KU LEUVEN

FACULTY OF ECONOMICS AND BUSINESS

Essays on Competition and Innovation in the Pharmaceutical Industry

Dissertation presented to obtain the degree of Doctor in Business Economics

by

Jan Málek

Number 799 2022

Essays on Competition and Innovation in the Pharmaceutical Industry

Dissertation presented to obtain the degree of Doctor in Business Economics

by

Jan Málek

Number 799 2022

Doctoral Committee

Supervisor:

Prof. Dr. Jo Seldeslachts	KU Leuven & DIW Berlin

Co-Supervisor:

Prof. Dr. Reinhilde Veugelers	KU Leuven

Doctoral committee:

Prof. Dr. Bruno Cassiman	KU Leuven
Prof. Dr. Tomaso Duso	TU Berlin & DIW Berlin
Prof. Dr. Florian Szücs	WU Vienna

Chairman:

Prof. Dr. Sam Arts	KU Leuven

Acknowledgments

Almost five years ago, I had decided to embark on an adventurous PhD journey. At that time, I knew that it was going to be a bumpy ride, but I also knew that there was hardly any better choice allowing me to develop the skills, expertise, and contacts needed to pursue my dreams in the field I love: competition economics.

Two people made this journey possible in the first place - Jo Seldeslachts and Reinhilde Veugelers. Jo, my supervisor, could be praised a lot for these acknowledgements. However, let me just highlight one of his best features as a supervisor - the fact that he truly cares and supports his students with exceptional dedication and generosity. Jo has always been very generous with the time he spends on guidance, feedback, coaching, informal chats (often in a nice restaurant) or simply listening to worries and making sure that all essentially works out. In other words, he is a great mentor on such an adventurous journey. Reinhilde, my co-supervisor, is a passionate researcher, and my learning and research progress benefited tremendously not just from her sharp mind and original ideas, but also from her dedication to always discuss the most peculiar problem and make our research better. Apart from Jo and Reinhilde, I also thank the DIW Graduate Center and the Flemish Research Foundation for funding my PhD.

I am very grateful to my committee members: Sam Arts, Bruno Cassiman, Florian Szücs, and Tomaso Duso - all fantastic researchers who took the time to ask the right questions and provide very helpful comments. On top of this, I must specifically thank Tomaso for his role as a department head. He maintains a great environment at the Firms and Markets department at DIW Berlin and has always been there to chat or give advice which kept me going. I also thank my coauthors, Melissa Newham and Marcel Wieting, who have accompanied me on this journey. Apart from learning a lot from them, the many laughs (and bits of gossip) made my PhD a much funnier experience. I thank all the other researchers at MSI and DIW for their valuable input during seminars and, of course, for being a nice company beyond research.

An inseparable part of my journey were the numerous activities I kept pursuing alongside it in the private sector and the public sector. Needless to say, these benefited me professionally a lot. More importantly, they also allowed me to meet great people who were always happy to share their experiences, and who inspired and supported me. I would

particularly like to mention Ulrich Puls and David Cayet, who showed me the secrets of an incredibly interesting monitoring trustee business and took me to fancy restaurants to always have a super interesting and refreshing talk. I would also like to mention Daniel Coublucq whom I met at DG Competition's Chief Economist Team. Being a great economist with whom I had a lot of fun working on cases, he also proved to be an incredibly nice person and another great mentor who is always happy to give me advice. Last, but not least, I would also like to mention Jakub Chini - a colleague turned friend - with whom I worked on a couple of cool projects and with whom I can spend endless hours laughing about the nitty-gritty details of our (Czech) competition world in attempts to cultivate it.

This naturally brings me to the groups of friends I met during the last five years who have always provided an incredible amount of support and were there for me in both good and bad times. To name a few, I thank Dennis Gaus, Shan Huang, Jan Hloušek, Boryana Ilieva, Stefano Piasenti, Raphael Salien, Thomas Seron, Stefano Selis, and, especially, Iuliia Grabova for being here.

Last but not least, the biggest thank you goes to my partner Peng, my brothers Ondra and Dan, my father Radek, and, especially, my mother Jana for their infinite love and everyday support throughout this journey. This would not have been possible without you on my side!

