclusive studies are even more susceptible to this problem. It could because that scientists are afraid of admitting failures or they think that only positive and conclusive results can diminish the off-label use of a licensed drug and be used to compile a clean narrative for investors. For example, several researchers in my interviews mentioned that the reporting and publication scheme is inadequate for failures in pharmaceutical research. One senior scientist said that it is not possible to publish failed clinical results since no journal is willing to accept negative results. Therefore, other members in the same firm may not learn from the potential mistakes and may approach future tasks in ways that are similar to the approach which leads to failure.

The reasons presented above help explain why it is not easy for biotechnology firms to learn from failures or successes and to build a circumstantial case in favour of interpreting the learning process during innovation. However, these reasons do not enable me to reject some potentially relevant versions of theory - such as the mindfulness of researchers in interpreting information - nor do they allow me to learn about the degree of each effect.

2.4.3 Robustness and Sensitivity Check

In addition to the main analyses reported above, I also conduct several supplemental tests to assess whether the patterns of results are robust to alternative specifications and samples. First, I use different estimation methods for missing data given that previous estimations might distort the data and generate false significant results, which would not be reflected in the real situation. As discussed above, my primary analyses assume that the missing data are random. Given that failure experience and success experience may vary in length, it is necessary to assess whether my arguments regarding the relative effects of learning from failures compared with learning from successes hold with respect to different estimations of the missing data. To rule out this possibility, I use different assumptions to calculate the missing data (see Appendix Table B.1).

- Minimum: Assume the missing stage date is one day later than the previous

stage date

- Maximum: Assume the missing stage date is one day earlier than the later stage date

- Random: Assume the missing stage date is a random date between the previous stage and later stage

All these estimation assumptions yield similar results.

Second, another concern regarding the missing data is that these missing data are not random but are related to innovative performance. One possibility is the decision makers pay less attention to some projects and choose not to record them if they think these projects are less likely to succeed, or the loss of key scientists or important achievement records could have caused the loss of date. Regardless of the reason, missing data indicates that particular knowledge might be lost. To rule out these possibilities, I test the models by including projects with non-missing data only. This sub-sample contains 1919 products with 1579 failures and 143 successes. The main results do not change qualitatively when projects with missing data are discarded.

Third, given the duration of successful drug development is ten years, generally, and a few projects in my dataset have less than a year of R&D duration, it is possible that these data are sometimes non-reliable since finding pre-clinical information record is problematic [Kola and Landis, 2004]. In addition, it is unlikely that organisations will allocate limited attention to these projects owing to their brevity and the lack of promising results. Therefore, to further test whether my findings are affected by this, I drop 51 projects that have a lifespan of less than one year. The results are similar to the main findings, suggesting that these short-duration projects do not twist my results.

Fourth, many researchers have found that the value of prior experience depreciates over time in such a way that recent experience is more valuable than the older experience (e.g., [Ingram and Baum, 1997]; [Madsen and Desai, 2010]; [Argote, 2012]). This is caused by the organisation members' exit [Argote et al., 1990]

or changes in organisational processes or structures [Pablo Martin de Holan, 2004]. Several methods have been developed to model knowledge depreciation (e.g. [Baum and Ingram, 1998]; [Haunschild and Sullivan, 2002]).

Typical values assigned to the discount factor are 1 (assuming that knowledge is non-depreciating), the age of experience (assuming that knowledge depreciates linearly), the age of experience squared (assuming that knowledge depreciates rapidly), and the square root of the age of experience (assuming that knowledge depreciates slowly). To examine whether knowledge depreciation can reduce the learning effect and therefore change my results, I replicate my analyses using the same sample but include the discount factors. None of the three commonly used discount factors change my main findings.

Last but not the least, to further examine whether the findings of the innovation inventor as a moderator are susceptible to the self-development of products, I also estimate the effect of primary ownership in a partial data set including only collaborative drug development. The results are similar to those of the main models, suggesting that the difference between organisational learning from success experience and failure experience can be driven by primary ownership.

2.5 Conclusion and Contribution

2.5.1 Conclusion

Despite the wealth of research on organisational learning which enables innovation and promotes economic development (e.g., [Alegre and Chiva, 2013]; [Moustaghfir and Schiuma, 2013]), we have limited knowledge of innovation learning and know little about how innovation strategies enable learning from past failures and successes. The present study aims to address this gap by analysing, first, how particular properties of innovation strategies - innovation novelty, co-development, primary ownership - improve learning from failure. In contrast to the majority of the empirical research in organisational learning, which discusses learning from general experience, I focus on improving learning from failure. Many

institutional systems, such as Total Quality Management and soliciting feedback from customers by using the United Stated Navy and Colour Code System by Boeing respectively, have been put in practice to give additional attention and resources to surface failures alongside the routine operations [Edmondson, 2011]. Learning from failure is more salient in innovation since 50-90% of innovation projects fail [Castellion and Markham, 2013] and diversified knowledge can be obtained from failures than from successes. This work, using Coarsened Exact Matching to improve the estimation of causal effects by reducing the imbalance in covariates between different groups, looks at the fine-grain project-level data and provides theoretical and practical implications for improving learning from failure.

Second, I also test whether these strategies improve more by learning from previous innovation failures than they do by learning from previous innovation successes. In doing so, this work explicitly distinguishes between success and failure experience in the context of innovation. The results yield strong evidence that all three innovation strategies help organisations learn by observing their own failure more than by observing their own successes.

Although the coefficients estimating the effect of failure experience and success experience on future performance are indistinguishable from zero, this should not be interpreted as evidence that organisations cannot learn from failures or successes to improve their performance. However, the fact that organisations do not experience demonstrable learning from failures or successes suggests that learning from innovation success and innovation failure separately is far from being an automatic process. Additional qualitative analysis explores the reasons underlying this.

2.5.2 Theoretical Contribution

This study contributes to existing theory in several ways. First, it introduces two innovation strategies as the accelerators of learning from innovation failures. Previous research shows that the complexity of an innovation task determines the knowledge strategy and subsequent performance of the firm [Bierly and Chakrabarti, 1996], but how different types of innovation strategies fa-

cilitate knowledge search between different performance outcomes is understudied. This paper argues and proves that conducting incremental innovation or acting as an inventor enhances learning from innovation failures. These factors not only improve learning from failure, but also lead to stabilizing learning from success since the results of current study do not show there is a considerable improvement in learning from success. The results suggests trade-off between learning from failure experience and learning from success experience and confirm that the aggregated experience may not represent the real learning result. This confirms the discussion of interpreting previous results cautiously in terms of learning from aggregated experience [Madsen and Desai, 2010]; [KC et al., 2013]; [Muehlfeld et al., 2012]; [Gong et al., 2017]. These two factors also drive entrepreneurship, since organisational learning acts as an important driving factor of entrepreneurial efforts and has been associated with firms' greater ability to innovate [Burgelman, 1983].

Second, this paper demonstrates how previous failure experience is expected to influence the efficiency of innovation. Failures motivate decision makers to challenge the status quo and engage in deep and mindful reflection involving complex thought processes [Madsen and Desai, 2010]. Organisational search for knowledge in response to failure helps to correct problems, challenge old assumptions and innovate [Sitkin, 1992], [Eggers, 2014]. During this process, innovation novelty and primary ownership not only provide resources to increase the chance of searching for new knowledge, but also bring mindfulness into the search process to attend to minor clues and the roadmap of the gap [Levitt and March, 1988]. In other word, these two innovation strategies are more likely to produce the necessary conditions for organisations to build capabilities that improve future performance.

Third, this work extends the literature on intelligent failures [Sitkin, 1992]. Intelligent failures are those in which expectations are not met, but something useful for the future is learned [McGrath, 2001]. These failures are necessary experimental steps for innovative outcomes, but learning from them is not straightforward. Organisations need a special schema to take advantage of them. For example, Eli Lillys failure parties since 1990, which honour intelligent experiments that fail to

achieve the desired results, redeploy valuable resources for new projects and kick-start many new discoveries [Edmondson, 2011]. This study attempts to uncover some strategic factors which can enhance learning from intelligent failures.

Finally, this work also presents a new analytical method to analyse organisational learning by taking project heterogeneity into consideration. Using aggregated data to infer individual behaviour by previous researchers means relying heavily on the representative agent assumption [Hsiao, 2005]. However, the diverse conditions of each innovative project render this assumption defective. This study constitutes the first project-level analysis of organisational learning and provides the micro foundation of learning variation.

2.5.3 Implications for Practice

This paper highlights that learning from what went wrong in past innovation processes is possible but challenging [Edmondson, 2011]; [Eggers, 2012]. Failure is sometimes difficult for organisation members to cope with because failures are often stigmatized and organisation members frequently refuse to acknowledge failure and refrain from communicating about it [March et al., 1991]; [Madsen and Desai, 2010]. This is especially salient in innovative industries because of the causal ambiguity.

Nonetheless, given the number of failures during innovation and their central role in organisations shown previously [Haunschild and Rhee, 2004]; [Baum and Dahlin, 2007]), the inability to learn from failure may deprive an organisation of the opportunities for improvement. Therefore, developing superior strategies for coping with failures should enable organisations to reap substantial gains. Thus this study provides some solutions to treat failures as invaluable learning opportunities and encourages the open sharing of information about them.

Furthermore, this study provides another way to overcome Erooms Law (drug discovery is becoming slower and more expensive over time, despite improvement in technology) [Scannell et al., 2012]. Although the paper published in Nature Reviews Drug Discovery in 2012 discusses the problems and solutions in great de-

tail, the authors dismiss the importance of using management to lessen the threat. Admittedly, a fundamental change in the business model is a painful process and the outcome is ambiguous. But based on the results from this study, organisations could learn more easily and effectively from previous failure experience and improve R&D efficiency by carefully choosing the projects. This could help reduce the growing panic over the status quo.

Last but not least, this study also has policy implications. Innovative companies are an important contributor to economic growth and government policy incentive research, development and innovation. There are various kinds of policies, and encouraging R&D knowledge spill-over is commonly used by many modern governments. However, no explicit policy has been targeting innovation learning. Instead of focusing collaboration to facilitate knowledge-sharing across organisations, other policies such as balancing incremental innovation and radical innovation or allocating more resource to inventors than to licensees could promote learning from previous failures and enhance innovation efficiency.

2.5.4 Limitations and Directions for Future Work

While I am able to investigate how to improve learning from failure during innovation, the results from this study should be interpreted with caution. Although incremental innovation can enhance learning from failure experience, organisations that only focus on incremental innovation will fall into the competence trap of just developing current and short-term competence, and will lose the chance to move to new and superior competence [Levinthal and March, 1981]. In this sense, radical innovation is necessary for organisations in the long term, although learning from failure is compromised. Therefore, a balance needs to be achieved between radical innovation and incremental innovation, and only by doing this, can organisations have a sustainable competitive advantage.

A similar condition applies to primary ownership. Although being an inventor is beneficial for learning from failure, it is impossible to invent everything due to limited resources and attention. Co-invention or co-development could be a com-

promise and organisations could benefit from vicarious learning. In addition, co-development also reduces uncertainty and cost and diversifies any organisation's portfolio, thereby increasing survival [Nelson and Winter, 1982]; [Powell, 1990].

In addition, not all failures make the same contribution to learning, since the magnitude of failure varies and future work could include the financial data of each drug development process to account for the magnitude of failure and see whether it will produce the same result as this study. This would not only control the magnitude of failure, but also test the learning curve framework in innovation learning, using the most acceptable outcome [Argote and Epple, 1990].

Last but not least, the research context of this study is drug development in biotechnology industry, and many people may argue that other technological innovation contexts may result in different conclusions. Admittedly, in all innovative industries, the ability to develop a new product quickly, effectively and efficiently is the most important factor driving a firm's success. Even though the cycle time of drug development is longer and the failure rate is higher than in other industries, reducing the innovation cycle time and reducing failure as a dependent variable, as used in this study, are common goals for all technological innovation. However, this does not guarantee that the conclusions in this study apply to other innovation contexts. So further studies on other research contexts, such as chemical or software development, to test the learning hypotheses in this paper, would be welcome.

This study demonstrates that innovation novelty and primary ownership can enhance learning from failure and improve learning from failure more than learning from success. Collectively, these findings suggest the need to further explore organisational learning practices associated with innovation failure and to determine how organisations may be able to reap the benefits of failure without exposing themselves to its undue cost.

Chapter 3

Towards an Understanding of Learning by Innovating: Evidence from Drug R&D

3.1 Introduction

Product innovation drives economic growth by raising product quality or increasing product price [Syverson, 2011]. It is aimed at entering new markets or at refocusing a firm's efforts towards growing demand segments [Acemoglu and Linn, 2004]; [Balasubramanian and Sivadasan, 2011]. Rapid technological changes and complex and unpredictable environments require that firms maintain the capacity to continuously improve in response to these conditions [Balasubramanian and Sivadasan, 2011]. The impact of the knowledge economy on today's business has promoted a growing recognition among the companies about the need to cultivate learning by doing. Knowledge spillovers and learning by doing are the workhorses of endogenous growth and the principle sources of passive productivity growth [Arrow, 1972]; [Thornton and Thompson, 2001]; [Levitt et al., 2013]; [Aghion and Jaravel, 2015]. Experience allows decision makers to identify opportunities for process improvement and shape the productivity trajectory. Hence, learning by innovating, learning by doing during product innovation, is increasingly being recognised as an important precondition for innovation

success [de Ven and Polley, 1992]; [Aghion and Jaravel, 2015].

Learning by innovating is crucial because it implies that both learning by doing and product innovation can reduce uncertainty and expand a firm's total factor productivity (TFP) [Thompson, 2010]. Despite this progress, there has been little empirical investigation into whether and under what conditions learning by innovating improves productivity. One important reason for this situation is the lack of a systematic instrument to measure productivity in the technological innovation context. One review article reveals that even though productivity is a relatively straightforward concept, the components and structure of the productivity construct are not well explicated [Syverson, 2011]. While the notion of learning to enhance productivity is generic in nature, its relationship in the technological innovation context might be different. Learning by innovating is a process of abduction instead of a routine-based process that responds to experience by repeating behaviours that have been found to be successful [Nooteboom, 2004]. Innovative firms project an existing practice into a context that is sufficiently similar to have a chance of success, and allow for some predication of likely results. The context also needs to be sufficiently different to yield a novel experience. Therefore, empirical studies that address the relationship between learning by innovating and productivity are needed.

Learning is bounded and knowledge gained in one setting may not be exploitable in another [Jovanovic and Nyarko, 1996]; [Thornton and Thompson, 2001] [Egelman et al., 2016], this may create problems for learning by innovating since different knowledge is generated along the innovation process. During the product innovation process, endogenous uncertainties create pressure to invest and resolve, and staging innovation becomes a common strategy to manage the risks associated with major innovation projects [Myers and Turnbull, 1977]. By staging the innovation process, firms can create options that convey the right (but not the obligation) to make further development. In industry practice, upstream research is most often separated from downstream development [Karlssona et al., 2004]. Upstream research is defined as the invention of new science and capturing of new know-

how; whereas downstream development is about applying proven technologies to commercialise products to achieve business objectives. Therefore, knowledge generated in different stages is unique but related, knowledge spillovers between stages may give firms an incentive to choose multi-stage options to reduce uncertainty and improve efficiency. Therefore, the boundary condition of learning by innovating, especially whether there is knowledge spillover across different stages during product innovation also need addressed.

In this paper, I gather rich data on R&D from the biotechnology industry to address questions regarding whether learning by innovating drives productivity growth and whether knowledge originating from one stage of learning by innovation could be exploited in another stage. To answer the first question, the role of innovation experience is evaluated in shaping and driving productivity. To answer the second question, the influence of experience from different stages on the likelihood of future failure in each stage is measured. In order to overcome barriers to measuring the project-level response to the experience from different stages, unique features of the biotechnological R&D setting, including the observability of development milestones, and the separability of R&D stage is leveraged. The core data set covers in incredible detail the innovation process of over 1,900 drugs over the courses of 25 years. There are numerous opportunities for learning, as drugs share similar therapeutic targets and production teams.

The US biotechnology industry is well suited to the investigation for several reasons. First, the biotechnology industry is under growing pressure of losing revenue due to patent expirations, increasingly cost-constrained healthcare systems and highly demanding regulatory requirements. Indeed, the industry's price/earnings ratio, a measure of the current valuation of the industry, has decreased below that of the S&P 500 index and has remained more or less flat for the last 10 years [Paul et al., 2010]. Therefore, without a drastic increase in innovation productivity, the industry cannot sustain the current business. Second, rapid technological advancement and unprecedented investment have resulted in clear stage division, and this is more apparent in biotechnology than in any other industry. Although

sequential development is not unusual in innovation, the clear division of new drug development stages, including upstream research process and downstream development process, followed by all decision makers, makes it easy to collect and analyse data. Third, learning is a potential source of productivity growth in this industry. Nearly 90% of newly developed drugs fail in the experimental stage, and it takes more than 10 years to develop a new drug [Kola and Landis, 2004]. The knowledge and skills acquired through experience could reduce the long development duration and sky-high cost.

The primary finding of this paper is that firms learn from their own innovation experience, but the benefits of experience mainly come from experience gained in the same stage of product innovation experience. Specifically, upstream research can only benefit from learning obtained from other upstream research processes but not from downstream development processes. The very act of conducting upstream research allows firms to identify opportunities for improvement of the same stage of the process to increase productivity. The same logic applies to the downstream development process.

In terms of magnitude, I estimate that upstream research experience decreases the failure rate of the research stage of innovation by 7.4% on average, and downstream development experience decreases the incidence of failure of the development stage of innovation by 3.8% on average. The enhancement of success gives biotechnology firms an incentive to simplify the R&D process inside the firms. Accordingly, both my data and industry anecdotes indicate that biotechnology firms tend to carry out only the upstream research process in house, and outsource downstream development or forming alliances to execute downstream development.

Finally, I examine the mechanism behind the observed stage-specific learning by innovating. Although I cannot rule out other possibilities, the boundary of knowledge spillover across stages during innovation appears to be driven partially by additional attention. Extra attention, which recognises and attends to weak cues to develop novel knowledge, improves cross-stage learning for upstream innovation from downstream development experience.

The findings in this paper add new insights to the literature that has been attempting to move beyond a progress function that simply relates learning by doing and product innovation to productivity. One distinction of this study is that I combine learning by doing and product innovation as reflected in productivity analysis. Previous research has demonstrated that both learning and product innovation drive economic growth [Arrow, 1972]; [Syverson, 2011], but few works have combined these two elements to illustrate their combined impact on productivity. This paper presents the first empirical study to analyse how learning by innovating affects subsequent productivity. The second distinction is my focus on the mechanism of learning during product innovation. Much of the economic and management research on learning by doing has focused on the finding of improved productivity with increasing experience, but what drives the improvement is still unclear [Thornton and Thompson, 2001]. This work, by using a unique dataset, separates experience and productivity into upstream and downstream, and addresses the relative effects of stage-specific learning. The third distinction is that this study illustrates the distinctive 'search modes' and knowledge spillover boundary [Levitt and March, 1988]. The fundamental distinction between upstream research and downstream development during product innovation is their search and experimentation process. The process of upstream research requires large scale experiments and variation, whereas the process of downstream development enhances productivity through choice, execution and variance reduction.

By fleshing out several details of stage-specific learning in the biotechnology industry, we understand more about the nature of knowledge spillovers. In this way, experience gained from one stage does not cause efficiency enhancement in another stage. Knowing more about such boundaries enables us to better understand the knowledge sharing and transfer process [Meyer and Goes, 1988]. Furthermore, it indicates that a firm not only needs to combine new product development with different types of knowledge at different stages [Madhavan and Grover, 1998], but also need to use a different process of knowledge sharing and knowledge transfer as a knowledge management process [Levitt et al., 2013].

Though this article focuses on the drug innovation process in the biotechnology industry, it seems that the prevalence of stage-specific learning extends beyond the drug R&D process. Many new product developments, such as new electronics, intricately involves multi-stage investment and decision making. In addition, while I focus on stage-specific learning within firm as the main construct in this paper, knowledge spillovers and specificities are also likely to determine the firm boundary. Azoulay (2004) and Macher and Boerner (2012) discuss how the knowledge-based view (KBV) can be applied to firm boundary decisions and efficient organisation approaches. During contractual arrangements, residual control rights over knowledge decide investment incentives and ownership rights [Grossman and Hart, 1986], [Hubbard, 2008]. These theories could in principle contribute to the firm boundary debate and complement the capability mode [Argyres and Zenger, 2012].

The remainder of the article is organised as follows: Section II provides background information on the biotechnology industry and the drug innovation process, and Section III discusses the mechanisms for learning by innovating and stage-specific learning. Section IV describes the data used in this study. Section V presents a model of staged-learning and discusses the empirical strategy. Section VI and VII provide the estimation results of learning by innovating and stage-specific learning respectively, and Section VIII examines the mechanisms behind staged-specific learning. Section IX presents the conclusion.

3.2 Institutional Background

The biotechnology industry is a young science-based industry with its foundation dating back to the pioneering work by Watson and Crick, who discovered the structure of DNA as a double helix in the early 1950s. Drug R&D, lying in the centre of innovation activities, is typically a sequential process. At several points in the process, a biotechnology firm tests and reviews the status of the drug and makes a decision on whether to continue with its development. In general, the decision

depends on the potential therapeutic benefits, the severity of adverse reactions and the projected estimates of a future revenue stream.

Drug R&D is a multi-stage process with a regulatory approval process that proceeds in well-known stages. Generally, a drug goes through the following stages to the point of the FDA approval: Initially, chemists and biologists synthesise new compounds using existing concepts and screen it for pharmacology activity. After identifying a set of promising compounds, researchers test their pharmacokinetic and pharmacodynamic properties in animals. If, after these pre-clinical tests, the drug is still considered to have a special molecular target and desirable therapeutic effects, it is filed with the FDA as an Investigational New Drug Application (IND). After receiving authorization from the FDA, the drug can undergo clinical testing 30 days after filing the application. Clinical trials normally occur over three distinct phases, each of which contributes information on safety, efficacy, and proper dosage strength and form.

In phase I clinical trial, the drug is tested with a small number of healthy volunteers to establish safe dosages and to gather information on the absorption, distribution, metabolic effects, excretion, and toxicity. In the next phase, phase II, the drug is administered to a larger number of patients who have the targeted disease or condition and evidence on safety and preliminary data on efficacy are reported. The final clinical trial, phase III, typically consists of large patients population and is designed to firmly establish the efficacy and to identify side-effects that occur infrequently. The large sample size (usually in thousands) increases the likelihood of the actual benefits being found to be statistically significant and the testing approximates the manner in which the drug would be utilised after marketing approval.