Berlin, July 2022

Contents

List of Figures

List of Tables

Chapter 1

General introduction

Innovation is essential for progress, long-term growth, and consumer welfare. As such, there is almost a universal consensus among economists that innovation must be not only protected but promoted. However, the role of competition in pursuing this goal remains unclear. The relationship between competition and innovation is complex and ambiguous (Aghion *et al.*, 2005; Shapiro, 2011). Discussions are thus still ongoing regarding policies that policymakers and competition authorities should adopt to spur innovation. Using the setting of one of the most important and innovative sectors - the pharmaceutical industry - this dissertation deepens our understanding and brings important empirical insights to these academic and policy debates by focusing on two specific topics. The first topic studies how changes in the competitive landscape through mergers and acquisitions (M&As) affect innovation. The second topic then explores how venture capital funds react to competition and shape the direction of early-stage innovation.

The pharmaceutical industry and diabetes

Since the pharmaceutical industry is the common setting for all chapters, this introduction proceeds with a brief description of the institutional setting and the pharmaceutical innovation process. Developing innovative and life-saving medical treatments, the pharmaceutical industry is one of the most innovative sectors worldwide (Grassano *et al.*, 2021). It is thus not surprising that the industry is very dynamic, with rich and varied activity in terms of research and development (R&D), mergers and acquisitions (M&A), and venture capital (VC) investments.[1]

Pharmaceutical innovation happens at the level of individual well-defined research projects (chemical molecules or therapeutic proteins). A project's development is always centred around a specific disease it should target (e.g., Diabetes type II) and the bio-

[1] Pharmaceutical companies invested $170 billion in R&D in 2019 (Deloitte, 2020). The value of pharmaceutical M&As reached $360 billion (Behner & Spence, 2020) and the investment activity of VC funds in the US biotech sector reached $17 billion in 2019 (BiopharmaDive, 2020).

chemical process through which the project produces the desired effect in the body - the so-called "Mechanism of action" (MoA).

This dissertation specifically focuses on projects undertaken by the private sector between 1997 and 2017 to treat diabetes. Diabetes is a disease caused by the insufficient ability (Diabetes type II) or complete inability (Diabetes type I) of the body to produce insulin - a hormone regulating levels of glucose in the body. Diabetes has no cure (yet), is widespread, and grows rapidly. In 2021, approximately 537 million adults were living with diabetes worldwide and projections indicate that by 2045 it will be up to 783 million (International Diabetes Federation, 2021). By focusing on one disease, this dissertation can concentrate efforts on the quality of the collected information.

Since the development process is regulated and milestones are disclosed to public authorities, each project in diabetes and pharma, in general, can be also tracked throughout its development from inception until termination (if unsuccessful) or launch (if successful). A project begins with discovery and testing in a laboratory. If promising, researchers undertake preclinical experiments and testing in animal subjects. If the preclinical results are satisfactory, the project continues to three phases of clinical trials in humans. In Phase I, the safety of the drug is tested with a small sample of healthy individuals. In Phase II, the efficacy of the drug is tested on a larger group of people. Phase III trials involve large groups of subjects and aim to provide a definitive assessment of how effective the drug is, often using randomized control trials. The drug development is funnel-shaped with many potential drug candidates entering the development process but only 6% passing the last phase of testing (Pammolli *et al.*, 2011).

Due to the huge development costs and low chances of success, nearly all projects are protected by patents. Patent filings typically happen early on during the preclinical phase, as soon as a target molecule is identified, and before entering clinical trials where detailed information about the project's nature is disclosed. The project's patents serve two prominent functions. First, these patents give companies the right to exclude others from using the substance or technology, granting them a temporary monopoly during which profits can be realized to recoup the upfront development costs. The patent lifetime lasts typically 20 years but can be extended to compensate for the lengthy development timelines, taking on average 12 years (Branstetter *et al.*, 2014). Second, being precise legal documents, patents contain rich information about the project's underlying technology, its exact scope as well as positioning concerning other known technologies, allowing assessments of the project's *technological nature* and position in the *technology space*.

The identity of patents behind specific projects is not disclosed if a project is not successful (launched). Therefore, it has not been so far possible to utilize this information in research and policy alike. This dissertation is the first to link projects under development to their underlying patents by developing a unique algorithm which uses the text of patent documents and the (bio)chemical properties of projects. This level of detail allows analysis

of the above topics at a much deeper level, opens up new research angles and also sets the stage for follow-on research.[2]

Mergers and acquisitions (M&As)

The first two chapters of this dissertation study how changes in the competitive landscape through mergers and acquisitions (M&As) affect innovation. New discussions on this topic have been sparked by a series of theoretical and empirical papers in industrial organization, demonstrating that M&As in highly innovative sectors can be used as a hindrance to innovation (to name a few, see for example Federico *et al.*, 2017; Motta & Peitz, 2020; Gilbert, 2018; Affeldt & Kesler, 2021b). Whilst the mechanisms differ and researchers have hinted at "killer acquisitions" (Cunningham *et al.*, 2021), "reverse killer acquisitions" (Caffarra *et al.*, 2020), or "kill-zone" (Kamepalli *et al.*, 2020), the common suspect driving the potentially negative innovation outcomes is the market power of the acquiring firms. These firms with market power - incumbents - wish to protect their existing markets and do so at the expense of innovation.

Emphasizing the position of acquirers in the product market space, this new stream of literature has however omitted to account for similar considerations in the *technology* dimension. This has been the case despite the literature in innovation, management, and finance repeatedly highlighting the role of *technology* as the key to understanding the relationship between M&As and innovation, especially in highly innovative sectors (see for example Ahuja & Katila, 2001; Cassiman *et al.*, 2005; Cloodt *et al.*, 2006; Ornaghi, 2009a). The first two chapters of this dissertation thus attempt to connect the ongoing M&A discussions about market power to the technology dimensions and close this gap.

Since detailed empirical evidence is scarce, Chapter 2 (*"Acquiring innovation: who, when, and what?"*, joint work with Melissa Newham, Jo Seldeslachts, and Reinhilde Veugelers) starts by analyzing patterns in M&A deals. Zooming inside of firms at the level of individual projects and considering R&D activities along the entire development pipeline, this paper studies the role of three characteristics as key determinants of M&As: targets' and acquirers' identity in terms of size and product market incumbency (who), timing of acquisitions (when), and type of projects in terms of their technological "high risk/high gain" potential (what). This detailed analysis uncovers that the transaction landscape is rich and varied and goes beyond the narrative of large incumbents buying young research-focused firms without marketed products. Our focus on technology characteristics allows showing that the so far unexplored technological uncertainty (early stages) and high-risk/high-gain technology profiles of projects are an essential part of the changing-of-ownership stories.

Chapter 3 (*"M&As spurring or stifling innovation?"*, joint work with Jo Seldeslachts and Reinhilde Veugelers) provides empirical evidence on which M&A deals involving small

[2]Whilst limited to antidiabetic R&D in this dissertation, the methodology for patent matching is universal and can be extended to other therapeutic markets and potentially the entire pharmaceutical industry with sufficient resources.

targets spur and which stifle innovation. Using the insights on the importance of both the *technology* and *product market* dimensions from Chapter 2, this paper looks at the position of targets and acquirers in existing and future product markets and also at the position of the parties in the technology space as key drivers of the M&A effects. Assessing impacts on projects of the acquirers, projects of the targets, and all projects combined, we show that M&As harm innovation on average, but the direction of the effects ranges from negative to positive, depending on the positions of the parties in both the product markets and the technology space. Negative cases are primarily linked to the absence of the acquirer's technological competence in areas where the target projects are developed. Positive cases are rare and only happen when large product market incumbents acquire technologically close projects in markets where they already operate, allowing the exploitation of technology synergies.

Venture capital

Rather than looking at the relationship between *competition* and *innovation* as such, the second topic explores another angle and investigates how venture capital funds shape the *direction of innovation* depending on *competition*. Venture capital funds are professional investors buying minority stakes in startups. Nevertheless, VCs are more than traditional financial intermediaries. Beyond liquidity, they also provide strategic guidance, are often actively involved in the management and thus play a prominent role in the early stages of startup's lives (Gompers *et al.*, 2020).

A substantial body of research has established a positive relationship between the activities of venture capitalists and innovation outcomes (see e.g. Kortum & Lerner, 2001; Bernstein *et al.*, 2016; Puri & Zarutskie, 2012). However, research remains largely silent on what is happening with innovation *inside* of firms. In other words, no empirical evidence exists on whether venture capital is likely to drive innovation in a specific direction, depending on competition from other players. Chapter 4 (*"Away from competition, away from 'defeat zone': VCs and R&D"*, single-authored) fills this gap and contributes to the literature by providing the first empirical evidence on how venture capitalists actively steer the early stage innovation activities of the startups they have invested in. VCs' behaviour follows a specific pattern, as VCs avoid *"defeat zone"* - an area around big incumbent firms in the product, R&D, and technology spaces where big incumbent firms already operate and where it is not worth competing. Instead of these projects, VCs pursue breakthrough projects, particularly in markets where they do not face product competition.