Once the clinical trials have been completed and the drug developer believes that there is sufficient evidence of safety and efficacy, the firm will compile the results and submit a New Drug Application (NDA) or a Biological License Application (BLA) to the FDA for review and approval. The drug can only be sold commercially after the FDA formally approves it. In this step, developers continue to provide post-approval safety and monitoring data (e.g. Phase

IV clinical trials, adverse event reporting, and other post-marketing surveillance). The entire drug R&D cycle usually takes over 10 years [DiMasi et al., 1991]; [Henderson and Cockburn, 1996].

Only around 10% drug R&D projects make it to the FDA approval stage and the process fails for a variety of reasons. No sign of success against the disease of interest and potential toxicity are the most common reasons for pre-clinical failure, whereas safety and efficacy concerns account for the vast majority of clinical trial project terminations. A research conducted by AstraZeneca found that about half of the clinical trial safety failures are related to the drug's primary biological target, while the other half of safety failures are attributed to off-target side effects. The most common reasons for efficacy failures are poor target validation (no causal linkage between drug target and clinical impact), dosage limitations, poor selection of indications and weak evidence from previous phases [Cook et al., 2014].

3.3 Theoretical Background

Product innovation is a process of combining equipment, work force, task specification, material, and information to produce a product or service [Utterback and Abernathy, 1975]. This process lies at the core of the creation and maintenance of competitive advantage. It is a highly risky process and most of the products fail along the way and cannot generate profit for companies [Mansfield, 1981].

Innovation in product quality is a stochastic process of exogenous random events that represents an independent and equally likely draw from an underlying probability distribution of possible actions [Hannan and Freeman, 1977]. Learning by doing, where firms create, retain and transfer knowledge, constitutes firms' capability of mitigating innovation uncertainty and serves as the driving force for firms' productivity growth [Lucas, 1993]; [Covert, 2014]. Learning by innovating has been documented in the literature since Schumpeterian (2010) 'creative destruction' and 'novel combination'. This departures from learning by doing in the sense

familiar to economists refers to the downward shift of an average cost curve as a function of cumulative uninterrupted production since it is dominated by uncertainty. If uncertainty precludes the 'substantive' rationality of choosing the best available option, we need a 'procedural' or heuristic rationality, in the form of some *modus operandi* that is likely to succeed [Simon, 1979]. In other words, a firm could abduct knowledge from adjacent possible projects and put that into a similar context for prediction while it is sufficiently different to yield novel experience and indications [Nooteboom, 2004].

Owing to the uncertainty during the process of product innovation, decision makers usually create options and stage the process to reduce risk. This option strategy secures a firm's claim to commercialise the product and helps mitigate loss at the same time. Different stages produce different experience that drives economic growth. These experiences are transformed into knowledge that is incorporated into routines and operating practices [Levitt and March, 1988]. Economists frequently make references to the limitation of knowledge spillovers as the principle source of the learning boundary [Thornton and Thompson, 2001]; [Macher and Boerner, 2012]; knowledge originating from a certain stage may not be exploitable in another stage. Therefore, aggregating experience from different stages may not only underestimate their contribution to productivity growth, but also create a no-growth trap if too little knowledge is transferable across innovation stages [Jovanovic and Nyarko, 1996].

The product innovation process can be divided into two distinct stages: upstream research and downstream development [Karlssona et al., 2004]. In the upstream research stage, decision makers usually experiment with a large quantity of alternatives to invent new science and to capture new know-how, and gain returns that are uncertain, distant and often negative. To solve problems and codify experiences from this stage, decision makers need to search for novel technologies in the area in which they have no prior experience. The main goal of this stage is to gain increased understanding of a phenomenon and utilise it in the current product and process. In the downstream development stage, on the other hand, decision

Table 3.1: The Key Difference Between Research and Development

	Research Stage	Development Stage
Time	Upstream	Downstream
Originality	Discontinuous 'jump' in knowledge	Continuous evolution of existing ideas
Knowledge Depth	Highly specialised	Broad Knowledge
Goal	To understand a phenomenon or to search for new elements of technology	To commercialize the product and to satisfy the customers' needs

makers tend to refine and extend the existing competence and paradigms and apply proven knowledge to commercialise products. They search for new dimensions in areas that enable them to build upon their established base. Therefore, the similar response generated by encountering routines during the downstream development stage is the source of continuity [Nelson and Winter, 1973]. The main goal of this stage is to apply knowledge from the previous stage to further develop and manufacture the products in order to improve product reliability and customer suitability. Therefore, research aims to develop new knowledge, whereas development aims to apply scientific or engineering knowledge to expand it and integrate the knowledge for commercial applications [Karlssona et al., 2004]. The distinction between upstream research and downstream development lies in four factors: time, originality, knowledge depth and goals (see Table 3.1 for more information).

In the drug development process, the upstream process is the pre-clinical stage, which focuses on the research process and explores new knowledge to create a prototype product that can undergo further testing and development. In this discovery process, researchers identify biological mechanisms that impact diseases and symptoms. For instance, they may want to develop a drug that inhibits or enhances the functioning of a particular target, such as a gene or its coded protein. Having identified the potential target, scientists then screen potential compounds which have some desired action on this target. Researchers then test the pharmacokinetic and pharmacodynamic properties of these promising compounds both *in vitro*(in a test tube) and *in vivo* (in a living organism). This process is

extremely uncertain and requires exploratory experimentation with new alterna-tives [Gilsing and Nooteboom, 2006]. Generally, scientists pursue things that might come to be known and face the risk of numerous failures.

On the other hand, the downstream process involves the clinical trials, which focus on the development process and exploit the knowledge gained through prior exploration. Clinical trials have three phases. Phase I and Phase II test the safety and dosage by using healthy volunteers and people who have the condition (disease) of interest, respectively. Phase III is essentially a large-scale version of Phase II trial, usually involving more participants who are tracked over a longer period of time. These clinical trials are less uncertain compared to pre-clinical research. Developers only need to refine and extend existing competencies and knowledge.

Literature attempting to investigate innovation productivity has mostly focused on measuring innovation activities as a whole [Karlssona et al., 2004]. However, the previous attempts to do so have failed for two main reasons. Firstly, the ex-pected outputs from research and development are totally different. The purpose of research is to develop new knowledge or technology, whereas the purpose of devel-opment is to combine knowledge to produce a new product to satisfy customers' needs. Secondly, research activities are very different from development activities, especially in knowledge elements; research requires specialised expertise in certain areas to expand current knowledge to new areas. Development requires broader knowledge bases for cross-functional understanding. Therefore, the learning pro-cess and outcome might be different in research and development stages.

3.4 Data

The main empirical goal of this paper is to identify how experience gained from different stages influences the productivity of each stage. Therefore, I needed a comprehensive data set with project development histories and disclosures. The primary data set was taken from Pharmaprojects, a commercial database that records information about the R&D process in the biopharmaceutical industry. The

Pharmaprojects database aggregates information from public records (e.g. patent fillings, company press releases, financial filings, clinical trial registries, FDA submissions, etc.), and employs professional analysts who curate the content. The database covers the progress of new drug candidates as they enter commercial pharmaceutical research and development programmes; it also tracks their progress from pre-clinical development up to market launch, or to discontinuation if the drug fails at any stage. For each innovation process, the database records the start and end dates of each R&D phase. Most projects fail, so Pharmaprojects also documents the date and stage of failure. The dataset also contains therapeutic class status, biological targets of the drugs and patent information.

My main analyses require this information to construct full research and development histories for each drug. The histories included which firms were actively developing the drug, and what stage of development (preclinical, phase I/II/III clinical trials, registration, approval and launch) the project was in at any given point in time. They also included event dates for development discontinuation, suspension, product withdrawal, and 'no development reported' if Pharmaprojects reported no change in development in 18 months.

However, the main purpose of this database is to provide critical information on the industry trends and benchmark the competitors' performance, so the project level data are not complete and accurate. To resolve this issue, I also used ClinicalTrial.gov [1], a central repository for publicly accessible information on current and past clinical trial information; and PubMed [2], a free search engine maintained

---

[1] The Federal Drug Administration Modernization Act (FDAMA) in 1997 led to the formation of ClinicalTrials.gov. While initial compliance with trial registration rules was low, registration rates accelerated after the International Committee of Medical Journal Editors (ICMJE) initiated a policy whereby trials must registered as a prerequisite for journal publication [Gill, 2012]). However, trial registries only require information about the stage of the projects, development time lines and other related documents concerning investor relationship and financial files. Only until recently, the Food and Drug Administration Amendments Act (FDAAA) of 2007 start pushing the disclosure of clinical trial results.

[2] PubMed, first released in January 1996, could access the Medical Literature Analysis and Retrieval System Online (MEDLINE). MEDLINE is compiled by the United States National Library of Medicine (NLM), and includes bibliographic information for articles from academic journals covering all biology and medicine field. PubMed automatically links to Medical Subject Headings (MeSH, a comprehensive controlled vocabulary for the purpose of indexing journal articles and books in the life sciences). MeSH is also used by ClinicalTrials.gov registry to classify th studied diseases.

by the National Institutes of Health (NIH) that consists of data primarily form the references and abstracts on life sciences and biomedical topics. These two data sources contained detailed pre-clinical and clinical data, which complemented the Pharmaprojects database.

I obtained data on strategic alliance from the Deloitte Recap database, which tracks alliance formation in the biopharmaceutical industry. The record is at the alliance level, stating the name of the companies forming the alliances and the date of the alliances. It also contains innovation target and therapeutic information of these alliances. I matched these data with the drug R&D data to construct a time series of drug innovation at the project and company level.

With these four data sets, it was possible to compile the drug R&D process and results across time and companies. After removing non-US biotechnology firms, small companies with unreliable data and projects with missing data, the final sample consisted of 1926 project observations from 37 firms in the period from 1987 to 2012. Of these projects, only 3% (62) were approved by the FDA eventually, whereas 82% (1585) were reported as failures. The descriptive statistics for the analysis sample are summarised in Table 3.2.

The success rate of drug R&D projects was relatively low (3%) compared with that reported in previous studies (10%) [Henderson and Cockburn, 1996]; [Kola and Landis, 2004]. This is mainly because many projects are still under development in my data set and the final outcomes are still unclear. One main limitation of the data set is the absence of firms that went bankrupt during 1987 to 2012. Unfortunately, I could not retrieve the innovation data for these firms. However, this study focused on project-level data, so firm survival was not highly relevant to the analysis. In addition, most failed firms were small start-ups, with limited resources and capabilities to complete the innovation process [Arrow, 1972]; [Kola and Landis, 2004]). They had no choice but to focus on the upstream research stage, so cross-stage spillover was not a viable situation for these firms. Therefore, although selection bias is presents in this study, it is not a prominent factor affecting the main results.

Table 3.2: Sample Summary Statistics

Variables	Mean	Median	SD	Minimum	Maximum
Experience	10.99	11.01	1.11	5.99	12.83
Upstream Experience	10.63	10.62	1.07	5.99	12.42
Downstream Experience	9.35	9.93	2.37	0	11.75
R&D Expense	11.19	10.91	1.63	6.51	14.13
Company Age	21.78	19	14.13	0	81
Annual Development	26.79	19	21.91	0	84
Total Assets	12.87	12.19	2.24	6.77	16.79
Industry Experience	9.36	9.36	0.04	9.24	9.41
Innovation Novelty	0.46	0	0.50	0	1
Co-development	0.46	0	0.50	0	1
Primary Ownership	0.88	1	0.32	0	1
Slack Resource	5.10	3.86	4.10	0.73	33.03

Notes: The analysis data set consist of 1926 drug innovation projects from 37 US biotechnology firms between 1987 and 2012. Approximately 3% of all drugs in my data set are approved by FDA eventually, whereas 82% of drug projects have been suspended or discontinued.

Unfortunately, at the present time, no data are available on the cost of the inputs in each stage of the drug R&D process such as capital investment and materials. The following section is going to illustrate how to use R&D efficiency data to proxy productivity. Project-level data were supplemented by firm financial data from Thomson One and registered drug launch data from the FDA's website. Using these supplementary data, I could construct the innovation history variable to control for industry learning and firm size to control for economies of scale.

To fully understand the data set, I conducted further analysis on the failed and launched products in my sample. Figure 3.1 shows the curve for the cumulative number of launched and failed products during the researched period. The number of failures far exceeds the number of successes, which indicates the sky-high attrition rate in drug R&D and the needs for improved productivity. Figure 3.2 illustrates the number of failures in each stage in my data set. The majority of the failures occurred in the pre-clinical stage, which economises on both time and resources. However, the number of failures in the development stage, which is the combination of all clinical trial stages, is still staggering, far greater than the

Figure 3.1: Product Successes and Failures
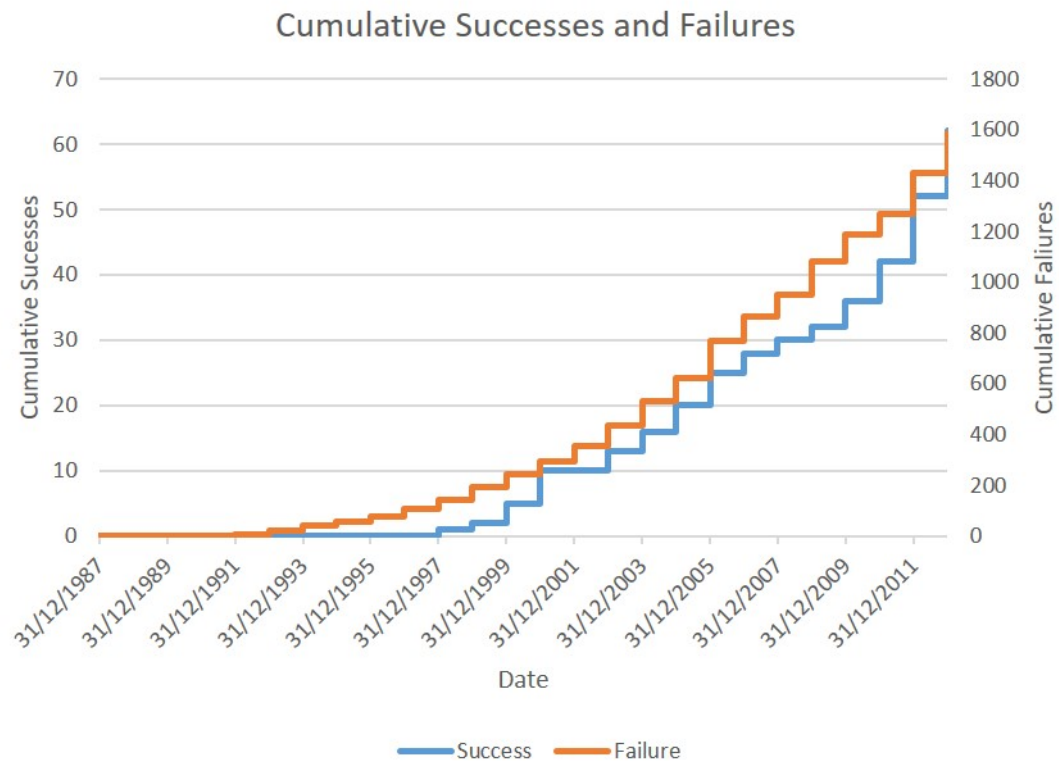
Notes: The figure represents a way of tracking the overall successes and failures in the data set. It graphs the cumulative number of projects launched or failed each year from 1987 to 2012. The left hand y-axis represents the cumulative number of launched projects since 1987, while the right hand y-axis indicates the cumulative number of failed projects.

number of launched products in my sample.

3.5   A Model of Learning by Innovation

Productivity is efficiency in production, i.e. how much outputs is obtained from a given set of inputs [Syverson, 2011]. The productivity of drug innovation can be simply defined as the relationship between the value created by a new medicine and the investments required to generate that medicine. As such, biotechnology firms strive to improve R&D's contribution to firm performance by increasing the volume of innovation and its value, while reducing cost [Paul et al., 2010]. In other words, it

Figure 3.2: Project Failure by Stage
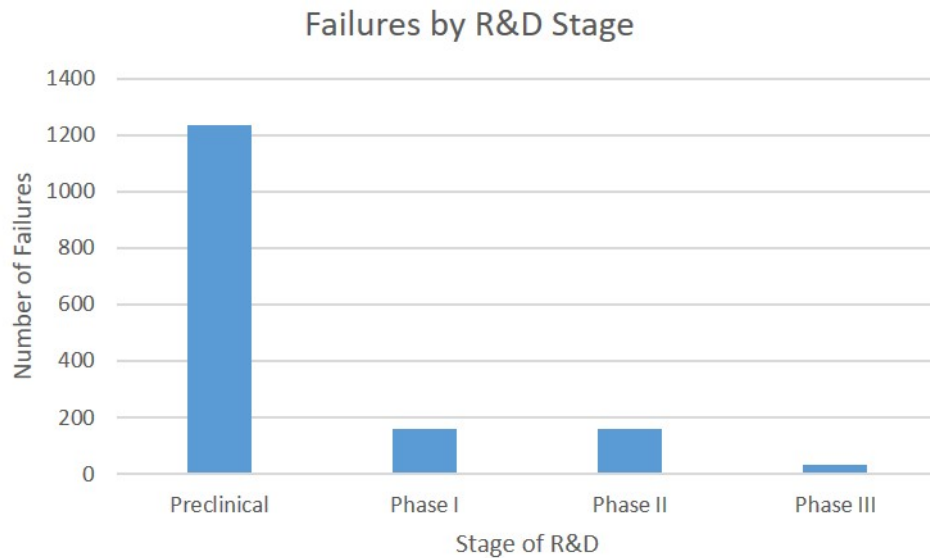
Notes: The bars display the number of failed projects in each stage.

can be elaborated in two important dimensions: innovation efficiency and financial effectiveness. Innovation efficiency represents the ability of a firm to translate inputs (for example, investments, efforts and materials) and deliver an innovative product within a defined period of time, whereas financial effectiveness can be defined as the ability to produce high value at minimum cost [Karlssona et al., 2004]. Specifically, biopharmaceutical firms need to raise R&D performance by increasing the number of new drugs each year, while reducing R&D cost and increasing the commercial and medical value of those new drugs. Thus, drug innovation productivity can be viewed as an aggregate representation of both innovation efficiency and financial effectiveness (See Figure 3.3).

With this definition of drug productivity in mind, I design a productivity relationship or 'drug R&D value equation', which includes the key elements that determine both R&D efficiency and financial effectiveness for any drug innovation project.

$$P = \frac{p(IS)}{CT} * \frac{V}{C}$$

Drug innovation productivity P can be viewed as a function of four elements:

Figure 3.3: Product Innovation Productivity

Product Innovation Productivity	=	Innovation Efficiency	*	Financial Effectiveness
Goal: Maximize		Success Rate	*	Value/Cost

Notes: Product innovation productivity is defined as the results of both innovation efficiency and financial effectiveness. To increase productivity, decision makers strives to increase the success rate and value, while reducing cost. For biotechnology industry, this simply amounts to raising drug development success rate and sales after launch, while keeping R&D cost low.

the probability of innovation success *p(IS)* divided by the cycle time *CT*, and the value of the drug divided by the cost of R&D. Each of these parameters is conceptualised and analysed on a per project basis in this paper. This equation can be also extended to company level by including the number of drug candidates in the R&D progress in each company.

Innovation success rate *p(IS)* is the most important determinant of overall R&D productivity. There is no doubt that increasing the success rate of drug R&D represents the greatest challenge and opportunity for drug innovation, and arguably for sustaining the viability of the entire industry. Since clinical trials account for the majority of the cost, it is clear that increasing innovation success rate is the priority. Unfortunately, several reports suggest that the attrition rate of clinical trials is increasing because more complex drugs are being pursued and because of heightened scrutiny and concerns about drug safety [Kola and Landis, 2004]; [Paul et al., 2010]. Therefore, more basic research is needed to uncover the underlying mechanisms of drug action on the primary target and off-target side effects to increase the innovation rate. In addition, since the majority of drug candidates are destined to fail, it is ideal to fail faster. Failing faster not only helps to reduce the rate of detrimental failures in the late stage of drug R&D, but also saves the cost.

Reducing the cycle time is an important means for improving innovation efficiency. It not only saves time for scientists, but also reduces the cost of drug development. Reduction in R&D cycle time can be achieved in several ways: firstly, developing similar drugs based on prior experience. Learning by doing is the core force for improving productivity, so companies need to reflect on their previous experience and create new knowledge, especially from their previous failures. Secondly, more resources can be dedicated to basic research. Basic research could help developers understand the underlying mechanisms of drug action and connections among different tasks, thereby reducing the number of random trials and errors and cycle time. This cannot be done by biopharmaceutical companies alone but requires the cooperation of the government and research institutions. Thirdly, state-of-the-art technology could also help achieve cycle time reduction, e.g. the use of automated robots to screen compounds and the use of digital humans to test drug samples. Last but not the least, the FDA needs to optimise its working procedure to reduce the time required for regulatory review (submission to launch).

Desirable health outcomes and economic profits determine the financial effectiveness of drug R&D. Obviously, there are two aspects of the value of drug R&D: (i) decreasing mortality and morbidity and reducing hospitalisation to enhance health outcomes and (ii) increasing the price of the medication to gain economic profit. The determinants of overall value are likely to be different depending on the perspective presented above. Therefore, it is crucial to balance these two views and formulate an optimal plan.

Reducing the overall operational expenses is necessary to deliver a successful product to the market. The cost reduction of an R&D project can be realised in several ways. Cutting the cost of R&D processes can provide important opportunities to reduce the overall R&D expense. However, special care needs to be taken to ensure that the production of knowledge is not restrained during this process of cost reduction. Leveraging new technology (such as software tools and laboratory automation) can also lead to cost reduction. In addition, suspending non-promising drug candidates as early as possible could also contribute to cost reduction. Fur-

thermore, restructuring the organisation or research teams helps cut overhead costs (salaries for employees that are not engaged in research and development activities but that are otherwise necessary to support R&D organisations). Such overhead costs are typically prevalent in larger, more mature firms, and it may be difficult to change this owing to the inertia in these organisations.

This paper models the objective of biotechnology firms to maximise the innovation efficiency of drug R&D, but does not consider the aspect of financial effectiveness. Although this approach is necessitated by the fact that the cost of project level data is unavailable, it parallels the way biotechnology firms actually view the drug innovation process. In practice, the decision makers of these firms have very little information about the cost and future revenue of potential drugs, and the survival and sustainability of the firms are mainly determined by the successful launch of new drugs [Kola and Landis, 2004]. In the case of a promising drug candidate, the likelihood of the project being abandoned purely on the basis of projected cash flow and cost of capital is very low. R&D is always the priority in this industry, partially because it is a science-based industry and scientists dominate the management teams. In addition, although the direct costs of drug innovation today have more than double that in the 1980s, the average sales of a recently launched drugs has also increased substantially, with more than 50 achieving the coveted 'blockbuster' status (more than US $ 1 billion revenue). Therefore, the most significant cause of reduced productivity is the movement in survival rates, in particular the attrition rate of R&D. The attrition of projects through the pipeline is therefore the central issue underlying the productivity challenge facing the industry today [Booth and Zemmel, 2004]. Furthermore, it is almost impossible to evaluate the value of a new drug since it does not only generate financial returns, but also creates social value by reducing the suffering of human beings. Last but not the least, it is problematic to allocate costs to each drug. Besides 'molecular cost', which is the direct cost of developing the new drug, there are also other costs. For example, the costs of basic research and overhead costs are hard to assign to a specific drug R&D project. For these reasons, using innovation efficiency to proxy productivity

is acceptable. Previous research also uses R&D efficiency as a performance metric [Henderson and Cockburn, 1996]; [Pammolli et al., 2011].

According to the discussion above, both probability of innovation success and cycle time contribute to the productivity of drug innovation. The average attrition rate of drug R&D is 90%; this is unacceptably high compared to other innovation projects [Kola and Landis, 2004]. Indeed, reducing the attrition rate of the innovation processes in both pre-clinical and clinical stages represents the greatest challenge and opportunity for drug R&D, and arguably for sustaining the viability of the entire industry. Given that the vast majority of drug candidates are destined to fail, failing fast could help redistribute the R&D resources from the later stage to the earlier stage. Shifting attrition from the later stage to the earlier stage and reducing cycle time for failure of projects, can reduce the cost of drug innovation and improve productivity. Reducing the cycle time of each phase of discovery and simultaneous development of other projects could also optimise time and resources.

The productivity of drug R&D can therefore be measured as the efficiency of drug innovation, which firms try to maximise subject to the constraints imposed by technology and firm capabilities. Specifically, I used hazard ratio to estimate innovation efficiency. Hazard ratio not only allowed me to predict whether and when innovation success will occur, but also allowed me to assess the conditional probability of success given that some of the projects in my sample were still under progress [Jenkins, 2005]. In addition, it also allowed me to use a competing risk model to account for heterogeneity in the events of 'success', 'failure' and 'ongoing'. For a given project, $h_t(t|x_{it})$ denotes the hazard ratio of the project's success/failure and is determined by the following equation:

$$h_t(t|x_{it}) = h_{0k}(t) * exp((x_{ik}\beta_k) * \gamma)$$

$h_{0k}(t)$ is the baseline hazard (i.e. the hazard when all covariates are equal to zero), and x_{ik} is a matrix of covariates, including factors such as drug characteristics and firm capabilities. γ denotes company and therapeutic class factors that impact the innovation. I considered both failure hazard ratio and success hazard ratio in this chapter. Assuming the risks are independent, previous research has

shown that the log-likelihood for the competing risks model is additively separable into K terms, each one being a function of the parameters of a single cause-specific hazard [Narendranathan and Stewart, 1991]. Thus, the estimation of a single risk hazard considers durations of completion for other reasons than the one of interest as censored at the point of completion.

Experience effects, denoted by E, are part of x_{ik}. Allowing for the baseline hazard $h_{0k}(t)$ that is independent of experience, the above equation can be transformed into:

$$log(h_t(t|x_{it})) = \beta_1 E_{it} + \theta C_{it} + \gamma$$

θC_{it} denotes a vector of observable variables that plausibly impact the productivity of drug innovation: (1) industry experience for vicarious learning; (2) total assets for firm size; (3) R&D expenditure for knowledge repository; (4) firm age for technology advancement; (5) drug delivery methods for difference in knowledge processing; (6) new chemical entity for innovation novelty; and (7) ownership and alliance formation for potential knowledge flow.

I measured E_{it} as the cumulative time spent by a firm on prior drug R&D in the same therapeutic class until time t . Rather than counting the number of previous drug R&D projects as the experience, my measurement captured the variation in experience because the units of time used to develop a drug vary extensively across products.

3.6 Results

3.6.1 Primary Estimation Results for Learning by Innovating

Table 3.3, column 1, presents the estimated learning model in drug innovation. The estimated coefficient of internal experience is above 1 and statistically significant. This point estimate implies that one year's experience will increase the success hazard ratio by 4.1%. Column 2 shows that internal experience could reduce future failure incidence, which confirm the conclusion in column 1. The estimated coefficients of industry experience in these two models are not statistically

significant. These results are in contrast with those of learning spillover studies [Irwin and Klenow, 1994]; [Thornton and Thompson, 2001], which identify modest cross-firm spillovers in the semiconductor and shipbuilding industries. The lack of knowledge spillover in biotechnology firms may be due to information secrecy: patent protection is rigorous in drug R&D and most companies are unwilling to share information with others owing to the underlining financial benefit even if they are in an existing alliance. In addition, clinical trials are usually strictly confidential processes and the knowledge acquired from drug innovation is not easy to articulate, so inter-organisation knowledge sharing and transfer is much more difficult in the biotechnology field than in other industries. Other studies such as Kellogg (2011) also report lack of spillover in oil and gas drilling companies due to the unwilling to share information.

Table 3.3: Estimation of Learning by Doing in Product Innovation

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Experience	1.03**	0.95***						
	(0.01)	(0.02)						
Industry Experience	0.99	1.00	1.07	0.95	1.09	0.93	1.07	1.00
	(0.15)	(0.17)	(0.14)	(0.19)	(0.20)	(0.18)	(0.37)	(0.26)
Experience with Depreciation			1.07***	0.91***	1.12***	0.87***	1.06**	0.93**
			(0.02)	(0.04)	(0.03)	(0.02)	(0.02)	(0.05)
Depreciation factor			age	age	age^2	age^2	$\sqrt{age}$	$\sqrt{age}$
loglikelihood	-2718.54	-2707.52	-2720.13	-2710.64	-2717.95	-2703.12	-2712.07	-2697.24
Chi-squared	471	492	463	486	475	494	484	498
Dependant Variable	Success	Failure	Success	Failure	Success	Failure	Success	Failure

Notes: Estimates stem from the partial likelihood method of estimation. Dependent variable is the hazard rate of future project success or failure as shown in the table. All models incorporate year fixed effect and firm fixed effect, as well as therapeutic class fixed effect. The depreciation of knowledge is ignored in model 1 and 2 as the basis analysis. In model 3 to 8, knowledge depreciates linearly, rapidly and slowly respectively. Model 1 and 2 are the preferred specification since knowledge depreciation does not affect the analysis. Robust standard errors in parentheses.

$***$p < 0.001; $**$p < 0.01; $*$p < 0.05

Previous works [Argote et al., 1990]; [Benkard, 2000]; [David and Brachet, 2011]; [Kellogg, 2011] focus on organisational forgetting: the decay of experience effects over time. This is mainly caused by members' exit or changes in a firm's structure or process. Since drug innovation techniques are designed by scientists, turnover and lay-offs may lead to loss of experience. I examined organisational forgetting of drug R&D experience by employing commonly used discount factors including 1 (assuming that knowledge is non-depreciating), the age of experience (assuming that knowledge depreciates linearly), the age of experience squared (assuming that knowledge depreciates rapidly), and the square root of the age of experience (assuming that knowledge depreciates slowly).

Column 3-8 examine this forgetting effect. All of them imply that drug R&D experience, no matter the depreciation rate, decreases the failure hazard ratio. More specifically, if knowledge depreciates linearly, one year's experience will decrease the failure hazard ratio by 5.2%; if knowledge depreciates rapidly, one year's experience will reduce the failure hazard ratio by 6.3%; if knowledge depreciates slowly, one year's experience will cutback the failure hazard ratio by 5.0%. Similarly, none of the coefficients of industry experience is significant.

Decision makers always create stage options to alleviate uncertainty during the process of product innovation and each stage has its own information process pattern. In the pre-clinical stage of drug innovation, firms focus on exploring new knowledge and transforming it into prototype products; whereas in the clinical stage of drug innovation, firms tend to test and exploit the knowledge gained through pre-clinical stage. To examine whether experience impacts the learning in different stages, I estimated the effect of experience on upstream success/failure hazard ratio and downstream success/failure hazard ratio separately. This analysis follows the previous procedure but separate success/failure hazard ratio into upstream and downstream as productivity indicators.

In Table 3.4, columns 1 and 2 predict the upstream success and failure hazard ratios, and columns 3 and 4 predict downstream success and failure hazard ratios. In models 1 and 3 for success hazard ratio, the coefficients of R&D experience

Table 3.4: Estimation of Learning by Doing for Each Stage

Variables	(1)	(2)	(3)	(4)
Experience	1.10**	0.92**	1.12***	0.90**
	(0.03)	(0.05)	(0.03)	(0.06)
Industry Experience	1.04	0.96	1.08	0.93
	(0.21)	(0.24)	(0.23)	(0.19)
loglikelihood	-2000.57	-2679.47	-2017.46	-2683.63
Chi-squared	121	484	130	485
Dependant Variable	Upstream Success	Upstream Failure	Downstream Success	Downstream Failure

Notes: Estimates stem from the partial likelihood method of estimation. Dependent variable is the hazard rate of upstream and downstream innovation success or failure as shown in the table. All models incorporate year fixed effect and firm fixed effect, as well as therapeutic class fixed effect. Robust standard errors in parentheses.

$***p < 0.001; **p < 0.01; *p < 0.05$

are above 1 and statistically significant. These results demonstrate that previous innovation experience increases the productivity of both upstream research and downstream development differently. The coefficients of R&D experience of models 2 and 4 show that previous innovation experience decreases both upstream research and downstream development failure rate, thereby increasing productivity at different magnitude. These results illuminate that there may be different learning process in different stages of R&D.

3.6.2 Empirical Analysis of Stage-specific Learning

This section empirically examines whether knowledge gained in one stage can be exploited in another stage, specifically, whether improved productivity with increasing experience is driven by learning from same stage experience, other stage experience, or a combination of the two.

From the above analysis, it is clear that learning in upstream and downstream R&D drive drug innovation experience differently. When a biotechnology firm completes the pre-clinical stage of a new drug, it will rarely continue to conduct clinical trials on its own; instead, the biotechnology firm will form alliances with pharmaceutical firms to maintain the development. Specifically, over 90% of the drugs are co-developed by biotechnology and pharmaceutical companies in the clinical trial stage [Azoulay, 2004]. Since patent rules are strictly enforced and knowledge sharing is anecdotal in drug innovation, why should biotechnology firms forgo the opportunity to increase their internal capability by restricting everything to in-

Table 3.5: Estimation of Stage-specific Learning

Variables	(1)	(2)	(3)	(4)
Upstream Experience	1.08***	0.92**	0.94	0.99
	(0.02)	(0.05)	(0.10)	(0.13)
Downstream Experience	1.05	0.97	1.12**	0.90***
	(0.17)	(0.12)	(0.03)	(0.03)
Industry Experience	1.10	0.87	1.06	0.91
	(0.19)	(0.24)	(0.21)	(0.20)
loglikelihood	-1994.37	-2754.58	-2010.85	-2786.54
Chi-squared	132	480	143	497
Dependant Variable	Upstream Success	Upstream Failure	Downstream Success	Downstream Failure

Notes: Estimates stem from the partial likelihood method of estimation. Dependent variable is the hazard rate of upstream and downstream innovation success or failure as shown in the table. All models incorporate year fixed effect and firm fixed effect, as well as therapeutic class fixed effect. Robust standard errors in parentheses.
$***p < 0.001; **p < 0.01; *p < 0.05$

house operations? Even if they lack the competence to conduct clinical trials in the earlier phase, building this capacity could save transaction costs and benefit further drug R&D, unless learning is bounded and knowledge spillover is not sufficiently large for learning to be a plausible source of long-term growth. I, therefore, tested the knowledge spillover effect by focusing on cross-stage learning.

The analysis still used the model mentioned before, but disaggregated experience into upstream research experience and downstream development experience in addition to separating the success and failure hazard ratio. Table 3.5 illustrates the results. Column 1 uses upstream hazard ratio as estimated outcome and upstream research experience increases the upstream success rate by 5.4%, but the coefficient of downstream development experience is not statistical significant. On the other hand, downstream development experience increases the downstream success hazard ratio by 5.6% although upstream research experience is no longer useful (See column 3, downstream hazard ratio is the outcome). Column 2 and 4 illustrate similar effect (reducing failure to improve productivity) but using upstream and downstream failure ratio as dependent variable.These results show that cross-stage knowledge spillovers are impossible (and sometimes even negative) during the drug innovation process. This explains why most biotechnology firms only focus on the pre-clinical stage of R&D alone and form alliances with other firms in the clinical trial stage.

3.6.3 Mechanisms

This section examines the mechanisms behind the observed stage-specific learning. Understanding the underlying mechanisms is important because these have implications for firms' optimal product innovation practices. If a firm can determine the essential criteria for learning across stages, it could allocate resources depending on whether it wants to be purely research focused firm or well-round innovative company. In addition, the government could use these benchmarks to make policies to encourage or dissuade consolidation among firms.

Resources, both human and physical, provide the basic material for knowledge generation. When resources are strained, decision makers have little time to explore better ways of working and therefore, standard procedures are followed, with little room for experimentation to improve the process [Greve, 2003]. Slack resources are the pool of resources in a firm that are in excess of the minimum necessary to produce a given product. Therefore, firms can use slack resources to engage in additional search and experimentation.

In columns 1 and 2 of Table 3.6, I investigate the impact of 'slack resources' by interacting it with upstream research and downstream development experience separately. I find that slack resources has no impact on cross-stage learning. At first, this finding appears to be surprising since although drug innovation is a sequential process, it is not always a linear procedure. There is a lot of feedback during the process, such as back to animal testing during clinical trial period for another therapeutic target or even to discovering the underlying molecular mechanism. However, if the process of drug R&D contains loops, why does my regression not suggest so? It could be because there are different types of slack resources, such as available slack, recoverable slack, and potential slack, and they have different effect on innovation performance [Greve, 2003]. Due to data restriction, I am not able to separate the slack resources and investigate in detail. As a result, it is unclear whether the results are due to the chosen factor or other possible explanations.

Another factor that contributes to enhanced knowledge spillovers is attention, which diminishes uncertainty. Extra attention could help scientists recognise and

address to weak cues, experiment with novel methods, and develop new knowledge. To test whether attention could help increase cross-stage learning, I use alliance formation as a proxy for extra attention. In the case of product innovation, the decision makers are normally very mindful and attentive [Gulati, 1998]; [Gulati et al., 2000]. The mindfulness could help them identify some insignificant links between upstream and downstream innovation and discover the underlying association for information sharing.

In columns 3 and 4 of Table 3.6, I investigate the effect of extra attention by interacting primary ownership and the experience in different stages. The results show that cross-stage learning happens only in upstream innovation but not downstream innovation. Although this provides additional support to the fact that cross-stage learning is possible but not straightforward, it also suffers the same problem as the previous factor, namely alternative explanation. For example, the alliance partners also possess a large scope of knowledge in the field to disentangle the causal relationship between information and actions.

The overall results presented above help build a circumstantial case in favour of interpreting cross-stage knowledge spillover boundaries and provide the possible policy implications for knowledge spillovers. However, other potential mechanisms, such as the current settings of collaboration between biotechnology firms and pharmaceutical companies are optimal for drug innovation and they are reluctant and it is unnecessary to learning across stages, cannot be rejected, nor do they provide a complete list of factors that alleviate cross-stage learning obstacles.

3.7 Conclusion

This article shows that learning by innovating is an important driver of productivity improvement. I find substantial learning in the productivity of drug innovation in the biotechnology industry. The patterns look qualitatively similar to those seen in other empirical studies documenting learning by doing in various production settings (e.g. [Syverson, 2011]; [Levitt et al., 2013]), albeit this paper focuses

Table 3.6: Empirical Analysis of Cross-stage Learning Mechanism

Variables	(1)	(2)	(3)	(4)
Upstream Experience	0.65	0.88	0.44**	0.96
	(0.33)	(0.16)	(0.15)	(0.12)
Downstream Experience	0.94*	2.81***	1.15	0.97
	(0.86)	(0.03)	(0.84)	(0.04)
Slack Resources	1.13	0.87		
	(0.25)	(0.09)		
Upstream Exp * Slack	0.97	1.02		
	(0.03)	(0.01)		
Downstream Exp * Slack	1.02	1.00		
	(0.03)	(0.01)		
Attention			1.50	0.67
			(2.10)	(0.45)
Upstream Exp * Attention			0.68*	1.05
			(0.25)	(0.08)
Downstream Exp * Attention			0.53**	0.91**
			(0.13)	(0.03)
loglikelihood	-1469.65	-2328.56	-1417.42	-2308.37
Chi-squared	126	340	140	391
Dependent Variable	Up Failure	Down Failure	Up Failure	Down Failure

Notes: Estimates stem from the partial likelihood method of estimation. All models incorporate year fixed effect and firm fixed effect, as well as therapeutic class fixed effect. Dependent variables in model 1 and model 3 are upstream failure rate, while dependent variables in model 2 and model 4 are downstream failure rate. The competing risk of success is not shown in the table due to space restriction. Robust standard errors in parentheses.
$***p < 0.001; **p < 0.01; *p < 0.05$

on product innovation instead of product production. This study combines both productivity drivers, i.e. learning by doing and product innovation, and explains multiple features of the economic phenomenon such as productivity difference and competitive advantage.

In addition, I also explore the specific knowledge that is built through different stages of the learning process. I find that knowledge spillovers are bounded in the same stage. In other words, upstream innovation can only learn from upstream research experience, whereas downstream innovation can only learn from downstream development experience. This work disaggregates the experience and productivity outcomes to addresses the relative effects of stage-specific learning, and illustrates the knowledge spillover boundary for long term economic growth [Thornton and Thompson, 2001].

Finally, I examine the mechanisms behind this stage-specific learning and the factors that can change it. I propose and evaluate - attention - could enhance knowledge spillover across stages, but only for learning in upstream innovation from downstream development experience. The mechanisms underlying stage-specific learning provide additional insights into expanding a firm boundary.

These results can be generalised to other knowledge intensive industries even though the importance of learning by innovating and stage-specific learning varies with industry settings and firm characteristics. For instance, software firms may need to stage a new software R&D and develop stage-specific experience throughout the R&D process. Moreover, the greater the complexity and uncertainty of the new product, the steeper the learning curve.

Beyond theoretical and empirical relevance, my findings also have managerial implications. In economic research, low duration of development and high success rate represent the key drivers of success [Thompson, 2010]. By using competing risk analysis, this paper grasps the main idea of learning-based productivity gains and provides a straightforward tool for managers to achieve high productivity during innovation. In addition, setting the knowledge boundary and understanding the sources of learning by innovating enable decision makers to resolve low efficiency of learning by matching knowledge and experience.

It is also worth mentioning that the government should actively encourage technological progress by following the process of technological externalities (or knowledge spillover). Although the Solow growth model takes technological progress as exogenous and does not explain the determinants of technological progress, many public policies are designed to stimulate private sectors to devote resources to technological innovation. For instance, the patent system gives a temporary monopoly to inventors over their new drugs; the tax code offers tax breaks for pharmaceutical firms engaging in R&D; government agencies, such as National Institute of Health, directly subsidise basic research in universities. As per the results shown in this paper, the government needs to realise that in the absence of knowledge spillover across stages, the social returns to capital accumulation may not exceed the private

returns, and the benefits of increased capital accumulation to facilitate cross-stage investment may not able to yield greater externalities. Therefore, the government should use tax laws to encourage investment in intra-stage technological progress instead of promoting cross-stage technological innovation.

As with all work, this study also has a number of limitations, each of which suggests opportunities for future research. First, stage division as upstream research and downstream development seems arbitrary. This study explicitly assumes that two stages of product innovation represent the R&D dichotomy, but the distinction may not be so obvious in other industries. Therefore, further research in other industries could find their own stage dimensions and test my conclusions.

Another significant critique is the fact that I do not have detailed cost data to calculate productivity. At present, there are no data on the cost of the inputs to each stage of drug development, and the research use aggregated data for clinical trial expense only [DiMasi et al., 1991]; [DiMasi et al., 2003]. However, firms may have data that provide insights beyond the scope of this study. While my analysis is limited to R&D data, firms have financial information about their own projects. Therefore, firms can use the financial data to document productivity gains through cross-stage knowledge spillovers, and evaluate the decisions on firm boundary.

Chapter 4

When the Choice of Governance Structure Overshadows the Competitive Advantage

4.1 Introduction

R&D alliances are an essential part of firm strategy to improve innovation performance in numerous industries (e.g. [Powell et al., 1996]; [Ahuja, 2000]; [Baum et al., 2000]; [Sampson, 2007]; [Zaheer and Bell, 2005]). Alliances function as 'pipelines' through which information and knowledge flow between firms, but while enjoying access to their partners' technological assets, firms in alliances also put their own technological assets at risk of appropriation [Owen-Smith and Powell, 2004]. Empirical evidence has indicated that network structure and interfirm governance create both potential strategic benefits and potential risks that influence a firm's learning and performance [Powell et al., 1996]; [Oxley, 1997]; [Ahuja, 2000]; [Polidoro et al., 2011]. Therefore, research has produced conflicting and confusing findings on how these two factors relate to firm performance. For example, Coleman (1988) argues that a closed or dense network provides social capital in that this structure gives rise to trust, reciprocity norms, and a shared identity, all of which may lead to a high standard of collaborative behaviour. In contrast, Burt (1992) argues that networks in which a focal actor's

partners are disconnected are ideal because they provide social capital in the form of access to timely, diverse information.

Similarly, interfirm governance also presents interesting challenges and opportunities for firms to promote knowledge sharing and protect core technologies from appropriation [Teece, 1986]; [Oxley, 1997]. Two common choices that organisation decision-makers face with regards to the form of governance are partner selection and ownership structure [Li et al., 2008]; [Meuleman et al., 2010] [1]. Partner selection involves whether to explore opportunities to collaborate with new partners or to continue working with old partners in different projects. New partners can bring novel information and additional innovative capabilities, which are essential for innovation [Phelps, 2010]. They also expand the existing boundary of business activities and knowledge domain, improving competitive advantage by enhancing the ability to develop innovation which is different from the existing process [Li et al., 2008]. However, new partners lack the trust, relational norms and routines which are indispensable for successful coordination and co-operation. This can lead to an unpredictable and high variance return of alliances [Reuer and Devarakonda, 2017], and raise the challenge of learning and integration knowledge from partners. In the same vein, firms generally have two alternatives in ownership structure choices: an equity-based joint venture and a non-equity-based market form of contract. A non-equity alliance saves money and time being spent in setting up a joint venture and has the flexibility of forming an alliance with other radically different partners, but at the same time they are vulnerable to potential opportunistic behaviours and uncertain outcomes [Oxley, 1997]; [Sampson, 2004].

These forms of network structure and interfirm governance are not necessarily contradictory, but rather play different roles, which are valuable for different populations or purposes [Burt, 1998]; [Li et al., 2008]; [Phelps, 2010]. In this paper,

---

[1]The scope of alliance is regarded as the third type of interfirm governance decisions to control the threat of knowledge leakage and to protect technological assets in R&D alliances. The scope of activities for a R&D alliance can be restricted to pre-competitive R&D activities only or be extended to include manufacturing and/or marketing. These scope decisions have important implications for the extent to which alliance partners expose valuable know-how to each other. However, in this research, I only focus on research and product development alliances (see the details in 4.3 Data and Sample Section), so the scope of alliance does not play a role here.