Chapter 2

Acquiring innovation: who, when, and what? [1]

Abstract. This paper analyzes patterns in M&As in the pharmaceutical industry involving drugs under development. Drawing on a detailed dataset of all corporate R&D activities related to antidiabetics over the period 1997-2017, we study the identity of targets and acquirers (who), timing of acquisitions (when) and which type of R&D projects change hands in terms of their technological novelty (what). Conversely to the narrative portraying large incumbent firms as typical acquirers and small firms as targets, we find that the majority of M&A activity takes place between small and research-focused firms. Further, it is also small and research-focused firms that engage in transactions involving novel or "high risk/high gain" projects. These potentially disruptive but highly uncertain projects are most likely to be acquired soon after their initiation, pointing towards an important yet so far unexplored pattern in M&A activity. By contrast, the largest incumbents in antidiabetics are less active acquirers, opting instead to conduct R&D in-house. Our findings have implications for where the scope of antitrust inquiry should be broadened when assessing the effects of M&As on innovation and competition in the pharmaceutical industry.

[1] This chapter is co-authored with Melissa Newham (ETH Zurich & KU Leuven), Jo Seldeslachts (KU Leuven & DIW Berlin), and Reinhilde Veugelers (KU Leuven). We thank seminar participants at KU Leuven, DIW Berlin, Bruegel, CET of DG COMP, MACCI 2022 and CISS 2019 summer school for valuable comments and suggestions. We are grateful to Ekaterina Khmelnitskaya, Manuel Gigena, Dennis Verhoeven, and Sam Arts for sharing data with us. Jan Malek acknowledges support from FWO through project 3H190094.

2.1 Introduction

Pharmaceutical companies are renowned for actively acquiring the research and development (R&D) portfolios of competitors. In fact, the value of pharmaceutical mergers and acquisitions (M&A) reached a record high of $360 billion in 2019 (Behner & Spence, 2020) - more than double the $170 billion that pharmaceutical companies invested in R&D (Deloitte, 2020). M&A deals in innovative sectors, such as pharma, have traditionally been seen as a means for firms to access new technologies or products (Wagner, 2011; Ornaghi, 2009b; Marco & Rausser, 2011; Yu *et al.*, 2016), replenish innovation pipelines (Higgins & Rodriguez, 2006; Danzon *et al.*, 2007; Arroyabe, 2021), or leverage synergies from combining innovation capabilities (Hoberg & Phillips, 2010; Bena & Li, 2014; Andersson & Xiao, 2016). Recent research has, however, highlighted that established firms with existing products ("incumbents") might have different motives. They may acquire promising or particularly novel innovations that have the potential to cannibalize existing profits to preemptively terminate their development - so-called "killer acquisitions" (Cunningham *et al.*, 2021; Federico *et al.*, 2020; Norbäck *et al.*, 2020). Whilst drawing a lot of attention, there is surprisingly little empirical evidence shedding light on such "killer" motives (Welch *et al.*, 2020).

To better understand the range of motives that underpin M&As in the pharmaceutical industry, this paper characterizes M&A in this sector along three dimensions: which companies are likely acquirers and targets (*who?*), at what stage of development do acquisitions take place (*when?*), and which type of projects, in terms of their novelty or "high-risk high-gain" nature, are likely to undergo ownership changes (*what?*), as well as interactions between these dimensions. Our analysis leverages detailed data on the innovation activities of firms in the market for antidiabetics. By focusing on one therapeutic market, we were able to connect numerous data sources and construct a unique database, covering all corporate activities in R&D related to antidiabetics at the project level over the period 1997-2017.[2] Whereas previous research has focused on the drivers of M&A for large firms,[3] our dataset also includes small and private firms which account for the lion's share of R&D undertaken but typically do not have any marketed drugs. Based on this detailed data, we aim to provide a more complete picture of the M&A landscape in the pharmaceutical industry.