I use this contingency approach to explore the interaction between these two factors as an important explanatory variable: whether firms should form their strategic alliances with new partners or old partners, and whether firms should form equity alliances or non-equity alliances, depending on how they are structurally embedded in the network. The multiplexity of governance decision-making within an alliance structure can in turn give an insight into developmental sequences for social relationships. When interfirm governance is explicitly included and alliances are viewed in terms of their relational and structural characteristics, additional sources of information and commitment emerge, which may improve the innovation outcome [Hoang and Yi, 2015].

Innovation search and transaction cost economics are the theoretical foundations for the issue I examine. At its core, innovation is a problem-solving process in which solutions to valuable problems are discovered via search [Dosi, 1988]. Historically, innovation search has focused on various processes that firms can use to develop knowledge that enhances performance. Some argue that innovation search leads to the creation of new knowledge, which typically involves the novel recombination of existing knowledge [Nelson and Winter, 1982]; [Fleming, 2001]; others argue that it is done by reconfiguring the ways in which knowledge elements are aligned [Henderson and Clark, 1990]. Yet whatever the perspective, innovation search is uncertain, costly and guided by prior experience [Dosi, 1988]. Therefore, it needs both knowledge elements and the build-up of absorptive capacity for understanding the novel combination [Gilsing et al., 2008]. Transaction cost economics theory informs this present work, in that both partner selection and ownership structure are transaction features of strategic alliances.

In the empirical work reported here, I studied the innovation performance of organisations which can choose different interfirm governance structures within dense networks: forming an alliance with a new partner or an old one; establishing an equity alliance or a non-equity one. I assessed the innovation and entrepreneurial outcomes based on the number of patent applications that the organisation files and the sum of patent forward citations, which are widely used and capture the R&D perfor-

mance well [Gomes-Casseres et al., 2006];[Phelps, 2010]. Using the NBER Patent Dataset and the SDC Database, along with other sources, I constructed a panel data set of 2,299 alliances from 200 biotechnology firms between 1975 and 2006, and I examined how the number of patents and patent citations change when organisations choose a different interfirm governance structure under a dense alliance network. One empirical challenge is that the interfirm governance decision-making is the outcome of a purposeful matching process to an alliance structure, making it difficult to uncover the causal effects. To overcome this endogeneity problem, I made use of the propensity score weighting and difference-in-differences (DD) techniques to make the treated and untreated groups comparable.

My results show that network closure has a curvilinear effect on innovation performance, and appropriately designed interfirm governance alleviates dense network risks. In particular, I found that when the network closure enhances innovation performance by providing more social capital, partnering with a new partner and engaging in a non-equity alliance, a positive effect and maximum value of innovation are achieved as the network density increases. In addition, forming alliances with new partners and engaging in non-equity form alliances in a closed network reduce the negative effects that occur when network closure lacks the advantages of open structure to access diverse information and information flows. These findings provide evidence which suggests interfirm governance is not independent from alliance structure; rather, it appears to shape the relationship between network structure and innovation performance.

This study contributes to the literature on both alliance and innovation by addressing significant gaps in the research on the influence of alliance network structures and interfirm governance on firm innovation. More specifically, I have bolstered the case for the dynamic process of strategic alliance by focusing on the interaction between network structure and interfirm governance. The results show that network structure and interfirm governance increase innovation, independently and in combination. The results also suggest that neither network structure nor interfirm governance guarantee superior performance. The dyadic-level factor (in-

terfirm governance) and the network-level factor (network structure) coexist in a firm's alliance network and the benefits of network closure and access to diverse information depend on the interaction between these factors. Both new partners and non-equity relationships enhance the positive effect of dense network, and interfirm governance is a portfolio level decision which needs to take the structure of the alliance into consideration.

The rest of the paper proceeds as follows. In the next section, I present the theoretical motivation for my hypothesis. Section 3 describes the construction of the sample and presents descriptive statistics. Section 4 lays out my econometric methodology. Section 5 reports and discusses the results of the analysis. Section 6 discusses the contribution and limitations, and then concludes.

4.2   Theoretical Background

4.2.1 Innovation Search

Innovation search refers to the attempts on the part of some actors to find a solution to a problem. In this way, innovation is characterized as a problem-solving process in which solutions to problems are discovered via search [Dosi, 1988]. The main aim of innovation search is to create knowledge and in general, search is a costly and partially routinised process that actors employ to solve problems and discover opportunities in uncertain and ambiguous environments [Cyert and March, 1963]; [Levinthal and March, 1981]; [Nelson and Winter, 1982]).

Numerous research articles have suggested that the creation of new knowledge is the result of a novel recombination of conceptual and physical materials that were previously in existence [Nelson and Winter, 1982]; [Penrose, 2009]; [Schumpeter, 2010]) or the reconfiguration of the ways in which such knowledge elements are linked [Henderson and Clark, 1990]. Therefore, an increase in search scope increases the number of knowledge elements available for recombination [Fleming, 2001]. In other word, the larger the set of knowledge elements,

the greater are their combinational possibilities [Fleming and Sorenson, 2004]. In addition, the increase in the variance of search materials also increases the expected improvement to current knowledge to be realized from such search [Levinthal and March, 1981].

On the other hand, firms need to ensure that such novel knowledge, once accessed, is evaluated, and when proven to be valuable is adequately absorbed. Innovation search is a costly activity, especially when directed at finding novel solutions [Phelps, 2010]. However, search costs may decline with experience as firms develop more efficient search competences and routines [Levinthal and March, 1981]; [Nelson and Winter, 1982]. In addition, search is often planned and guided by routine and heuristics [Nelson and Winter, 1982]; [Dosi, 1988]. Firms use heuristics in their recombinant efforts, which are often embedded in organisational routines. The institutionalization of past search experiences into organisational routines improves the efficiency and effectiveness of similar, subsequent search efforts [Dosi, 1988]; [March, 1991].

Innovation search emphasizes the importance to firms of having access to diverse domains of knowledge and the build-up of absorptive capacity for understanding novel recombinations. Strategic alliances are both a mechanism and medium for such search. However, alliances per se do not guarantee access or facilitate the search process. In fact, prior research demonstrates that increasing alliance diversity can impede access and inhibit the learning associated with search [Phelps, 2010]; [Phelps et al., 2012].

Both network structure and interfirm governance have potential value to technological innovation. The extent to which these factors in an alliance network relationship can increase a firm's access to its partners' knowledge stock and the capacity to assimilate the knowledge facilitate the firm's search and combination efforts and subsequently its technological innovation.

4.2.2 Network Structure

Strategic alliances facilitates information and knowledge flows between firms

[Owen-Smith and Powell, 2004]. Firms are embedded in varying degrees in networks of relations, and the nature or extent of this structure embeddedness influences firms' access to information as well as their behaviour. Existing theories and research contain opposite claims on network structure. On the one hand, network closure facilitates knowledge transfer by deterring opportunistic behaviour [Coleman, 1988]; [Polidoro et al., 2011]. In addition, dense network also advocates a generalized reciprocity norm and boosts trust, so as to alleviate the concerns about free riders and knowledge leaking [Uzzi, 1997]; [Gulati et al., 2000]. On the other hand, the network closure structure also demonstrates enormous costs, notably lacks novel information and restricts flexibility in a fast-changing environment [Burt, 1992]; [Bogenrieder and Nooteboom, 2004]; [Zaheer and Bell, 2005]. Network density affects the relative novelty of knowledge available in a network and the ease with which a firm can recognize, assimilate and utilize this knowledge [Phelps, 2010].

Increasing network density enhances a firm's relative absorptive capacity. Alliances function as a facilitator of knowledge access, however, knowledge from a direct partner may not be readily understandable by the firm, and if they are not able to understand information from the innovation search, they may need another partner to complement their own absorptive capacity [Gilsing and Nooteboom, 2006]. In other words, redundant ties from dense networks are sometimes needed in order to understand and absorb knowledge acquired in the other relations. This is especially salient for tacit knowledge. A dense network could enhance the absorptive capacity of the firm by acting as a device for screening and interpreting novel information on its potential relevance and value [Gilsing et al., 2008].

A dense alliance network also deters opportunistic behaviour, promotes trust and creates a new reciprocity norm between partners. In a dense network, alliance partners are likely to have common partners, which act as an effective informal social control mechanism over wrongful behaviours. The presence of common partners generates deterrence-based trust by making opportunistic behaviour more visible [Raub and Weesie, 1990]; [Polidoro et al., 2011]. Density and trust build-

ing enable a proliferation of triangulation to judge the reliability of information. The focal firm could develop a richer understanding and a better evaluation of the accrued knowledge since the information from any partner is richer and more realizable to the extent that the partner also profits from triangulation among its partners [Rowley et al., 2000]; [Gilsing et al., 2008].

However, the dense network also embodies tremendous costs, because it lacks the structure holes which exist in a sparse network to access diverse information [Burt, 1992]; [Walker et al., 1997]. Alliance networks rich in structure holes have access to mutually unconnected partners and consequently to much broader and non-redundant information. Thus, maximizing the number of structure holes constructs an efficient and information-rich network. On the other hand, firms with a dense network structure are ill-positioned for learning efficiently and having access to new and novel information [Zaheer and Bell, 2005]. In an R&D alliance, the lack of novel knowledge and distinguished capability is detrimental, demotivating engagement in the innovation process and in the formation and maintenance of the alliance.

A dense network also opens up more avenues for undesirable spillovers, which may increase risk of loss of competitive advantage. In a dense network, knowledge and information reaching the company through its alliance network also reach its partners. From a competence perspective, such knowledge spillovers are desirable, but from a governance perspective spillover to competitors may lead to loss of competitive position. Such diffusion of novel information throughout the network may also deter firms from conducting further innovation search, due to the search costs and appropriation issues [Gilsing and Nooteboom, 2006]; [Devarakonda and Reuer, 2018].

In addition, a dense network functions as an obstacle for firms to create impacting innovation value, as it restricts firms behavioural flexibility and the ability to form alliances with other novel partners. Dense networks often generate strong social norms that largely define the accepted form of behavioural routines and decision choices. Deep embeddedness in dense networks restrains firms from entering

into new, more innovative relationships, since they do not want to be radically different from their existing partners [Gilsing et al., 2008]; [Phelps, 2010].

Given these benefits and costs of a dense network, it has a curvilinear effect on a firm's innovation performance. A dense network is especially beneficial for knowledge transfer during the cooperation process in an R&D alliance. The successful transfer of knowledge and the co-development of new products between alliance parties require the flow of necessary information and the build-up of absorptive capacity to process information [Reuer and Devarakonda, 2017], and partners are more vulnerable and sensitive to each other's opportunistic behaviour, such as shirking, since the partner's input is hard to specify and monitor [Teece, 1986]; [Phelps, 2010]. At low levels of network density, alliance partners lack an established routine and trust, and absorptive capacity for knowledge access and evaluation is still at its infancy. At high levels of network density, although trust is well established and opportunistic behaviour is likely to be deterred, rigid routines and obsolete and familiar knowledge become prominent. Research has shown that as networks become more closed, the chances of adapting to a changing environment and inventing new products declines [Orsenigo et al., 2001]; [Bogenrieder and Nooteboom, 2004]. In contrast, at a moderate level of network density, a firm's innovation performance benefits from a balance of access to novel information and the build-up of absorptive capacity. Therefore, some degree of network density is valuable for innovation, but either too much or too little can be detrimental.

The logic of the above argument can be understand as meaning that novel information decreases with network density, but absorptive capacity increases with it. Both effects are linear since no theoretical and empirical arguments suggest otherwise. In this way, alliance innovation performance is hypothesized to arise from the interaction of novel information and absorptive capacity. The basic idea here is that there is an inverted-U shaped relationship. Mathematically, it can be formulated as follows:

$$AC = a_0 + a_1 * ND (a_0, a_1 > 0) \tag{4.1}$$

and

$$NI = b_0 - b_1 * ND (b_0, b_1 > 0). \tag{4.2}$$

where *AC* is the absorptive capacity, *NI* is the novel information, and *ND* is the network density.

The innovation performance of the alliance is defined as the product of the two linear effects:

$$IP = AC * NI \tag{4.3}$$

Replacing *AC* and *NI* by using equation (1) and equation (2) yields:

$$IP = a_0 * b_0 + (a_0 * b_1 - b_0 * a_1) * ND - a_1 * b_1 * ND^2 \tag{4.4}$$

Equation (4) results in an inverse U-shaped effect if and only if

$$a_0 * b_1 > b_0 * a_1 \tag{4.5}$$

Hypothesis 1. The network density has an inverted U-shaped relationship with the firm's subsequent innovation performance

4.2.3 Partner Selection

Prior research in transaction cost economics suggests that choosing an appropriate interfirm governance is one mechanism that firms use to promote knowledge sharing and protection in alliances [Pisano, 1989]; [Oxley, 1997]; [Sampson, 2004]). Interfirm governance must become align with transaction to reduce appropriability hazards and leakage of valuable intellectual property [Teece, 1986]; [Williamson, 1991]; [Levinthal, 1997]. Transaction cost economics treatment of alliance network has focused exclusively on the two aspects of interfirm governance: partner selection and ownership structure [Gulati et al., 2012].

Both of these are options to reduce transaction cost and uncertainties involved in knowledge sharing and transfer [Li et al., 2008]; [Reuer and Devarakonda, 2017].

Although an alliance can facilitate innovation search and knowledge transfer to create innovation, it does not guarantee the effective combination, retention and assimilation of that knowledge [Teece, 1986]. The tacit and embedded nature of technological knowledge make it difficult for partners to learn. Dense network is a closed network which has no access to broader and non-redundant information [Zaheer and Bell, 2005]. Increasing network density worsens the problem, since there is no open network structure to enable access to novel information. In addition, a dense network also increase the chance of undesirable knowledge spillover since it enhances the likelihood that knowledge and information reach partners which are not supposed to have access to the information [Gilsing and Nooteboom, 2006]. Furthermore, a dense network has strong social norms and behaviour routines, rendering firms unable to form alliances with more innovative partners [Gilsing et al., 2008].

Partner selection is strategically critical because it has the potential to ease knowledge transfer between partners and reduce potential transaction hazards stemming from opportunism [Reuer et al., 2002]. Depending on previous alliance experience, partner selection preference can be categorized into new partner and old partner [Li et al., 2008]. A new partner facilitates the access to novel knowledge, which mitigates redundant information problem, decreases the appropriation of undesirable knowledge spillovers, and increases partners' flexibility to be radically different. These problems become more challenging, and thus more important to resolve, as network density grows.

A new partner reduces novel information problems related to growing network density. It brings novel information and provides a unique combination of distinct capabilities and knowledge assets in order to innovate, therefore innovation arises from new combinations of capabilities [Li et al., 2008]; [Schumpeter, 2010]. Since unique combinations of different assets are required to generate innovation, the addition of similar capabilities by an old partner is no longer useful beyond a critical

minimum level of R&D activities [Sampson, 2007]. In this sense, partners who have no prior experience of cooperation are more likely to access to novel information and learning opportunity , and contribute to innovation.

A new partner not only brings novel information, but can also facilitates searching beyond the organisation's existing boundary of business activities and knowledge domain [Levinthal, 1997]. This enhances a firm's ability to develop timely radical innovation that represents a clear departure from the existing processes [Tushman and Rosenkopf, 1996]; [Li et al., 2008]. A dense network has strong social norms and routines which constrain the cognitive capacity and the ability to look beyond their existing pools of social relationships. A new partner could rescue the deep embeddedness of the dense network in fast-changing, fiercely competitive industries [Wuyts et al., 2005].

A new partner can also reduce knowledge appropriation owing to information asymmetry [Li et al., 2008]. This is especially beneficial for reducing the risk of undesirable knowledge spillover. A new partner alliance means that each partner is less likely to understand the other's know how and learning routine. Therefore, this creates an information asymmetry between the alliance partners such that knowledge leakage is less likely to happen. Even if knowledge and information diffuse unintentionally in a dense network, it is difficult for these partners to appropriate their partner firms' core technologies.

In short, forming an alliance with a new partner improves a firm's ability to alleviate the downside of a dense network and moderates the curvilinear effect of network density on the firm's innovation in several ways by enhancing the potential for novel creation. First, it will increase the magnitude of the positive relationship between network density and innovation. Second, it will increase the maximum value of innovation performance achieved by the firm. In other words, the amplitude of the curvilinear effect of network density is augmented. Third, after the effect of network density on performance turns negative, partnering with a new firm will dampen the negative effect.

The effect of network structure and partner selection occur simultaneously, and

hence there are interaction effect between them. For the combined effects, the full model then becomes:

$$AC = a_0 + a_1 * ND - a_2 * NP \tag{4.6}$$

$$NI = b_0 - b_1 * ND + b_2 * NP \tag{4.7}$$

Multiplying equation (6) and (7) yields:

$$
\begin{aligned}
IP = a_0 * b_0 + (a_1 * b_0 - a_0 * b_1) * ND + (a_0 * b_2 - a_2 * b_0) * NP + \\
(a_1 * b_2 + a_2 * b_1) * ND * NP - a_1 * b_1 * D^2 - a_2 * b_2 * NP^2
\end{aligned}
\tag{4.8}
$$

where *NP* is new partner, and

$$a_1 * b_0 > a_0 * b_1, a_0 * b_2 > a_2 * b_0 \tag{4.9}$$

The interaction effect is positive since the variables have the opposite effect on novel information and on absorptive capacity. An increase in novel information, due to forming an alliance with a new partner, is accompanied by an increase in the ability to absorb it due to an increase in network density.

Hypothesis 2a. Forming alliance with new partners increases the magnitude of the positive effect of dense network when the density is low.

Hypothesis 2b. Forming alliance with new partners reduces the negative effect of the dense network when the density is high.

Hypothesis 2c. Forming alliance with new partners increases the amplitude of the effect of the dense network.

4.2.4 Ownership Structure

Another transaction cost economics treatment of an alliance network has been to focus on ownership structure. Key findings from the literature show that certain characteristics of alliance activities are associated with the adoption of more hier-

archical or protective governance structure - most notably the equity joint venture [Teece, 1986]; [Pisano, 1989]; [Sampson, 2004]. When firms choose an equity-based joint venture, they create a new entity that is dedicated to the alliance activities and is jointly owned and operated by the alliance partners. Transaction cost analysis suggests that a joint venture helps to reduce the hazard of opportunism, especially in technological innovation projects where information is complex and knowledge is tacit. On the other hand, non-equity market contract has no equity investment and only knowledge transfer between partners [Oxley, 1997], [Sampson, 2004]. By adopting a non-equity contract form, firms have the flexibility to form innovative alliances with other partners and the incentive to access diverse information and increased knowledge flow.

The non-equity form mitigates some of the costs and amplifies some of the benefits of increasing network density, thus positively moderating its effect on innovation. Non-equity alliances facilitate the access to novel knowledge, which decreases the frigidness of knowledge protection, enhances the flexibility of the alliances, and decreases the appropriation of undesirable knowledge spillovers.

Non-equity alliances facilitate knowledge access and knowledge sharing, reducing the frigidness of knowledge protection problems related to growing network density. Unlike non-equity alliance, joint venture is a well-understand knowledge protection strategy designed to reduce knowledge leakage concerns and thereby encourage knowledge transfer [Li et al., 2008]. The 'mutual hostage' positions offered by equity joint venture also create behavioural pressure for partners to conform to certain routines and pre-empt them from entering into new, more novel relationships [Teece, 1986]; [Pisano, 1989]; [Kraatz, 1998]; [Gilsing et al., 2008]. On the other hand, a non-equity form with loose alliance structure and less bureaucratic in terms of decision-making, provides new information, search opportunities and novelty creation [Gulati et al., 2012]. The loose structure and novel information benefits of non-equity alliances promote innovation search, which improves the recombination and reconfiguration of knowledge elements [Nelson and Winter, 1982]; [Henderson and Clark, 1990]. Therefore, the non-

equity form reduces organisational frigidness and brings novel knowledge, especially under network closure.

The non-equity form also expands the learning activity boundaries and enhances adaptability. It allows firms to search beyond the formal equity structure to acquire new information. It also accelerates the decision-making process without going through reviews and consensus of joint boards of directors, reduces the sunk costs and commitment, and thereby frees limited capability of joint ventures [Rowley et al., 2000]. In addition, while a joint venture alliance needs a new venture with its own legal structure, a non-equity alliance is a market form of contract, which is easy to dissolve under an unfavourable situation, in order to access other information [Sampson, 2007]. These flexibility and knowledge-flow issues become more challenging, and thus more crucial to resolve, as network density grows.

In addition, the non-equity form can reduce knowledge appropriation owing to lack of hierarchical controls [Gulati and Singh, 1998]. Equity alliance is the likely governance structure for organising alliances to manage uncertainty and reduce appropriation concerns. It can assert control by fiat, provide monitoring, and align incentives. On the other hand, non-equity alliance lacks contractual monitoring and enforcement; therefore, it is less likely to capture profits generated by the innovative activities [Teece, 1986]; [Oxley, 1997]. Therefore, even if knowledge and information disseminate undesirably among partners in a dense network, the structure of non-equity alliance renders their ability to learn effectively from the knowledge spillovers.

To sum up, forming a non-equity alliance reduces the downside of a dense network and moderates the curvilinear effect of network density on a firm's innovation in several ways. Firstly, it will increase the slope of the positive relationship between network density and innovation. Secondly, the amplitude of the curvilinear effect of network density is increased. More specifically, the maximum value of patent applications and patent citations are boosted. Lastly, it will reduce the slope of the negative relationship between network density and innovation.

Like partner selection, ownership structure also needs to be decided simulta-

neously with network density, and hence the interaction model then becomes:

$$AC = a_0 + a_1 * ND - a_2 * NE \tag{4.10}$$

$$NI = b_0 - b_1 * ND + b_2 * NE \tag{4.11}$$

Multiplying equation (10) and (11) yields:

$$IP = a_0 * b_0 + (a_1 * b_0 - a_0 * b_1) * ND + (a_0 * b_2 - a_2 * b_0) * NE +$$
$$(a_1 * b_2 + a_2 * b_1) * ND * NE - a_1 * b_1 * D^2 - a_2 * b_2 * NE^2 \tag{4.12}$$

where *NE* is non-equity alliance, and

$$a_1 * b_0 > a_0 * b_1, a_0 * b_2 > a_2 * b_0 \tag{4.13}$$

The interaction effect is also positive since the variables are complemented by novel information and on absorptive capacity. An increase in novel information, due to the forming of the non-equity alliance, is accompanied by an increase in the ability to absorb it due to an increase in network density.

Hypothesis 3a. Forming a non-equity alliance increases the magnitude of the positive effect of dense network when the density is low.

Hypothesis 3b. Forming a non-equity alliance reduces the negative effect of the dense network when the density is high.

Hypothesis 3c. Forming a non-equity alliance increases the amplitude of the effect of the dense network.

4.3 Data and Sample

The setting for my empirical work is biotech-pharmaceutical alliances. In this industry, fierce competition and fast-changing technology trends motivate firms to seek out various R&D partnerships [Levinthal and March, 1981];

[Powell et al., 1996]. Since the late 1970s, this industry experienced a significant growth in the use of R&D alliances, and by the late 1990s, all industry players were directly or indirectly interconnected to one another [Roijakkers and Hagedoorn, 2006]. The ubiquitousness of R&D alliances in this industry enabled me to closely examine the network structure and interfirm governance at the same time. Additionally, since I use patent data to measure innovation performance, the unique role of patents in drug development is found to be helpful. Given the unusually strong appropriability of knowledge associated with drug development, organisations actively patent their innovations [Fleming, 2001].

4.3.1 Alliance data

To minimize right censoring and undesirable heterogeneity, I limited the study period to 1975-2006. I only collected data from public companies to ensure the availability and reliability of the financial data. I selected only the top 200 firms based on market capitalization because complete and accurate alliance data are more available for industry leaders than for smaller firms [Phelps, 2010]. To minimize the influence of right censoring, I ended the study period in 2006 to allow sufficient time for the approval of patent applications that sample firms made during the period.

To avoid undesirable heterogeneity [Levinthal and March, 1981], I also eliminated alliances in which:

- One of the parties is a university, medical centre, other non-profit organisation, or government agency

- The two parties have a previous alliance covering the same set of technologies, and consequently were primarily renegotiating the terms of an earlier alliance (this inflates the old partner measurement)

- There is neither a research nor a product development component, but the alliance simply involves the marketing of an existing product (marketing alliance is less likely to affect innovation)

- The alliance is focused on non-human therapy or medicine development

The alliance data are downloaded from SDC Platinum. Overall the data set from which this study draws includes 2,299 alliances initiated during 1975-2006. To ensure the data quality, I reviewed every record from the SDC data and correct duplicate entries and other errors and omissions using secondary sources such as annual reports, 10K filings and Lexis-Nexis.

4.3.2 Measuring innovation outcome

I used the number of patents and patent citations to measure innovation performance as patents have been recognized as a rich and potentially fruitful source of data for the study of innovation and entrepreneurship [Griliches, 1990]; [Jaffe et al., 1993]. Even though several limitations of patent data have been mentioned in previous literature (e.g. [Jaffe and Palmer, 2006]), most of them are void in the drug industry and patent references provide one of the most accepted and reliable resources to measure innovation. I obtained the list of all patents assigned to the biotechnology firms from 1975 through 2006 using data from the National Bureau of Economic Research (NBER) Patent Data Project [Hall et al., 2001]. This database links patent data from the U.S. Patent Office to COMPUSTAT.

The raw patent count provides a first approximation of the innovation performance of each firm. However, previous research has shown a high variance in patent value [Griliches, 1990], thus I used the number of forward citations as an alternative measure. Forward citations are a well-established proxy for innovation since they correlate with patent quality, intellectual property value and the market value of the firms [Hall et al., 2001]. To account for the truncation of the citation measurement, I deployed the exponentially decaying factor ($e^{-\frac{2007-Y_t}{C}}$) to discount the older citation count, where Y_t is the year in which the patent was granted and C is a constant of knowledge loss, which is set at 5 years, following Fleming (2001).

4.3.3 Measuring Network Density

To measure network density, I used Burt's indicator of constraint (1992), a measure of triadic closure. This captures the extent of connections that a focal firm forms with its network partners that also have their own level of connectedness. Higher constraint values indicate higher connectedness and density in a focal firm's ego network. Many researchers have used this constraint measure in alliance network studies (e.g. [Shipilov and Li, 2008]; [Phelps, 2010]), as a substitute for ego network density which is defined as the percentage of all possible ties among an egos alters that have been formed.

4.3.4 Descriptive statistics

To minimize alternative explanations and isolate the marginal effects of the explanatory variables, I controlled for several firm-level and alliance-level factors whose influence on innovation might be confused with the explanatory variables. For each biotechnology firm, I gathered data from Compustat on each firm's age, R&D expenditure, and sales. I also recorded each firm's current ratio as the proxy for slack resources from the same source, since previous research has shown that slack resources lead to greater innovative performance [Nohria and Gulati, 1996]. Current ratio is calculated as current assets divides by current liabilities [Singh, 1986].

I calculated alliance experience as the cumulative alliances of the past five years based on the alliance data from SDC Platinum, since alliance experience enhances the collaborative capability of a firm, which in turn facilitates knowledge transfer [Sampson, 2007]. This five-year window is widely accepted as acknowledging knowledge depreciation [Katila and Ahuja, 2002]. Therapeutic areas, R&D stages and alliance formation years were also obtained from SDC Platinum. Firm patent stock was also included, so as to control for the absorptive capacity. The patent stock was measured using the number of patents a firm had obtained in the previous four years. I also used a modified Herfindahl index [Phelps, 2010] as the proxy for technological diversity since diverse firms may be more innovative with

Variables	Mean	Median	SD	Minimum	Maximum
New Partner					
Firm Age	16.85	14.32	15.07	4	36
R&D Expenditure	1.36	1.56	1.03	0.46	6.94
Firm Sales	4.89	5.24	7.21	0	21.68
Current Ratio	2.76	3.03	1.49	0.06	15.27
Alliance Experience	7.82	8.91	6.36	2	28
Therapeutic Area	21.34	18.93	10.24	1	34
R&D Stage Alliance Formation	4.01	3.78	3.75	1	9
Year	1996.21	1997.48	9.28	1980	2006
Patent Stock	13.21	10.28	8.57	2	28
Herfindahl Index	0.39	0.43	0.11	0.15	0.85
Network Density	24.51	17.46	29.67	0	100
Old Partner					
Firm Age	18.97	16.21	13.24	6	41
R&D Expenditure	1.49	1.98	1.46	0.35	7.62
Firm Sales	6.73	6.05	7.36	0	24.57
Current Ratio	2.12	3.24	1.28	0.02	13.29
Alliance Experience	7.46	9.56	7.31	4	34
Therapeutic Area	20.31	18.27	9.87	1	34
R&D Stage Alliance Formation	4.24	3.91	3.65	1	9
Alliance Formation Year	1997.14	1997.1	8.02	1980	2006
Patent Stock	10.53	14.27	10.73	2	31
Herfindahl Index	0.48	0.54	0.12	0.17	0.79
Network Density	25.19	22.76	27.45	0	100

Table 4.1: Descriptive Statistics of Partner Selection

the diverse knowledge flow.

New partner and old partner alliance samples at the baseline. Table 4.1 presents baseline descriptive statistics. Companies that choose new partners as alliance partners are approximately two years younger than companies that choose old ones. Firms with more R&D investment, firm sales and slack resources tend to choose new partners instead of old partners. They also have more alliance experience and patent stock than their counterparts.

In summary, characteristics that determine selection into the new partner alliance are not especially well balanced at the baseline comparing with old partner alliances. However, the region of common support is wide, indicating that it should be possible to create comparable samples on these important dimensions.

Equity and non-equity alliance samples at the baseline. Table 4.2 presents

Variables	Mean	Median	SD	Minimum	Maximum
Equity					
Firm Age	18.23	14.19	16.23	5	38
R&D Expenditure	1.49	1.63	1.06	0.35	6.94
Firm Sales	4.89	5.24	7.21	0	24.57
Current Ratio	3.02	3.46	1.63	0.06	15.27
Alliance Experience	8.69	10.43	6.15	3	34
Therapeutic Area	20.16	19.56	10.54	1	34
R&D Stage Alliance Formation	3.06	3.65	3.97	1	9
Alliance Formation Year	1998.45	1997.34	8.78	1975	2006
Patent Stock	14.62	12.68	8.52	2	31
Herfindahl Index	0.51	0.47	0.13	0.16	0.85
Network Density	26.79	20.67	25.45	0	100
Non-equity					
Firm Age	19.23	17.31	15.32	4	41
R&D Expenditure	1.12	1.48	1.57	0.41	7.62
Firm Sales	4.78	5.04	6.15	0	22.27
Current Ratio	2.15	2.96	1.43	0.02	13.65
Alliance Experience	7.86	8.54	9.48	2	32
Therapeutic Area	22.15	19.03	8.78	1	34
R&D Stage Alliance Formation	4.12	4.34	3.27	1	9
Alliance Formation Year	1997.38	1996.34	7.59	1975	2006
Patent Stock	9.78	15.64	11.57	2	30
Herfindahl Index	0.47	0.41	0.14	0.15	0.81
Network Density	24.99	23.89	25.98	0	100

Table 4.2: Descriptive Statistics of Ownership Structure

baseline descriptive statistics. Companies that choose equity alliance form are approximately one year younger than companies choosing non-equity alliance. Firms with more R&D investment and slack resources also tend to choose the equity alliance. They also have more alliance experience and patent stock than their counterparts. However, the sale for organisations choosing different equity forms are quite similar.

To sum up, although there is a selection bias between equity alliance and non-equity alliance, the overlapping of the important characteristics facilitates the matching to generate analogous samples.

4.4   Econometric Considerations

Governance choices are sometimes driven by expectations about the innovative potential of each alliance, and the innovation performance might occur irrespective of the governance that has been selected. Therefore, traditional econometric techniques, which assume that assignment into different governances is random, cannot determine causal effects. I follow Azoulay and his colleagues' methods (2010, 2011) to estimate the innovation outcomes of different governance alliances.

4.4.1 Propensity-score weighting

To overcome the above-mentioned challenge, I estimate the effects of alliance governance using inverse probability of treatment-weighted estimation [Rosenbaum and Rubin, 1983]; [Hirano and Imbens, 2001]; [Abadie and Imbens, 2016]. Suppose we are interested in the causal impact of a treatment T_i (governance choice) on the outcome Y_i (innovation performance), and that we have a vector of exogenous control variables Xi. Matching methods assume that the treatment is strongly ignorable or conditionally independent of outcomes. This is often written as $Y \perp T \mid X$ with $0 < P(T_i = 1 \mid X_i) < 1$.

The unconfounded assumption depends on other three conditions: first, a rich list of covariates is used to model the probability of treatment; second, units are drawn from similar circumstances; third, outcomes are measured in the same way for both treatment and control groups [Dehejia and Wahba, 2002]. The latter two conditions are trivially satisfied, but the first one, namely the extent to which the analysis accounts for the relevant determinants of governance choice, is arguable. To address this issue, I include most of the variables in the alliance database which divide the alliances into difference governance based on my personal communication with senior scientists from those biotechnology companies.

According to the findings from the interviews with these senior scientists, organisations appear to focus on their scientific resources and innovative capabilities when choosing governance form. Scientific resources include therapeutic areas, stage of the alliance project and firm R&D resources (using current ratio to proxy

firm slack resources). I capture the innovation capability by counting the cumulative patent stock, cumulative alliance experience, and Herfindahl Index of technological diversity. I also include firm age, firm sale, and R&D expenditure to control for the scale in the selection equation. Last but not least, since partner selection and ownership structure are simultaneous decisions, I separate these two elements in propensity score weighting.

Rosenbaum and Rubin (1983) show that if treatment is strongly ignorable conditional on *X*, it is also strongly ignorable after conditioning on a balancing score *b(X)*. The coarsest balancing score (i.e. the *b(X)* with fewest dimensions) is the propensity score *p(X)*, i.e. the conditional probability of being assigned to the treatment group.

Propensity scores are often used to check the common support assumption. They can also be used as sampling weights in an ordinary least squares (OLS) regression. The choice of weights depends on whether we wish to estimate the average treatment effect (ATE), the treatment effect for the treated (ATT) or some other quantity [Abadie and Imbens, 2006]. ATE elucidates the average effect of treatment for an individual picked at random from the population, and ATT measures the average effect for the sub-population that is likely to receive treatment. Suppose we stratify based on estimates of the propensity score, the strata are indexed by j and defined by a set of lower and upper cut-points $(\underline{p_j}, \overline{p_j})$. Let n_{1j} represent the number of treatment (control) observations in each group and $N_j = n_{1j} + n_{0j}$. For each strata, $\alpha_j = \mathbb{E}[Y_1 - Y_0 | \hat{p} \in (\underline{p_j}, \overline{p_j})]$ is an estimate of the local treatment effect. To calculate ATE, we can take a weighted average of the α_j's

$$ATE = \sum_j \frac{N_j \alpha_j}{N} = \frac{1}{N} \sum_j \left\{ \left(\frac{N_j}{n_{1j}}\right) \sum Y_{1j} - \left(\frac{N_j}{n_{0j}}\right) \sum Y_{0j} \right\} = \mathbb{E}\left[\frac{Y_{i1}}{p_j} - \frac{Y_{i0}}{1 - p_j}\right]$$

where $p_j = n_{1j}$ is the probability of treatment (propensity weighting) for group j. As we allow the size of the strata to diminish, this leads to an estimator where each observation is given its own inverse probability of treatment weight.

A similar argument shows that we can estimate

$$ATT = \sum_j \frac{n_{ij}\alpha_j}{N_1}$$

by assigning a weight of 1 to treated observations, and a weight of $\frac{p}{1-p}$ to control observations. Therefore,

$$ATT = \mathbb{E}\left[Y_{i1} + \frac{Y_{i0} * p_j}{1 - p_j}\right]$$

4.4.2 Semiparametric difference in difference

An alternative econometric analysis relies on within-organisation variation to identify the governance's treatment effect. The organisation fixed effects eliminate any influence of unobserved heterogeneity that is constant over time. However, the assumption of difference in differences (DD) estimation is that the average outcomes of the treated and control groups follow parallel paths over time in the absence of treatment.

This assumption is implausible since some drug projects require specific governance structures to reduce risk or improve innovation efficiency. Therefore, some alliance characteristics that are associated with the innovation performance are unbalanced between different governance modes. To address this issue, I use a semiparametric difference in differences (SDD) estimator [Abadie, 2005] that combines adjustment for observed heterogeneity with difference in differences. The idea is to apply propensity-score weighting to the differences in outcomes between post- and pre-treatment periods so that a pseudo-population of old-partner (equity) alliance is created which follows similar dynamics to the new-partner (non-equity) alliance in the pre-treatment period. The SDD estimator calculation is similar to the DD estimator calculation but with the pseudo-population and the inference is obtained using a non-parametric pairwise bootstrap procedure with 500 replications. The ATT of SDD is calculated as follows:

$$SDD = \mathbb{E}\left[\frac{(Y_{i1} - Y_{i0}) * (T_i - P_j)}{\pi * (1 - p_j)}\right]$$

where p_j is still the propensity weighting, T is the indicator of the treatment, and π denotes the unconditional odds of treatment $Prob(T_i = 1)$.

This method is still vulnerable to the critique that time-varying sources of unobserved heterogeneity could produce biased effects. However, the selection concerns will not jeopardise my analysis since it is unlikely that any innovative tendency would be recognised and selected before any governance decision is made. The fact that treated and untreated organisations are well matched at baseline along the dimension of innovation output provides the required support.

4.5 Results

4.5.1 Determinants of Governance Design

Table 4.3 shows the result of logit models of selecting into different governance designs. The regressions include scientific resources, innovation capabilities and scale factors. Among the scientific resource characteristics, only R&D expenditure is significant for both governance selections. I also found that alliance experience, which is an innovation capability factor, has a consistent pattern of significant positive for these two types of governances. However, the scale factors and other scientific resource and innovation capability measurements have no role in the odds of selection.

Hypothesis 1 is the baseline hypothesis which predicts an inverted U-shaped effect of network density on firm innovation performance. Both Table 4.4 and 4.5 provide support for this hypothesis. In these two tables, network density exhibits a positive and significant effect on innovation performance; and the squared term is negative and also significant. Thus, I find evidence of a curvilinear effect.

4.5.2 Effects of Governance Design on Patent Application

With regard to the effect of governance design on patent application, the first

Table 4.3: Determinants of Selection into Governance Design in Alliance Network

Variables	New Partner Selection	Non-Equity Selection
Current Ratio	0.030	0.032
	(0.103)	(0.114)
Alliance Experience	0.009***	0.008***
	(0.002)	(0.002)
Patent Stock	0.011	0.013
	(0.007)	(0.007)
Herfindail Index	−0.113	−0.109
	(0.067)	(0.073)
Firm Age	0.017	0.015
	(0.013)	(0.021)
Firm Sale	−0.003	−0.006
	(0.021)	(0.019)
R&D Expenditure	0.007*	0.006*
	(0.003)	(0.003)
New Partner		0.151
		(0.971)
Non-equity	0.147	
	(0.145)	
Pseudo-R^2	0.343	0.361
Number of Alliances	2299	2299

The dependent variable is the probability of an alliance being formed with a new partner or a non-equity form structure being set up. Estimates correspond to marginal effects from logit specification, with robust standard errors in parentheses. All models incorporate year fixed effect, therapeutic class dummy and R&D Stage indicators.

$***p < 0.001; **p < 0.01; *p < 0.05$

point to be considered is partner selection. Table 4.4 reports the effect of a new partner alliance on the number of patents which firms apply for. There are five models corresponding to the different ways of assessing the partner selection effect. The first model presents the naive regression result, which ignores the selection process. The second and third model include the propensity score weighting to recover the ATE, which emphasizes the average effect of partner selection for a firm picked at random from the population, and ATT, which measures the average effect for the sub-population that is likely to select a new partner in order to form an alliance. The fourth model reports simple DD regression. The last model reports results cor-

Table 4.4: Effects of Partner Selection on Patent Counts

Variables	Native	ATE	ATT	DD	SDD
Density	0.593**	0.412*	0.403*	0.278***	0.342*
	(0.17)	(0.18)	(0.18)	(0.04)	(0.15)
$Density^2$	−0.605*	−0.475*	−0.442*	−0.289**	−0.278*
	(0.28)	(0.24)	(0.21)	(0.10)	(0.11)
New Partner	0.654*	0.517**	0.493**	0.193**	0.208**
	(0.27)	(0.19)	(0.18)	(0.06)	(0.07)
Density * New Partner	−0.943*	−0.877*	−0.584*	−0.598***	−0.477*
	(0.40)	(0.43)	(0.23)	(0.13)	(0.12)
$Density^2$ * New Partner	0.865**	0.843*	0.519*	0.471*	0.518**
	(0.31)	(0.41)	(0.24)	(0.19)	(0.18)

The dependent variable is the number of patent applications. Each coefficient corresponds to the treatment effect of forming an alliance with a new partner in a specification that regresses patent applications on treatment status, firm age and year indicators in all models. The first three models also include therapeutic area, R&D stage, current ratio, Herfindahl Index and firm size measurements (coefficients not reported). Estimates derive from a quasi-maximum likelihood (QML) Poisson estimation, with robust standard errors in parentheses, clustered around firms; bootstrapped standard errors are reported for the semi-parametric difference-in-differences estimates. All models except the naive and the plain difference-in-differences include regression weights computed by using fitted values for the probability of forming an alliance with a new partner, as estimated in Table 3. For the differences between these models see 4.4 Econometrics Consideration section for details.
$***p < 0.001; **p < 0.01; *p < 0.05$

responding to SDD estimates. The SDD model is my preferred one since it adjusts the selection effects while purging the estimates of time-invariant unobserved heterogeneity.

The naive model is always the largest in magnitude, whereas the DD estimate is the smallest. The fact that the DD estimate is systematically lower than the SDD estimates implies that they are on different output trends even before partner selection. This further proves the necessity of using SDD model instead of simple DD one.

Table 4.5 presents the effect of non-equity alliance on the number of patents that firms apply for. It seems that a non-equity alliance has a larger impact on patent application than forming alliance with new partner does. Both Table 4 and Table 5

Table 4.5: Effects of Governance Choice on Patent Counts

Variables	Native	ATE	ATT	DD	SDD
Density	0.612**	0.477**	0.399*	0.301*	0.354***
	(0.21)	(0.18)	(0.19)	(0.13)	(0.19)
Density^2	−0.647***	−0.463*	−0.478*	−0.317*	−0.379**
	(0.14)	(0.22)	(0.20)	(0.14)	(0.11)
Non-Equity	1.113*	0.940**	1.013**	0.204*	0.274*
	(0.53)	(0.31)	(0.34)	(0.08)	(0.10)
Density * Non-equity	−1.347**	−1.275*	−1.047*	−0.893***	−0.814**
	(0.52)	(0.47)	(0.41)	(0.14)	(0.26)
Density^2 * Non-equity	1.004***	1.134***	1.114***	0.817***	0.979***
	(0.17)	(0.21)	(0.31)	(0.20)	(0.27)

The dependent variable is the number of patent application. Each coefficient corresponds to the treatment effect of forming a non-equity alliance in a specification that regresses patent applications on treatment status, firm age and year indicators in all models. The first three models also include therapeutic area, R&D stage, current ratio, Herfindahl Index and firm size measurements (coefficients not reported). Estimates derive from a quasi-maximum likelihood (QML) Poisson estimation, with robust standard errors in parentheses, clustered around firms; bootstrapped standard errors are reported for the semi-parametric difference-in-differences estimates. All models except the naive and the plain difference-in-differences include regression weights computed by using fitted values for the probability of forming a non-equity alliance, estimated in Table 3. For the difference between these models see 4.4 Econometrics Consideration section for details.

$***$p < 0.001; $**$p < 0.01; $*$p < 0.05

show that network density increases innovation performance in the first place, but reduces the patent application afterwards. Partnering with a new partner not only enhances the initial positive effect of a dense network, but also amplifies the amplitude of the effect. Furthermore, it reduces the negative effect of the dense network on the innovation outcome when the density is high. Choosing a non-equity based alliance has a similar effect but different magnitude.

4.5.3 Effects of Governance Design on Patent Citation

The effect of partner selection and governance choice on patent citation is presented in Tables 4.6 and 4.7. As in the previous sections, the results show that patent citation increases with the network density first, but decreases after a certain threshold. Both partnering with a new partner and choosing a non-equity form alliance moderate the inverted U-shaped relationship by improving the positive effect of the

Table 4.6: Effects of Partner Selection on Patent Citation

Variables	Native	ATE	ATT	DD	SDD
Density	2.034***	1.536***	1.457***	0.987***	1.324***
	(0.58)	(0.35)	(0.38)	(0.24)	(0.24)
$Density^2$	−1.978**	−1.324**	−1.438**	−1.132***	−1.003**
	(0.64)	(0.49)	(0.51)	(0.31)	(0.38)
New Partner	1.896**	1.347***	1.515***	1.096**	1.434***
	(0.61)	(0.32)	(0.37)	(0.29)	(0.44)
Density * New Partner	−3.763***	−2.987***	−3.014***	−2.433***	−1.895***
	(0.73)	(0.54)	(0.52)	(0.34)	(0.32)
$Density^2$ * New Partner	3.123***	1.642***	2.014***	1.207***	1.783***
	(0.76)	(0.45)	(0.53)	(0.29)	(0.41)

The dependent variable is the patent citation. Each coefficient corresponds to the treatment effect of forming an alliance with a new partner in a specification that regresses patent applications on treatment status, firm age and year indicators in all models. The first three models also include therapeutic area, R&D stage, current ratio, Herfindahl Index and firm size measurements (coefficients not reported). Estimates derive from a quasi-maximum likelihood (QML) Poisson estimation, with robust standard errors in parentheses, clustered around firms; bootstrapped standard errors are reported for the semi-parametric difference-in-differences estimates. All models except the naive and the plain difference-in-differences include regression weights computed by using fitted values for the probability of forming an alliance with a new partner, estimated in Table 3. For the difference between these models see 4.4 Econometrics Consideration section for details.
$***p < 0.001; **p < 0.01; *p < 0.05$

dense network when network density is low, reducing the negative effect of dense network when network density is high and enhancing the amplitude of the effect.

4.6 Discussion and Conclusion

R&D alliances promote access to their partners' technological assets, but also place their own valuable technological assets at risk of appropriation. Previous research has suggested two factors that affect this appropriation risk: network structure and interfirm governance, but such literature has largely ignored the potential interaction between these two factors. This study is motivated by this important limitation of research conducted on alliance networks and firm innovation. In addition, this research also draws on the seemingly incompatible theoretical arguments by Burt (1992) and Coleman (1988) which have produced conflicting empirical results regarding the influence of network structure. These conflicts stem from an assumption that a firm's access to novel information and the innovation benefits of

Table 4.7: Effects of Governance Choice on Patent Citation

Variables	Native	ATE	ATT	DD	SDD
Density	2.536***	1.312**	1.423***	0.988*	1.253***
	(0.60)	(0.43)	(0.32)	(0.44)	(0.36)
Density^2	−2.203**	−1.521**	−1.637**	−1.013**	−1.145**
	(0.52)	(0.57)	(0.61)	(0.38)	(0.42)
Non-Equity	2.873***	2.076***	2.164***	1.247***	1.538***
	(0.79)	(0.47)	(0.57)	(0.36)	(0.37)
Density * Non-equity	−3.245***	−2.745***	−2.632**	−2.103***	−1.675***
	(0.87)	(0.63)	(0.82)	(0.63)	(0.41)
Density^2 * Non-equity	2.104**	1.648**	1.528**	1.843***	1.947***
	(0.74)	(0.51)	(0.54)	(0.34)	(0.35)

The dependent variable is the patent citation. Each coefficient corresponds to the treatment effect of forming non-equity alliance in a specification that regresses patent application on treatment status, firm age and year indicators in all models. The first three models also include therapeutic area, R&D stage, current ratio, Herfindahl Index and firm size measurements (coefficients not reported). Estimates derive from a quasi-maximum likelihood (QML) Poisson estimation, with robust standard errors in parentheses, clustered around firms; bootstrapped standard errors are reported for the semi-parametric difference-in-differences estimates. All models except the naive and the plain difference-in-differences include regression weights computed by using fitted values for the probability of forming non-equity alliance estimated in Table 3. For the difference between these models see 4.4 Econometrics Consideration section for details.
∗∗∗p < 0.001; ∗∗p < 0.01; ∗p < 0.05

network closure are mutually exclusive.

This study addresses these limitations by examining the influence of the interaction between network structure and interfirm governance of R&D alliances on firms' innovation performance. The study draws on innovation search and transaction cost economics literature and proposes that network structure and interfirm governance play different, yet complementary, roles in innovation. The results support the prediction of the theoretical framework. More specifically, this research has found that a new partner - which requires a trust-building process, but brings novel information - increases innovation performance when the network density is low, reduces the negative effect of network density when it is high, and boosts the maximum value of innovation performance. It has also found that the non-equity form - which facilitates knowledge flows, alleviates the behavioural pressure to conform, and optimizes search and learning cost - has a similar effect on innovation to that of a new partner.

This study has important implications for research and practice. First, it contributes to the debate between Burt's (1992) structure hole and Coleman's (1988) closure form of social capital by suggesting a contingency factor which moderates the relationship between network structure and innovation performance. The prior research assumption has been that a dense network promotes trust and cooperation, whereas firms embedded in a sparsely connected network will enjoy efficiency and brokerage advantages based on the ability to access novel information. Because network closure and structure holes are inversely related, this argument implies that the trust and cooperation benefits of network closure must come at the expense of the benefits of structure holes, and vice versa. Several recent efforts have been made to reconcile these differences. Burt (1998) suggests that these two forms of social capital are not necessarily contradictory, but rather play different roles, which are valuable and depend on the different populations and purposes. This paper utilizes this contingency approach to explore the conditions under which network structure and interfirm governance are positively related to innovation performance. It is suggested that a new partner and a non-equity alliance form could provide diverse and novel information, independent of the network structure. The benefits of network closure and access to diverse information can coexist in a firm's alliance network, and the combination of the two enhances its innovation.

Second, this study contributes to the innovation search literature. Much of this literature stresses the outcome of such search, notably knowledge creation [Dosi, 1988]; [Schumpeter, 2010]. Little research explores how firms conduct innovation search to access their partners' knowledge. This study suggests that innovation search is a two-part task. On the one hand, firms need to develop access to heterogeneous sources of knowledge and in this way create potential forms of novel combination. This emphasis on diversity, is related to Burt's argument (1992), which stresses the benefits of accessing non-redundant contracts to obtain novel information. New partners and non-equity forms also provide new knowledge when network closure lacks information access. On the other hand, firms need to ensure that such novel knowledge, once accessed, is evaluated, for absorption and recombi-

nation. This process is more in line with Coleman's view (1988), which highlights the benefits of redundant network structure. This emphasize show network structure could rescue the lack of absorptive capacity of a new partner and non-equity form. So an important contribution of this paper is that it illustrates two essential components of innovation search - novel information and the build-up of absorptive capacity - and how these two are played out in the interaction of network structure and interfirm governance.

Third, this study contributes to the multi-level network research scheme. Although research into social capital and social network analysis has been conducted at multiple levels of analysis, researchers have largely limited their studies to a single level of analysis and have failed to fully recognize that different levels of structure may interact with each other [Gilsing et al., 2008]. This study considers both the role of global network density and dyadic interfirm governance choices. The focus on both levels illuminates the under-socialized view of alliances and illustrates how far the dyadic level of interfirm governance choices could be rescued or amplified by the entire network structure [Gulati et al., 2012]; [Powell et al., 2005]. Thus in this research it has been found that the network structure, in terms of its density, indeed plays an important role and conditions the potential benefits and cost of interfirm governance.

Fourth, the research results show that complementary relationships exist between network structure and interfirm governance. This study suggests that partner selection and ownership structure have opposite effects on information novelty and on absorptive capacity compared to network structure. Here the interaction effects complement each other. An increase in novelty, due to a new partner or non-equity form, is complemented by an increase in the ability to absorb the novel knowledge due to an increase in network density. Thus this work demonstrates a dynamic and endogenous system of alliance network.

In terms of implications for managerial practice, an analysis in the present study suggests that partner selection and ownership structure are not merely dyadic-level alliance decisions, but rather portfolio-level decisions. Decision-makers

should devise their partner selection strategy and ownership structure choices in conjunct consideration of the network structure they are embedded in, which is determined at the alliance portfolio level. Unless decision-makers look beyond their direct ties and recognize the potential performance implications of such contingencies, they might end up achieving an inferior innovation performance. In other words, decision-makers have to be able to consider the whole network in order to optimize their alliance strategy and in order to achieve better innovation performance.

Despite the significant theoretical and practical contributions of this study, however, there are several limitations. First, no mechanism story could be provided due to the data limitations and the complex system. Rather, these effects appear to be driven, at least in part, by the balance between benefits and costs of both governance and structure. When organisation decision-makers facing the choice between a new partner or an old one, an equity form or a non-equity one, they also need to take their network structure into consideration, because these two decisions are interdependent on each other to determine innovation and entrepreneurship performance. The increasing mean and decreasing variance of choosing these two types of governance also suggest that there is always a trade-off between reward and uncertainty.

In addition, this paper raises many questions and needs further work on the decision-making process of alliance formation. It would be interesting to conduct a qualitative study to see how organisational decision-makers decide on the interfirm governance. The decision-making process not only provides an explanation for the results of the present study, but also provides a bench-mark for other research in decision science. Future work could also usefully match performance data to each alliance. This could present evidence at the alliance level instead of the organisational one.

Furthermore, because patents were used in this study as a proxy for innovation performance, the measurement may not capture all of the innovation in firms. Therefore, other measurements of innovation could be used to supplement the re-

sults. Although patent data is regarded as a reliable measure that can present a fairly broad range of technical innovation, the data cannot capture process innovation and commercialisation performance.

Last but not least, this research only considered two party alliances, while it ignored multilateral R&D alliances. A multilateral alliance is a single cooperative arrangement involving three or more partners [Lavie et al., 2007]. Increasing the number of alliance partners introduces more complexity to knowledge exchange, so trust-based governance is more suitable for overcoming the heightened challenge associated with knowledge exchange [Li et al., 2011]. In addition, uncooperative behaviour is more likely to be anonymous in a multilateral alliance, and the monitoring and penalty of this kind of behaviour is more salient. So the results from bilateral alliances may not be applicable to multilateral alliances, and so the topic deserves separate research and investigation.

In sum, the research presented here suggests that firms can use both network structure and interfirm governance to share knowledge and protect knowledge leakage. The complementary nature of these two factors is beneficial to innovation performance. The results suggest that this line of inquiry has potentially important implications for the theory and management of interfirm alliance.

Chapter 5

Conclusion

A large amount of scientific information internal and external to the firms has created challenges for knowledge management, resulting in a loss of innovation and decrease in productivity if the information cannot be employed effectively for R&D. Firms that manage knowledge more effectively may be more successful in bringing new products to the market, and could be in a prominent position in the marketplace by recouping development expenses at an increased pace and subsequently achieving greater profit. Ineffective knowledge management may cause decisions to not be made, to be delayed, to be made in error, or to be made without thorough analysis, which may further lead to budget overruns and time delays for new product launches. The intent of this dissertation is to identify knowledge management factors that contribute to technological innovation and business value.

As discussed in the introduction, the three studies in this dissertation provide different perspectives on how firms manage knowledge to enhance technological innovation and improve competitive performance. Understanding exactly how experience and network structure affect both learning behaviour and innovation productivity is important to our knowledge of organisational learning, knowledge spillover and alliance networks. In the context of drug innovation, these three studies address how firms design innovation strategies to enhance learning and capability; how they manage the new product development (NPD) process and knowledge spillovers and define firm boundaries, and how they organise network structures and interfirm governance to protect technological assets.

The value of any technological innovation is only partly determined by what the technology can do. A large part of the value of an innovation is determined by the degree to which decision makers can manage and transform the knowledge. Crafting an innovation strategy is not just a way for firms to realise innovation, but it is a core part of the knowledge management and innovation process itself. Chapter 2 attempts to illustrate that innovation strategy can enhance learning from failure and improve capability. The findings show that two innovation strategy choices - novelty of innovation and primary ownership - not only improve learning from failure experience but also enable prior experience of failure to better reduce the incidence of future failure than prior experience of success.

Another important consideration regarding knowledge management in technological innovation relates to the combination of product innovation and learning by doing to promote the productivity of innovation. Both factors are the principle sources of productivity growth. Knowledge spillover is a source of increasing returns in R&D and enhances technological development [Griliches, 1991]. Chapter 3 demonstrates that learning by innovating is a driving force of productivity and explores the boundary conditions of learning by innovating. Most innovation firms deploy stage-development processes instead of parallel-development processes, especially with the increasing uncertainty and risks involved in the innovation process. Stage-development process can reduce commitment and shift the innovation direction with minimum cost. The findings show that drug research cannot fully exploit the productivity gains acquired through experience with drug development. In other words, learning is stage-specific owing to the presence of limited attention.

The structure and governance of innovation networks can significantly influence the capability of firms' innovation, the effectiveness of that innovation, and the speed of new knowledge production. A vast majority of firms use some types of network structure to organise their NPD process. Chapter 4 examines the network effect, especially dealing with the interaction between network structure and interfirm governance. The findings show that interfirm governance contingently determines the outcome of patent application and patent citation under dense net-

works.

Together, these three studies analyse the determinants of innovation behaviour and competitive performance. They show that both experience and structure play an important role in developing capabilities and improving innovation productivity.

5.1 Research Setting and Generalisability

Even though the three studies are answering different questions, they focus on the same industry setting - the biotechnology industry. While the general research setting is the same, there are some differences in each setting to test specific hypotheses addressed in each study. Table 5.1 summarises the setting, dependent variables, independent variables and key descriptive information on the research settings of these three studies.

	Chapter 2	Chapter 3	Chapter 4
Setting	Drug R&D	Drug R&D	Drug Innovation Network
Dependent Variables	Failure Hazard Ratio Success Hazard Ratio	Failure Hazard Ratio Success Hazard Ratio	Patent Application Patent Citation
Independent Variables	Failure Experience Success Experience Innovation Strategy	Research Experience Development Experience	Network Density Partner Selection Equity Form
Level of Analysis	Project	Project	Network
NPD Stage	R&D	R&D	Pre-product

Table 5.1: Key Characteristics of Settings

Table 5.1 presents some interesting similarities and differences among the three studies. The primary similarity is the complicated nature of the drug innovation process. All three chapters focus on the drug innovation process; chapter 4 measures the pre-product performance - patent, whereas chapters 2 and 3 consider the whole drug research and development process. Another important similarity is that all three studies focus on non-market based outcomes (success or failure hazard ratio, patent application, and citation). For non-market outcomes, there is no requirement for financial data, and the determinants of outcomes are more salient to technologi-

cal innovation. This helps limit the influence of external factors, such as economic conditions and industry cycles. The reduced importance of these outcomes of external factors makes the empirical contexts easily comparable with other types of technological innovation.

There are also a number of differences among the research settings. First, these three studies focus on different stages of technological innovation. As mentioned above, chapter 4 tests hypotheses in the pre-product stage, whereas chapters 2 and 3 focus on the entire technological innovation process. A second difference is the analytical level. Chapters 2 and 3 focus on project level, but chapter 4 addresses the network level.

These two similarities and two differences provide some insights into the level of uniqueness in each setting. Another issue with the research context is that it is possible to test the hypotheses in other technological innovation contexts, such as chemical or software development. In many industries, the ability to develop new products quickly, effectively, and efficiently is now the single most important factor driving success. In industries such as computer hardware and software, telecommunications, auto mobiles, and consumer electronics, the failure rate for new product development is very high. Per many estimates, more than 90% of all new product developments fail to result in an economic return [Berggren and Nacher, 2001]; with drug R&D, many projects are never completed. Therefore, making new product development more effective and efficient in other industry settings is also salient and deserve attention.

In addition, for NPD process to be successful, technological innovation needs to minimise the cycle time. Shortening innovation cycle time can help a firm build brand loyalty, pre-emptively capture scarce assets, and reduce costs. Additionally, a company that is able to complete innovation quickly has more time to develop complementary goods that enhance the value and attractiveness of the main product [Schilling, 1998]. Therefore, reducing innovation cycle time is a common goal for all technological innovation.

Furthermore, to avoid escalating commitments that lead managers to support

and push bad projects forward, many innovators use stage decision in the NPD process. At each stage of the process, they are required to gather vital technological, market, and financial information to use in the decision making process to forward, recycle, hold or abandon a project. According to studies by the Product Development and Management Association, nearly 60% of the firms use the stage decision process to manage their new product development process.

Last but not the least, like biotechnology firms, all firms frequently face difficult decisions about the scope of activities to perform in-house, and whether to perform them alone as a solo venture or to perform them collaboratively with one or more partners. A significant portion of innovations can be attributed not to any single individual or organisation, but to the collaborative efforts of multiple parties. Collaboration can often enable firms to achieve more at a faster rate and with low costs or risk.

The findings of the dissertation can be generalised based on the reasons discussed above. Of course, these statements are merely general claims considering the mechanism and details of technological innovation across industries; additional work would need to be done to test the hypotheses in this dissertation under other industry contexts.

5.2 Contributions

The overall purpose of innovation management theory and research is to explain how and why there are differences in innovation performance. Some scholars, thus, have a collective interest in understanding what role knowledge plays in the innovation process. To determine the extent to which knowledge management contributes to technological innovation, this dissertation contributes in the following ways:

First, this dissertation illustrates that two factors, experience and alliance structure, explain the heterogeneity in firms' behaviour and competitive performance. Chapter 2 empirically confirms that failure experience and success experience have differential impact on innovation performance, and organisations learn to improve

their performance more significantly through experience with failure than through experience with success. Chapter 3 demonstrates the boundary condition of experience in various stages and the scope of learning from these experiences. Chapter 4 explains the dynamic interaction of network structure and interfirm governance, and shows that their interaction increases innovation.

Second, this dissertation complements transaction cost economics in three ways. Coase suggests that firms exist because they are more effective than markets in applying the price mechanism to negotiate contracts [Coase, 1937]. Williamson explains that the control opportunism associated with specificity of assets, and costs associated with the negotiation of contracts, determine the firm form or the market form [Williamson, 1999]. Despite this progress, TCE cannot explain heterogeneous firm performance [Kaplan et al., 2001]. This dissertation demonstrates that knowledge activities, which are partially explained by TCE [Grant, 1996], contribute to innovation heterogeneity. In addition, Chapter 3 discusses how the KBV can be applied to firm boundary decisions and efficient organisation approaches besides TCE's claim that boundaries depend on the bureaucratic costs associated with the decision to take a transaction out of the market or organise it internally [Williamson, 1999]; [Azoulay, 2004]. Furthermore, although TCE suggests that a joint venture helps to reduce the hazard of opportunism, especially in technological innovation projects where knowledge is complex, Chapter 4 proves that non-equity venture increases innovation performance when firms are already embedded in dense network.

Third, this dissertation shows the relationship between experience and resources. Recent research in resource-based view has attempted to demonstrate that the years of experience that firms have is an indicator of the presence or absence of capabilities. Several research studies suggest that complementary assets are useful in the commercialisation of a new innovation [Teece, 1986], and product market experience can build these complementary assets [Eggers, 2012]. Researchers also use years of experience to track the accumulation of complementary assets and measure product quality [Levin, 2000]; [Nerkar and Roberts, 2004]. The above re-

search linking experience and the development of resources (complementary assets and capabilities) generally suggests that years of experience is the best available measure of resource development. This implies that a year of experience for one firm is identical for any two different firms in the market. One of the findings of this dissertation is that there is heterogeneity in the experiences of different firms, and that those experiences lead unevenly to the development of knowledge. In addition, the strategic choices firms make also dictate how experience translates into resources.

5.3 Limitation

While I use Coarsened Exact Matching and Propensity Score Weighting to overcome the endogeneity of strategic choices in Chapter 2 and governance choices in Chapter 4, and to uncover causal effects, these methods suffer some drawbacks. Both matching methods rest on the strong assumption that assignment to the treatment condition can be ignored after conditioning on X. Unfortunately, this assumption cannot be tested. But I am able to examine the balance of exogenous pre-treatment covariates across treated and control observations to make sure that they are not imbalanced. Although we can view matching methods as a generalisation of 'kitchen sink' regression that is somewhat more likely to yield unbiased estimates of causal parameters, they lack an explicit conceptual model of the process that assigns observations to the treatment and control condition.

Even though I make every endeavour to understand the logic behind the findings of each chapter, no complete mechanism story could be provided due to the data limitation and the complex system. The underlying processes of strategy making or interfirm governance decision are not fully explored in the dissertation. When organisation decision-makers face these choices, they also need to take the type of experience or network structure into consideration, which increases the complexity of discovering a mechanism story. Future research should use qualitative methods to uncover the process story and complement the results here.

Last but not least, although I use product launch as the proxy for innovation

performance, I do not have sales data of these products to illustrate the underlying economic significance of corporate incentive efforts. Admittedly, using product innovation output is complementary to the literature that examines more intermediate indicators of innovation outcomes, such as R&D intensity or patent applications [Stuart and Podolny, 1996], it is far from the perfect 'market-relevant' measure of innovation output; many new products are not valued by the stock market. The use of products as a measure of innovation performance is further undermined by the observation that many innovations cannot influence firm performance until the product has been introduced and tested in the market for a while [Kogan et al., 2017].

5.4 Future Opportunities for Research

The above discussions - evaluating how the findings of my current work would play out in other empirical settings, and the limitations of this dissertation - provide some further clarity on how I could leverage my current research work to initiate important related projects. One stream relates to the reconfiguring of findings with settings; another relates to the methods, and the third relates to the link between experience and decision making.

As noted above, there are some potentially interesting opportunities to exploit the findings discussed in this dissertation in different empirical settings. One of the ideas noted above would be to repeat the three chapters in a very different setting from the biotechnology industry, for example, in software development. Software development also features multiple new product development projects for most firms over a number of years. Both failure and success experience exist during the product development process even though the failure rate is much lower than in the biotechnology industry, research and development are clearly separated although the R&D process takes less time, and strategic alliances prevail in the R&D process. Support for the finding in the dissertation would further back the relationship between knowledge management and technological innovation discussed here.

The second potential path for future work would be to further articulate the knowledge and innovation link by exploring the knowledge management process.

First, in future research I would like to look at how organisational experience is stored and accessed later to create useful capabilities. Qualitative work looking at the routines that firms and managers use to decide which existing knowledge and memories to use for new projects within the firm may help illuminate this process. Second, it may be instructive to look at how managers choose innovation strategies or interfirm governance to match the experience or network structure. Archival data on the managerial attention and decision-making process could help to illustrate the process.

The third path involves the experience as the 'context' for decision-making. An experiment on exactly how experience, especially failure experience, plays a role in subsequent decision-making. In the experiment, I could look at how different types of failures or different stages of experience shape the relevant context for learning and firm boundary. A simulation study on decision-making processes could also shed light on how previous experience may lead to a firm to optimise on decision-making processes, thus limiting the firm's ability to retrieve knowledge.

5.5 Conclusion

Overall, this dissertation has investigated various aspects of the links between knowledge management and technological innovation in the context of new drug development. Understanding exactly how knowledge management processes affect both capabilities and innovation performance is crucial to our knowledge of organisational learning, knowledge boundary and knowledge appropriation. By placing the studies in the context of new drug development in the biotechnology industry, these three studies address more complicated knowledge creation and sharing than product manufacturing, which rely in large part on similar past experience. In order to increase innovation performance, decision makers must identify innovation strategies, understand the difference between failure and success experience, figure out the knowledge spillovers boundary in R&D projects, and implement a proper alliance network structure to facilitate knowledge flow. These three studies each address one or more aspects of these factors.

Chapter 2 deals with the link between different experience (failure and success experience) and innovation strategies, and the link between experience and innovation outcomes (failure leading to superior product-level outcomes). Chapter 3 deals with the link between different stages of experience and innovation outcomes (only same-stage experience leading to superior innovation performance). Chapter 4 deals with the links between network structure and interfirm governance, and the link between their interaction and innovation performance. Together, these three studies analyse the complicated relationship between knowledge management and innovation outcomes, and lead to three distinct conclusions.

First, success and failure experience contribute differently to innovation outcomes. Although innovation failures are devastating, organisations are able to learn more from the failure experience. In addition, two types of innovation strategy choices - innovation novelty and primary ownership - can enhance learning from failure and improve learning from failure more than learning from success.

Second, while many prior researches have demonstrated that learning by innovation is a precondition for innovation success, this dissertation points to the effect that learning by innovating has on productivity. Similar to findings in other production settings [Syverson, 2011], this dissertation demonstrates that learning by innovating is a key driver of innovation productivity. In addition, this dissertation also discusses the boundary condition of learning by innovating, and finds that knowledge spillovers are generally bounded in the same stage, that is, upstream innovation can only learn from upstream research experience, whereas downstream innovation can only facilitate learning from downstream development experience.

Third, this dissertation also expands from firm level to alliance network level, and suggests that firms use both network structure and interfirm governance to share and protect knowledge. New partner alliance or non-equity alliance complements dense network in novel information access and the build-up of absorptive capacity, therefore, enhances innovation performance.

Appendix A

Missing Values

Three assumptions are illustrated here to calculate missing data: minimum, maximum and middle point.

Table A.1: Missing Value Calculation

Assumption	Missing Data	Additional Conditions	Remedy
Middle Point	Dc	Mc ≠ Mp or Mc ≠ Mn	Dc = 15
	Dc	Yc = Yp and Mc= Mp	Dc = (Dp + Dldm)/2
	Dc	Yc = Yn and Mc= Mn	Dc = Dn/2
	Mc, Dc	Yc ≠ Yp or Yc ≠ Yn	Mc = 6, Dc = 30
	Mc, Dc	Yc = Yp	McDc = (MpDp + Dldy)/2
	Mc, Dc	Yc = Yn	McDc = MnDn/2
	Yc, Mc, Dc		YcMcDc = (YpMpDp - YnMnDn) * ratio
Minimum	Dc	Mc ≠ Mp or Mc ≠ Mn	Dc = 1
	Dc	Yc = Yp and Mc= Mp	Dc = Dp + 1
	Dc	Yc = Yn and Mc= Mn	Dc = 1
	Mc, Dc	Yc ≠ Yp or Yc ≠ Yn	Mc = 1, Dc = 1
	Mc, Dc	Yc = Yp	Mc = Mp, Dc = Dp + 1
	Mc, Dc	Yc = Yn	Mc = 1, Dc = 1
	Yc, Mc, Dc		YcMcDc = (YpMpDp - YnMnDn) * ratiomin
Maxmium	Dc	Mc ≠ Mp or Mc ≠ Mn	Dc = 30
	Dc	Yc = Yp and Mc= Mp	Dc = Dp - 1
	Dc	Yc = Yn and Mc= Mn	Dc = 30
	Mc, Dc	Yc ≠ Yp or Yc ≠ Yn	Mc = 12, Dc = 31
	Mc, Dc	Yc = Yp	Mc = 12, Dc = 31
	Mc, Dc	Yc = Yn	Mc = Mn, Dc = Dn - 1
	Yc, Mc, Dc		YcMcDc = (YpMpDp - YnMnDn) * ratiomax

Notations:

- Current stage year: Yc; Current stage month: Mc; Current stage day: Dc;

- Previous stage year: Yp; Previous stage month: Pc; Previous stage day: Pc;

- Next stage year: Yn; Next stage month: Mn; Next stage day: Dn;

- Last day of the year: Dldy; Last day of the month: Dldm

Stage duration ratio: ratio (this ratio is calculated based on existing data by using previous stage duration divided by the previous stage duration plus next stage duration)

Appendix B

List of Abbreviations

Table B.1: List of Abbreviations

ATE	average treatment effect
ATT	treatment effect for the treated
BLA	biological license application
CEM	Coarsened Exact Matching
DD	difference-in-differences
FDA	Federal Drug Administration
FDAAA	Food and Drug Administration Amendments Act
FDAMA	Federal Drug Administration Modernization Act
GDP	gross domestic product
ICMJE	International Committee of Medical Journal Editors
IND	investigational new drug
KBV	knowledge-based view
MEDLINE	Medical Literature Analysis and Retrieval System Online
MeSH	Medical Subject Headings
NBER	National Bureau of Economic Research
NCE	new chemical entity
NDA	New Drug Application
NIH	National Institute of Health
NLM	National Library of Medicine
NPD	new product development
OLS	ordinary least squares
QML	quasi-maximum likelihood
R&D	research and development
SDD	semiparametric difference-in-difference
TCE	transaction cost economics
TFP	total factor productivity

Bibliography

[Abadie, 2005] Abadie, A. (2005). Semiparametric Difference-in-Differences Estimators. *The Review of Economic Studies*, 72(1):1–19.

[Abadie and Imbens, 2006] Abadie, A. and Imbens, G. W. (2006). Large Sample Properties of Matching Estimators for Average Treatment Effects. *Econometrica*, 74(1):235–267.

[Abadie and Imbens, 2016] Abadie, A. and Imbens, G. W. (2016). Matching on the Estimated Propensity Score. *Econometrica*, 84(2):781–807.

[Acemoglu and Linn, 2004] Acemoglu, D. and Linn, J. (2004). Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry. *The Quarterly Journal of Economics*, 119(3):1049–1090.

[Adams, 1990] Adams, J. D. (1990). Fundamental Stocks of Knowledge and Productivity Growth. *Journal of Political Economy*, 98(4):673–702.

[Aghion and Howitt, 1998] Aghion, P. and Howitt, P. (1998). *Endogenous Growth Theory*. The MIT Press.

[Aghion and Jaravel, 2015] Aghion, P. and Jaravel, X. (2015). Knowledge Spillovers, Innovation and Growth. *The Economic Journal*, 125(583):53–573.

[Ahuja, 2000] Ahuja, G. (2000). The Duality of Collaboration: Inducements and Opportunities in the Formation of Interfirm Linkages . *Strategic Management Journal*, 21(3):317–343.

[Ahuja and Lampert, 2001] Ahuja, G. and Lampert, C. M. (2001). Entrepreneurship in the Large Corporation: a Longitudinal Study of How Established Firms Create Breakthrough Inventions. *Strategic Management Journal*, 22(6-7):521–543.

[Alavi and Leidner, 2001] Alavi, M. and Leidner, D. E. (2001). Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly*, 25(1):107–136.

[Alegre and Chiva, 2013] Alegre, J. and Chiva, R. (2013). Linking Entrepreneurial Orientation and Firm Performance: The Role of Organizational Learning Capability and Innovation Performance . *Journal of Small Business Management*, 51(4):491–507.

[Argote, 2012] Argote, L. (2012). *Organizational Learning: Creating, Retaining and Transferring Knowledge*. Springer Science & Business Media.

[Argote et al., 1990] Argote, L., Beckman, S. L., and Epple, D. (1990). The Persistence and Transfer of Learning in Industrial Settings. *Management Science*, 36(2):140–154.

[Argote and Epple, 1990] Argote, L. and Epple, D. (1990). Learning Curves in Manufacturing. *Science*, 247(4945):920–924.

[Argote et al., 2003] Argote, L., McEvily, B., and Reagans, R. (2003). Managing Knowledge in Organizations: An Integrative Framework and Review of Emerging Themes. *Management Science*, 49(4):107–136.

[Argyres and Zenger, 2012] Argyres, N. S. and Zenger, T. R. (2012). Capabilities, Transaction Costs, and Firm Boundaries. *Organization Science*, 23(6):1643–1657.

[Arora and Gambardella, 1994] Arora, A. and Gambardella, A. (1994). The Changing Technology of Technological Change: General and Abstract Knowledge and the Division of Innovative Labour. *Research Policy*, 23(5):523–532.

[Arrow, 1972] Arrow, K. J. (1972). Economic Welfare and the Allocation of Resources for Invention. *Readings in Industrial Economics*, pages 219–236.

[Azoulay, 2004] Azoulay, P. (2004). Capturing Knowledge within and across Firm Boundaries: Evidence from Clinical Development. *American Economic Review*, 94(5):1591–1612.

[Balasubramanian and Sivadasan, 2011] Balasubramanian, N. and Sivadasan, J. (2011). What Happens When Firms Patent? New Evidence from U.S. Economic Census Data. *Review of Economics and Statistics*, 91(1):368–385.

[Barley et al., 2018] Barley, W. C., Treem, J. W., and Kuhn, T. (2018). Valuing Multiple Trajectories of Knowledge: A Critical Review and Agenda for Knowledge Management Research. *Academy of Management Annals*, 12(1):278–317.

[Baum et al., 2000] Baum, J. A. C., Calabrese, T., and Silverman, B. S. (2000). Don't Go It Alone: Alliance Network Composition and Startups' Performance in Canadian Biotechnology. *Strategic Management Journal*, 21(3):267–294.

[Baum and Dahlin, 2007] Baum, J. A. C. and Dahlin, K. B. (2007). Aspiration Performance and Railroads Patterns of Learning from Train Wrecks and Crashes. *Organization Science*, 18(3):368–385.

[Baum and Ingram, 1998] Baum, J. A. C. and Ingram, P. (1998). Survival-Enhancing Learning in the Manhattan Hotel Industry, 18981980. *Management Science*, 44(7):996–1016.

[Benkard, 2000] Benkard, C. L. (2000). Learning and Forgetting: The Dynamics of Aircraft Production. *American Economic Review*, 90(4):1034–1054.

[Berggren and Nacher, 2001] Berggren, E. and Nacher, T. (2001). Introducing New Products can be Hazardous to Your Company: Use the Right New-solutions Delivery Tools. *Academy of Management Perspectives*, 15(3):92–101.

[Bierly and Chakrabarti, 1996] Bierly, P. and Chakrabarti, A. (1996). Generic Knowledge Strategies in the U.S. Pharmaceutical Industry. *Strategic Management Journal*, pages 123–135.

[Blackwell et al., 2010] Blackwell, M., Iacus, S., King, G., and Porro, G. (2010). cem: Coarsened Exact Matching in Stata. *Stata Journal*, 9(4):524–546.

[Bogenrieder and Nooteboom, 2004] Bogenrieder, I. and Nooteboom, B. (2004). Learning Groups: What Types are there? A Theoretical Analysis and an Empirical Study in a Consultancy Firm . *Organization Studies*, 25(2):287–313.

[Booth and Zemmel, 2004] Booth, B. and Zemmel, R. (2004). Prospects for Productivity. *Nature Reviews Drug Discovery*, 3(5):451–456.

[Burgelman, 1983] Burgelman, R. A. (1983). A Process Model of Internal Corporate Venturing in the Diversified Major Firm. *Administrative Science Quarterly*, 28(2):223–244.

[Burt, 1992] Burt, R. S. (1992). The Social Structure of Competition. *Networks and organizations: Structure, form, and action*, 57:13–56.

[Burt, 1998] Burt, R. S. (1998). The Gender of Social Capita. *Rationality and Society*, 10(1):5–46.

[Cannon and Edmondson, 2005] Cannon, M. D. and Edmondson, A. C. (2005). Failing to Learn and Learning to Fail (Intelligently): How Great Organizations Put Failure to Work to Innovate and Improve. *Long Range Planning*, 38(3):299–319.

[Castellion and Markham, 2013] Castellion, G. and Markham, S. K. (2013). Perspective: New Product Failure Rates: Influence of Argumentum ad Populum and SelfInterest. *Journal of Product Innovation Management*, 30(5):976–979.

[Chuang and Baum, 2003] Chuang, Y.-T. and Baum, J. A. C. (2003). It's All in the Name: Failure-induced Learning by Multiunit Chains. *Administrative Science Quarterly*, 48(1):33–59.

[Coase, 1937] Coase, R. H. (1937). The Nature of the Firm. *Economica*, 16(4):386–405.

[Cohen and Levinthal, 1990] Cohen, W. M. and Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 35(1):128–152.

[Coleman, 1988] Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, 94:95–120.

[Colville et al., 2014] Colville, I., Hennestad, B., and Thoner, K. (2014). Organizing, Changing and Learning: A Sensemaking Perspective on an Ongoing Soap Story. *Management Learning*, 45(2):216–234.

[Cook et al., 2014] Cook, D., Brown, D., Alexander, R., March, R., Morgan, P., Satterthwaite, G., and Pangalos, M. N. (2014). Lessons Learned from the Fate of AstraZeneca's Drug Pipeline: a Five-dimensional Framework. *Nature Reviews Drug Discovery*, 13(6):419–431.

[Covert, 2014] Covert, T. R. (2014). Experiential and Social Learning in Firms: The Case of Hydraulic Fracturing in the Bakken Shale. *SSRN*.

[Cyert and March, 1963] Cyert, R. and March, J. (1963). *A Behavioral Theory of the Firm*. Englewood Cliffs.

[Daft and Weick, 1984] Daft, R. L. and Weick, K. E. (1984). Toward a Model of Organizations as Interpretation Systems. *Academy of Management Review*, 9(2):284–295.

[Dahlin et al., 2018] Dahlin, K. B., Chuang, Y.-T., and Roulet, T. J. (2018). Opportunity, Motivation, and Ability to Learn from Failures and Errors: Review, Synthesis, and Ways to Move Forward. *Academy of Management Review*, 12(1):252–277.

[David and Brachet, 2011] David, G. and Brachet, T. (2011). On the Determinants of Organizational Forgetting. *American Economic Journal: Microeconomics*, 3(3):284–295.

[de Ven and Polley, 1992] de Ven, A. H. V. and Polley, D. (1992). Learning While Innovating. *Organization Science*, 3(1):92–116.

[Dehejia and Wahba, 2002] Dehejia, R. H. and Wahba, S. (2002). Propensity Score-Matching Methods for Nonexperimental Causal Studies. *Review of Economics and statistics*, 84(1):151–161.

[Devarakonda and Reuer, 2018] Devarakonda, S. V. and Reuer, J. J. (2018). Knowledge Sharing and Safeguarding in R&D Collaborations: The Role of Steering Committees in Biotechnology Alliances. *Strategic Management Journal*, 39(7):1912–1934.

[DiMasi et al., 2003] DiMasi, J. A., Hansen, R. W., and Grabowski, H. G. (2003). The Price of Innovation: New Estimates of Drug Development Costs. *Journal of health economics*, 22(2):151–185.

[DiMasi et al., 1991] DiMasi, J. A., Hansen, R. W., Grabowski, H. G., and Lasagna, L. (1991). Cost of Innovation in the Pharmaceutical Industry. *Journal of health economics*, 10(2):284–295.

[Dosi, 1988] Dosi, G. (1988). Sources, Procedures, and Microeconomic Effects of Innovation. *Journal of Economic Literature*, 26(3):1120–1171.

[Dougherty, 1992] Dougherty, D. (1992). Interpretive Barriers to Successful Product Innovation in Large Firms. *Organization Science*, 3(2):179–202.

[du Plessis, 2007] du Plessis, M. (2007). The Role of Knowledge Management in Innovation. *Journal of knowledge management*, 11(4):20–29.

[Edmondson, 2011] Edmondson, A. C. (2011). Strategies for Learning from Failure. *Harvard Business Review*, 89(4):48–55.

[Egelman et al., 2016] Egelman, C. D., Epple, D., Argote, L., and Fuchs, E. R. H. (2016). Learning by Doing in Multiproduct Manufacturing: Variety, Customizations, and Overlapping Product Generations. *Management Science*, 62(2):405–423.

[Eggers, 2012] Eggers, J. (2012). All Experience is not Created Equal: Learning, Adapting, and Focusing in Product Portfolio Management. *Strategic Management Journal*, 33(3):315–335.

[Eggers, 2014] Eggers, J. (2014). Competing Technologies and Industry Evolution: The Benefits of Making Mistakes in the Flat Panel Display Industry. *Strategic Management Journal*, 35(2):499–514.

[Eisenhardt and Martin, 2004] Eisenhardt, K. M. and Martin, J. A. (2004). Dynamic Capabilities: What are They? *Strategic Management Journal*, 21(10-11):1105–1121.

[Fiol and Lyles, 1985] Fiol, C. M. and Lyles, M. A. (1985). Organizational Learning. *Academy of Management Review*, 10(4):803–813.

[Fleming, 2001] Fleming, L. (2001). Recombinant Uncertainty in Technological Search. *Management Science*, 47(1):117–132.

[Fleming and Sorenson, 2001] Fleming, L. and Sorenson, O. (2001). Technology as a Complex Adaptive System: Evidence from Patent Data. *Research Policy*, 30(7):1019–1039.

[Fleming and Sorenson, 2004] Fleming, L. and Sorenson, O. (2004). Science as a Map in Technological Search. *Strategic Management Journal*, 25(8-9):909–928.

[Freeman, 1991] Freeman, C. (1991). Networks of Innovators: A Synthesis of Research Issues. *Research Policy*, 20(5):499–514.

[Garcia and Calantone, 2002] Garcia, R. and Calantone, R. (2002). A Critical Look at Technological Innovation Typology and Innovativeness Terminology: a Literature Review . *Journal of Product Innovation Management*, 19(2):1545–1560.

[Gavetti and Levinthal, 2000] Gavetti, G. and Levinthal, D. (2000). Looking Forward and Looking Backward: Cognitive and Experiential Search. *Administrative Science Quarterly*, 45(1):109–122.

[Gill, 2012] Gill, C. J. (2012). How Often do US-based Human Subjects Research Studies Register on Time, and How Often do They Post Their Results? A Statistical Analysis of the Clinicaltrials.gov Database. *BMJ open*, 2(4).

[Gilsing and Nooteboom, 2006] Gilsing, V. and Nooteboom, B. (2006). Exploration and Exploitation in Innovation Systems: The Case of Pharmaceutical Biotechnology. *Research Policy*, 35(1):1–23.

[Gilsing et al., 2008] Gilsing, V., Nooteboom, B., Vanhaverbeke, W., Duysters, G., and den Oorda, A. (2008). Network Embeddedness and the Exploration of Novel Technologies: Technological Distance, Betweenness Centrality and Density. *Research Policy*, 37(10):1717–1731.

[Gomes-Casseres et al., 2006] Gomes-Casseres, B., Hagedoorn, J., and Jaffe, A. B. (2006). Do Alliances Promote Knowledge Flows? *Journal of Financial Economics*, 80(1):5–33.

[Gong et al., 2017] Gong, Y., Zhang, Y., and Xia, J. (2017). Do Firms Learn More From Small or Big Successes and Failures? A Test of the Outcome-Based Feedback Learning Perspective. *Journal of Management*.

[Grant, 1996] Grant, R. M. (1996). Toward a Knowledgebased Theory of the Firm. *Strategic Management Journal*, 17(S2):109–122.

[Greve, 2003] Greve, H. R. (2003). A Behavioral Theory of R&D Expenditures and Innovations: Evidence from Shipbuilding. *Academy of Management Journal*, 46(6):685–702.

[Griliches, 1990] Griliches, Z. (1990). Patent Statistics as Economic Indicators: a Survey. *National Bureau of Economic Research*, (w3301).

[Griliches, 1991] Griliches, Z. (1991). The Search for R&D Spillovers. *National Bureau of Economic Research*, (w3768).

[Grossman and Hart, 1986] Grossman, S. J. and Hart, O. D. (1986). The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy*, 94(4):691–719.

[Guedj and Scharfstein, 2004] Guedj, I. and Scharfstein, D. (2004). Organizational Scope and Investment: Evidence from the Drug Development Strategies and Performance of Biopharmaceutical Firms. *National Bureau of Economic Research*, (w10933).

[Gulati, 1998] Gulati, R. (1998). Alliances and networks. *Strategic Management Journal*, 19(4):293–317.

[Gulati et al., 2000] Gulati, R., Nohria, N., and Zaheer, A. (2000). Strategic networks. *Strategic Management Journal*, 21(3):203–215.

[Gulati and Singh, 1998] Gulati, R. and Singh, H. (1998). The Architecture of Cooperation: Managing Coordination Costs and Appropriation Concerns in Strategic Alliances. *Administrative Science Quarterly*, 43(4):781–814.

[Gulati et al., 2012] Gulati, R., Wohlgezogen, F., and Zhelyazkov, P. (2012). The Two Facets of Collaboration: Cooperation and Coordination in Strategic Alliances. *The Academy of Management Annals*, 6(1):531–583.

[Hall et al., 2001] Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. *National Bureau of Economic Research*, (w8498).

[Hannan and Freeman, 1977] Hannan, M. T. and Freeman, J. (1977). The Population Ecology of Organizations. *American Journal of Sociology*, 82(5):929–964.

[Hargadon and Sutton, 1997] Hargadon, A. and Sutton, R. I. (1997). Technology Brokering and Innovation in a Product Development Firm. *Administrative Science Quarterly*, 42(4):716–749.

[Haunschild and Rhee, 2004] Haunschild, P. R. and Rhee, M. (2004). The Role of Volition in Organizational Learning: The Case of Automotive Product Recalls. *Management Science*, 50(11):1545–1560.

[Haunschild and Sullivan, 2002] Haunschild, P. R. and Sullivan, B. N. (2002). Learning from Complexity: Effects of Prior Accidents and Incidents on Airlines' Learning. *Administrative Science Quarterly*, 47(4):609–643.

[Henderson and Stern, 2004] Henderson, A. D. and Stern, I. (2004). Selection-Based Learning: The Coevolution of Internal and External Selection in High-Velocity Environments. *Administrative Science Quarterly*, 49(1):39–75.

[Henderson and Clark, 1990] Henderson, R. M. and Clark, K. B. (1990). Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms. *Administrative Science Quarterly*, 35(1):9–30.

[Henderson and Cockburn, 1994] Henderson, R. M. and Cockburn, I. (1994). Measuring Competence? Exploring Firm Effects in Pharmaceutical Research. *Strategic Management Journal*, 15(S1):63–84.

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

[Hirano and Imbens, 2001] Hirano, K. and Imbens, G. W. (2001). Estimation of Causal Effects using Propensity Score Weighting: An Application to Data on Right Heart Catheterization. *Health Services and Outcomes Research Methodology*, 2(3-4):259–278.

[Hoang and Yi, 2015] Hoang, H. and Yi, A. (2015). Network-based Research in Entrepreneurship: A Decade in Review. *Foundations and Trends in Entrepreneurship*, 11(1):525–539.

[Holt and Cornelissen, 2014] Holt, R. and Cornelissen, J. (2014). Sensemaking revisited. *Management Learning*, 45(5):525–539.

[Hsiao, 2005] Hsiao, C. (2005). Why panel data? *The Singapore Economic Review*, 50(02):143–154.

[Hsiao et al., 1993] Hsiao, C., Appelbe, T. W., and Dineen, C. R. (1993). A General Framework for Panel Data Models with an Application to Canadian Customer-dialed Long Distance Telephone Service. *Journal of Econometrics*, 59(1-2):63–86.

[Hubbard, 2008] Hubbard, T. N. (2008). Empirical Research on Firms' Boundaries. *Canadian Journal of Economics*, 41(2):341–359.

[Huber, 1985] Huber, G. P. (1985). Temporal Stability and Response-Order Biases in Participant Descriptions of Organizational Decisions. *Academy of Management Journal*, 28(4):943–950.

[Ingram and Baum, 1997] Ingram, P. and Baum, J. A. C. (1997). Opportunity and Constraint: Organizations' Learning from the Operating and Competitive Experience of Industries. *Strategic Management Journal*, 18(S1):75–98.

[Irwin and Klenow, 1994] Irwin, D. A. and Klenow, P. J. (1994). Learning-by-Doing Spillovers in the Semiconductor Industry. *Journal of Political Economy*, 102(6):1200–1227.

[Islam, 1995] Islam, N. (1995). Growth Empirics: a Panel Data Approach. *The Quarterly Journal of Economics*, 110(4):1127–1170.

[Jaffe and Palmer, 2006] Jaffe, A. B. and Palmer, K. (2006). Environmental Regulation and Innovation: A Panel Data Study. *Review of Economics and Statistics*, 79(4):610–619.

[Jaffe et al., 1993] Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics*, 108(3):577–598.

[Jassawalla and Sashittal, 1998] Jassawalla, A. R. and Sashittal, H. C. (1998). An Examination of Collaboration in High-Technology New Product Development Processes. *Journal of Product Innovation Management*, 15(3):237–254.

[Jenkins, 2005] Jenkins, S. P. (2005). *Survival analysis*. Institute for Social and Economic Research, University of Essex.

[Jovanovic and Nyarko, 1996] Jovanovic, B. and Nyarko, Y. (1996). Learning By Doing and the Choice of Technology. *Econometrica*, 64(6):1299–1310.

[Kaplan et al., 2001] Kaplan, S., Schenkel, A., von Krogh, G., and Weber, C. (2001). Knowledge-Based Theories of the Firm in Strategic Management: A Review and Extension. *International Journal of Project Management*, pages 143–158.

[Karlssona et al., 2004] Karlssona, M., Trygga, L., and Elfstrmb, B.-O. (2004). Measuring R&D Productivity: Complementing the Picture by Focusing on Research Activities. *Technovation*, 24(3):179–186.

[Katila and Ahuja, 2002] Katila, R. and Ahuja, G. (2002). Something Old, Something New: A Longitudinal Study of Search Behavior and New Product Introduction. *Academy of Management Journal*, 45(6):1183–1194.

[Katz and Allen, 1982] Katz, R. and Allen, T. J. (1982). Investigating the Not Invented Here (NIH) syndrome: A look at the performance, tenure, and communication patterns of 50 R & D Project Groups. *R&D Management*, 12(1):7–20.

[Kauffman, 1993] Kauffman, S. A. (1993). *The Origins of Order: Self Organization and Selection in Evolution*. Oxford University Press.

[KC et al., 2013] KC, D., Staats, B. R., and Gino, F. (2013). Learning from My Success and from Others' Failure: Evidence from Minimally Invasive Cardiac Surgery. *Management Science*, 59(11):2435–2449.

[Kellogg, 2011] Kellogg, R. (2011). Learning by Drilling: Interfirm Learning and Relationship Persistence in the Texas Oilpatch. *The Quarterly Journal of Economics*, 126(4):1961–2004.

[Kim and Miner, 2007] Kim, J.-Y. J. and Miner, A. S. (2007). Vicarious Learning from the Failures and Near-Failures of Others: Evidence from the U.S. Commercial Banking Industry. *Academy of Management Journal*, 50(3):687–714.

[Kogan et al., 2017] Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological Innovation, Resource Allocation, and Growth. *The Quarterly Journal of Economics*, 132(2):665–712.

[Kogut and Zander, 1996] Kogut, B. and Zander, U. (1996). What Firms do? Coordination, Identity, and Learning. *Organization Science*, 7(5):469–592.

[Kola and Landis, 2004] Kola, I. and Landis, J. (2004). Can the Pharmaceutical Industry Reduce Attrition Rates? *Nature Reviews Drug discovery*, 3(8):711–716.

[Kraatz, 1998] Kraatz, M. S. (1998). Learning by Association? Interorganizational Networks and Adaptation to Environmental Change. *Academy of Management Journal*, 41(6):621–643.

[Lant, 1992] Lant, T. K. (1992). Aspiration Level Adaptation: An Empirical Exploration. *Management Science*, 38(5):609–755.

[Lant and Mezias, 1992] Lant, T. K. and Mezias, S. J. (1992). An Organizational Learning Model of Convergence and Reorientation. *Organization Science*, 3(1):47–71.

[Lavie et al., 2007] Lavie, D., Lechner, C., and Singh, H. (2007). The Performance Implications of Timing of Entry and Involvement in Multipartner Alliances. *Academy of Management Journal*, 50(3):578–604.

[Levin, 2000] Levin, D. Z. (2000). Organizational Learning and the Transfer of Knowledge: An Investigation of Quality Improvement. *Organization Science*, 11(6):630–647.

[Levinthal, 1997] Levinthal, D. A. (1997). Adaptation on Rugged Landscapes. *Management Science*, 43(7):895–1045.

[Levinthal and March, 1981] Levinthal, D. A. and March, J. G. (1981). A Model of Adaptive Organizational Search. *Journal of Economic Behavior and Organization*, 2(4):307–333.

[Levitt and March, 1988] Levitt, B. and March, J. G. (1988). Organizational Learning. *Annual Review of Sociology*, 14:319–340.

[Levitt et al., 2013] Levitt, S. D., List, J. A., and Syverson, C. (2013). Toward an Understanding of Learning by Doing: Evidence from an Automobile Assembly Plant. *Journal of Political Economy*, 121(4):643–681.

[Li et al., 2008] Li, D., Eden, L., Hitt, M. A., and Ireland, R. D. (2008). Friends, Acquaintances, or Strangers? Partner Selection in R&D Alliances. *Academy of Management Journal*, 51(2):315–334.

[Li et al., 2011] Li, D., Eden, L., Hitt, M. A., Ireland, R. D., and Garrett, R. P. (2011). Governance in Multilateral R&D Alliances. *Organization Science*, 23(4):1191–1210.

[Lucas, 1993] Lucas, R. E. (1993). Making a Miracle. *Econometrica*, 61(2):251–272.

[Macher and Boerner, 2012] Macher, J. T. and Boerner, C. (2012). Technological Development at the Boundaries of the Firm: a Knowledgebased Examination in Drug Development. *Strategic Management Journal*, 33(9):1016–1036.

[Madhavan and Grover, 1998] Madhavan, R. and Grover, R. (1998). From Embedded Knowledge to Embodied Knowledge: New Product Development as Knowledge Management. *The Journal of marketing*, 62(4):1–12.

[Madsen and Desai, 2010] Madsen, P. M. and Desai, V. (2010). Failing to Learn? The Effects of Failure and Success on Organizational Learning in the Global Orbital Launch Vehicle Industry. *Academy of Management Journal*, 53(3):451–476.

[Maitlis and Sonenshein, 2010] Maitlis, S. and Sonenshein, S. (2010). Sensemaking in Crisis and Change: Inspiration and Insights From Weick (1988). *Journal of Management Studie*, 47(3):551–580.

[Mansfield, 1981] Mansfield, E. (1981). Composition of R and D Expenditures: Relationship to Size of Firm, Concentration, and Innovative Output. *The Review of Economics and Statistics*, 63(4):610–615.

[March, 1982] March, J. G. (1982). Theories of Choice and Making Decisions. *Society*, 20(1):118–131.

[March, 1991] March, J. G. (1991). Exploration and Exploitation in Organizational Learning. *Organization Science*, 2(1):71–87.

[March and Olsen, 1975] March, J. G. and Olsen, J. P. (1975). The Uncertainty of the Past: Organizational Learning Under Ambiguity. *European Journal of Political Research*, 3(2):147–171.

[March and Shapira, 1992] March, J. G. and Shapira, Z. (1992). Variable Risk Preferences and the Focus of Attention. *Psychological Review*, 99(1):551–580.

[March et al., 1991] March, J. G., Sproull, L. S., and Tamuz, M. (1991). Learning from Samples of One or Fewer. *Organization science*, 2(1):1–13.

[Marsili and Salter, 2005] Marsili, O. and Salter, A. (2005). Inequality of Innovation: Skewed Distributions and the Returns to Innovation in Dutch Manufacturing. *Economics of Innovation and New Technology*, 14(1-2):83–102.

[McGrath, 2001] McGrath, R. G. (2001). Exploratory Learning, Innovative Capacity, and Managerial Oversight. *Academy of Management Journal*, 44(1):118–131.

[Meuleman et al., 2010] Meuleman, M., Lockett, A., Manigart, S., and Wright, M. (2010). Partner Selection Decisions in Interfirm Collaborations: The Paradox of Relational Embeddedness. *Journal of Management Studies*, 47(6):995–1019.

[Meyer and Goes, 1988] Meyer, A. D. and Goes, J. B. (1988). Organizational Assimilation of Innovations: A Multilevel Contextual Analysis. *Academy of Management Journal*, 31(3):897–923.

[Moustaghfir and Schiuma, 2013] Moustaghfir, K. and Schiuma, G. (2013). Knowledge, Learning, and Innovation: Research and Perspectives. *Journal of Knowledge Management*, 17(4):495–510.

[Muehlfeld et al., 2012] Muehlfeld, K., Sahib, P. R., and Witteloostuijn, A. V. (2012). A Contextual Theory of Organizational Learning from Failures and Successes: A Study of Acquisition Completion in the Global Newspaper Industry, 19812008. *Strategic Management Journal*, 33(8):938–964.

[Myers and Turnbull, 1977] Myers, S. C. and Turnbull, S. M. (1977). Capital Budgeting and the Capital Asset Pricing Model: Good News and Bad News. *The Journal of Finance*, 32(2):321–333.

[Narendranathan and Stewart, 1991] Narendranathan, W. and Stewart, M. B. (1991). PRACTITIONERS' CORNER: Simple Methods for Testing for the Proportionality of CauseSpecific Hazards in Competing Risks Models. *Oxford Bulletin of Economics and Statistics*, 53(3):331–340.

[Nelson, 1982] Nelson, R. R. (1982). The Role of Knowledge in R&D Efficiency . *The Quarterly Journal of Economics*, 97(3):453–470.

[Nelson and Winter, 1973] Nelson, R. R. and Winter, S. G. (1973). Toward an Evolutionary Theory of Economic Capabilities. *The American Economic Review*, 63(2):440–449.

[Nelson and Winter, 1982] Nelson, R. R. and Winter, S. G. (1982). *An Evolutionary Theory of Economic Change*. Harvard University Press.

[Nerkar and Roberts, 2004] Nerkar, A. and Roberts, P. W. (2004). Technological and Productmarket Experience and the Success of New Product Introductions in the Pharmaceutical Industry. *Strategic Management Journal*, 25(8-9):779–799.

[Nickson and Zenger, 2004] Nickson, J. A. and Zenger, T. R. (2004). A Knowledge-Based Theory of the FirmThe Problem-Solving Perspective. *Organization Science*, 15(6):617–734.

[Nohria and Gulati, 1996] Nohria, N. and Gulati, R. (1996). Is Slack Good or Bad for Innovation? *Academy of Management Journal*, 39(5):1245–1264.

[Nooteboom, 2004] Nooteboom, B. (2004). Innovation, Learning and Industrial Organisation. *Cambridge Journal of Economics*, 23(2):127–150.

[Orsenigo et al., 2001] Orsenigo, L., Pammolli, F., and Riccabonic, M. (2001). Technological Change and Network Dynamics: Lessons from the Pharmaceutical Industry. *Research Policy*, 30(3):485–508.

[Owen-Smith and Powell, 2004] Owen-Smith, J. and Powell, W. W. (2004). Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community. *Organization Science*, 15(1):5–21.

[Oxley, 1997] Oxley, J. E. (1997). Appropriability Hazards and Governance in Strategic Alliances: A Transaction Cost Approach. *The Journal of Law, Economics, and Organization*, 13(2):5–21.

[Pablo Martin de Holan, 2004] Pablo Martin de Holan, N. P. (2004). Remembrance of Things Past? The Dynamics of Organizational Forgetting. *Management Science*, 50(11):1603–1611.

[Pammolli et al., 2011] Pammolli, F., Magazzini, L., and Riccaboni, M. (2011). The Productivity Crisis in Pharmaceutical R&D. *Nature Reviews Drug Discovery*, 10(6):428–438.

[Paul et al., 2010] Paul, S. M., Mytelka, D. S., Dunwiddie, C. T., Persinger, C. C., Munos, B. H., Lindborg, S. R., and Schacht, A. L. (2010). How to Improve R&D

Productivity: the Pharmaceutical Industry's Grand Challenge. *Nature Reviews Drug Discovery*, 9(3):203–214.

[Penrose, 2009] Penrose, E. (2009). *The Theory of the Growth of the Firm*. Oxford University Press.

[Phelps, 2010] Phelps, C. C. (2010). A Longitudinal Study of the Influence of Alliance Network Structure and Composition on Firm Exploratory Innovation. *Academy of Management Journal*, 53(4):890–913.

[Phelps et al., 2012] Phelps, C. C., Heidl, R., and Wadhwa, A. (2012). Knowledge, Networks, and Knowledge Networks A Review and Research Agenda. *Journal of Management*, 38(4):1115–1166.

[Pisano, 1989] Pisano, G. P. (1989). Using Equity Participation to Support Exchange: Evidence from the Biotechnology Industry. *Journal of Law, Economics, & Organization*, 5(1):109–126.

[Pisano, 2006] Pisano, G. P. (2006). *Science Business: The Promise, the Reality, and the Future of Biotech*. Harvard University Press.

[Polidoro et al., 2011] Polidoro, F., Ahuja, G., and Mitchell, W. (2011). When the Social Structure Overshadows Competitive Incentives: The Effects of Network Embeddedness on Joint Venture Dissolution. *Academy of Management Journal*, 54(1):203–223.

[Posen et al., 2018] Posen, H. E., Keil, T., Kim, S., and Meissner, F. D. (2018). Renewing Research on Problemistic SearchA Review and Research Agenda. *Academy of Management Annals*, 12(1):208–251.

[Powell, 1990] Powell, W. W. (1990). Neither Market nor Hierarchy: Network Forms of Organization. *Research in Organizational Behavior*, 12:295–336.

[Powell et al., 1996] Powell, W. W., Koput, K. W., and Smith-Doerr, L. (1996). Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. *Administrative Science Quarterly*, 41(1):116–145.

[Powell et al., 2005] Powell, W. W., White, D. R., Koput, K. W., and OwenSmith, J. (2005). Network Dynamics and Field Evolution: The Growth of Interorganizational Collaboration in the Life Sciences. *American journal of sociology*, 110(4):1132–1205.

[Raub and Weesie, 1990] Raub, W. and Weesie, J. (1990). Reputation and Efficiency in Social Interactions: An Example of Network Effects. *American Journal of Sociology*, 96(3):626–654.

[Rerup, 2009] Rerup, C. (2009). Attentional Triangulation: Learning from Unexpected Rare Crises. *Organization Science*, 20(5):876–893.

[Reuer and Devarakonda, 2017] Reuer, J. J. and Devarakonda, R. (2017). Partner Selection in R&D Collaborations: Effects of Affiliations with Venture Capitalists. *Organization Science*, 28(3):574–595.

[Reuer et al., 2002] Reuer, J. J., Zollo, M., and Singh, H. (2002). Postformation Dynamics in Strategic Alliances. *Strategic Management Journal*, 23(2):135–151.

[Roberts, 1999] Roberts, P. W. (1999). Product Innovation, Productmarket Competition and Persistent Profitability in the U.S. Pharmaceutical Industry. *Strategic Management Journal*, 20(7):655–670.

[Roijakkers and Hagedoorn, 2006] Roijakkers, N. and Hagedoorn, J. (2006). Interfirm R&D Partnering in Pharmaceutical Biotechnology Since 1975: Trends, Patterns, and Networks. *Research Policy*, 35(3):431–446.

[Romer, 1990] Romer, P. M. (1990). Endogenous Technological Change. *Journal of Polotical Economic*, 98(5):71–102.

[Rosenbaum and Rubin, 1983] Rosenbaum, P. R. and Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70(1):41–55.

[Rothaermel and Deeds, 2004] Rothaermel, F. T. and Deeds, D. L. (2004). Exploration and Exploitation Alliances in Biotechnology: a System of New Product Development. *Strategic Management Journal*, 25(3):201–221.

[Rowley et al., 2000] Rowley, T., Behrens, D., and Krackhardt, D. (2000). Redundant Governance Structures: an Analysis of Structural and Relational Embeddedness in the Steel and Semiconductor Industries . *Strategic Management Journal*, 21(3):369–386.

[Sampson, 2004] Sampson, R. C. (2004). The Cost of Misaligned Governance in R&D Alliances . *The Journal of Law, Economics, and Organization*, 20(2):484–526.

[Sampson, 2007] Sampson, R. C. (2007). R&D Alliances and Firm Performance: The Impact of Technological Diversity and Alliance Organization on Innovation. *Academy of Management Journal*, 50(2):364–368.

[Scannell et al., 2012] Scannell, J. W., Blanckley, A., Boldon, H., and Warrington, B. (2012). Diagnosing the Decline in Pharmaceutical R&D Efficiency. *Nature Reviews Drug discovery*, 11(3):191–200.

[Scherer and Ross, 1980] Scherer, F. M. and Ross, D. (1980). *Industrial market structure and economic performance*. Houghton Mifflin.

[Schilling, 1998] Schilling, M. A. (1998). Technological Lockout: An Integrative Model of the Economic and Strategic Factors Driving Technology Success and Failure. *Academy of Management Review*, 23(2):267–284.

[Schumpeter, 2010] Schumpeter, J. A. (2010). *Capitalism, Socialism and Democracy*. Routledge.

[Shiffrin and Schneider, 1977] Shiffrin, R. M. and Schneider, W. (1977). Controlled and Automatic Human Information Processing: II. Perceptual Learning, Automatic Attending and a General Theory. *Psychological Review*, 84(2):127–190.

[Shipilov and Li, 2008] Shipilov, A. V. and Li, S. X. (2008). Can you have your Cake and Eat it too? Structural Holes' Influence on Status Accumulation and Market Performance in Collaborative Networks. *Administrative Science Quarterly*, 53(1):73–108.

[Simon, 1979] Simon, H. A. (1979). Rational Decision Making in Business Organizations. *The American Economic Review*, 69(4):127–190.

[Singh, 1986] Singh, J. V. (1986). Performance, Slack, and Risk Taking in Organizational Decision Making. *Academy of Management Journal*, 29(3):562–585.

[Sitkin, 1992] Sitkin, S. B. (1992). Learning Through Failure: The Strategy of Small Losses. *Research in Organizational Behavior*, 14:231–266.

[Solow, 1994] Solow, R. M. (1994). Perspectives on Growth Theory. *Journal of Economic Perspectives*, 8(1):45–54.

[Starbuck et al., 1978] Starbuck, W. H., Greve, A., and Hedberg, B. (1978). Responding to Crises. *Journal of Business Administration*, 9(2):111–137.

[Steven et al., 2010] Steven, P. M., Mytelka, D. S., Dunwiddie, C. T., Persinger, C. C., Munos, B. H., Lindborg, S. R., and Schacht, A. L. (2010). How to Improve R&D Productivity: the Pharmaceutical Industry's Grand Challenge. *Nature reviews Drug discovery*, 9(3):203–214.

[Stuart and Podolny, 1996] Stuart, T. E. and Podolny, J. M. (1996). Local Search and the Evolution of Technological Capabilities. *Strategic Management Journal*, 17(S1):21–38.

[Syverson, 2011] Syverson, C. (2011). What Determines Productivity? *Journal of Economic literature*, 49(2):326–365.

[Teece, 1986] Teece, D. J. (1986). Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy. *Research Policy*, 15(6):285–305.

[Thompson, 2010] Thompson, P. (2010). Learning by Doing. *Handbook of the Economics of Innovation*, 1:285–305.

[Thornton and Thompson, 2001] Thornton, R. A. and Thompson, P. (2001). Learning from Experience and Learning from Others: An Exploration of Learning and Spillovers in Wartime Shipbuilding. *American Economic Review*, 91(5):1350–1368.

[Tushman and Anderson, 1986] Tushman, M. L. and Anderson, P. (1986). Technological Discontinuities and Organizational Environments. *Administrative Science Quarterly*, 31(3):439–465.

[Tushman and Rosenkopf, 1996] Tushman, M. L. and Rosenkopf, L. (1996). Executive Succession, Strategic Reorientation and Performance Growth: A Longitudinal Study in the U.S. Cement Industry. *Management Science*, 42(7):939–953.

[Utterback and Abernathy, 1975] Utterback, J. M. and Abernathy, W. J. (1975). A Dynamic Model of Process andPproduct Innovation. *Omega*, 3(6):639–656.

[Uzzi, 1997] Uzzi, B. (1997). Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness . *Administrative Science Quarterly*, 42(1):35–67.

[Walker et al., 1997] Walker, G., Kogut, B., and Shan, W. (1997). Social Capital, Structural Holes and the Formation of an Industry Network. *Organization Science*, 8(2):109–125.

[Warsh, 2007] Warsh, D. (2007). *Knowledge and the Wealth of Nations: A Story of Economic Discovery*. W.W. Norton&Company.

[Weick, 1995] Weick, K. E. (1995). *Sensemaking in Organizations*. Sage Publications.

[Wildavsky, 1988] Wildavsky, A. (1988). *Searching for Safety*. Transaction Publishers.

[Williamson, 1991] Williamson, O. E. (1991). Comparative Economic Organization: The Analysis of Discrete Structural Alternatives. *Administrative Science Quarterly*, 36(2):269–296.

[Williamson, 1999] Williamson, O. E. (1999). Strategy Research: Governance and Competence Perspectives. *Strategic Management Journal*, 20(12):1087–1108.

[Wuyts et al., 2005] Wuyts, S., G.Colombo, M., Dutta, S., and Nooteboom, B. (2005). Empirical Tests of Optimal Cognitive Distance. *Journal of Economic Behavior & Organization*, 58(2):277–302.

[Zaheer and Bell, 2005] Zaheer, A. and Bell, G. G. (2005). Benefiting from Network Position: Firm Capabilities, Structural Holes, and Performance. *Strategic Management Journal*, 26(9):809–825.