In our analysis, we categorize firms both with regard to their overall market presence in the pharmaceutical industry (size) and their market presence in antidiabetics (incumbency). We link firms with their R&D projects and track the progress of projects over time as they pass through the typical development stages in the pharmaceutical industry. All projects start in preclinical development and then, if successful, pass through three phases

[2] We focus on the market for antidiabetic drugs but the framework we develop can be extended to other pharmaceutical markets.

[3] For example, Bena & Li (2014) assess the drivers of M&A using an economy-wide sample of large M&A deals between public firms.

of clinical trials. Uniquely, we link R&D projects to patents using an original algorithm that makes use of the (bio)chemical properties of projects. We make use of the project-patent link and information provided by patents to characterize the "high-risk high-gain" profile of all projects in our sample using the Novelty in Technological Origins (NTO) indicator developed by Verhoeven *et al.* (2016).[4]

We find that 50% of acquisitions occur in the earliest and the most uncertain development stage of preclinical development. The majority of M&A deals take place between small firms and purely research-focused pipeline firms, both of which do not have any marketed antidiabetic drugs and hence are "non-incumbents" in our framework. This finding alone is striking because much of the focus of previous studies on M&A in innovative industries has been on large companies with launched products. In general, we find that firms tend to acquire the projects of firms of similar or relatively smaller sizes. That is, small firms tend to acquire other small and pipeline firms, and pipeline firms primarily acquire other pipeline firms.

Beyond accounting for the majority of M&A deals, it is also the research-focused firms ("pipeline firms") that engage in transactions involving novel or "high risk/high gain" projects. In general, we find that "high risk/high gain" projects are significantly more likely to change hands. "High risk/high gain" projects typically originate from pipeline firms, and in particular pipeline firms that have some level of experience in research already ("mature pipeline firms"). These potentially disruptive, but highly uncertain, projects are most likely to be acquired by other mature pipeline firms at the very beginning of their development in the preclinical stage. This suggests that novel projects with high risks tend to be acquired soon after their initiation, pointing towards an important yet so far unexplored pattern in M&A activity.

In contrast to the small and pipeline firms which are highly active acquirers of R&D projects, we find that large pharmaceutical firms are less frequent acquirers in absolute terms. Big acquirers are also generally less likely to take on "high risk/high gain" projects. Although, one exception to this rule is a small group of fast-growing and risk-taking "star" firms, such as Gilead, which target the youngest pipeline firms and buy "high risk/high gain" projects. We find that the four market leaders in antidiabetics (Merck & Co., Eli Lilly, Sanofi, and Novo Nordisk) are substantially less likely to be acquirers despite being a large source of R&D in this market suggesting that these incumbent firms prefer to conduct innovation in-house. Big non-incumbent pharmaceutical firms (i.e. those without launched antidiabetic drugs but with substantial sales in other therapeutics) are most likely to acquire projects after they have progressed to the late development phase (phase 2 and phase 3 clinical trials) - and therefore at a time when there is much less uncertainty about

[4] A project is considered "high risk/high gain" if its associated patents draw on technological knowledge from domains that were previously not used in the technological domain of the invention. Technological novelty increases the variance of technological impact and the likelihood of being among the positive outliers concerning impact (Verhoeven *et al.*, 2016).

market outcomes. This suggests that larger firms without a presence in the antidiabetics product market use M&A as an entry strategy.

Our findings have important policy implications. The effects of M&A involving drugs under development and the potential for "killer" motives have recently caught the attention of policymakers.[5] Our results show that the transaction landscape is rich and varied, going beyond the narrative of large companies buying small companies to preempt future competition. While recent academic work has argued that dominant firms might be particularly willing to acquire disruptive firms with novel projects to prevent deterioration of existing profits (Federico *et al.*, 2020), our findings suggest that acquiring "high risk/high gain" projects in the early development stages is not a strategy that incumbents and big firms generally pursue. Our results provide evidence of significant M&A activity between small and research-focused firms and show that "high risk/high gain" projects tend to change hands between mature pipeline firms in the early development phases. Policymakers and academic researchers should broaden the scope of their inquiry to consider the implications of these transactions.

This paper contributes to the large body of literature that studies the characteristics and drivers of M&A in the pharmaceutical industry and innovative sectors in general. A major contribution of this study is that we are the first to link the M&A decisions of firms to the novelty of the drugs involved in the transaction.[6] In the biotech and pharmaceutical sector, several studies indicate that firms which are active in similar technological and therapeutic fields are more likely to merge,[7] however previous analyses are conducted at the firm level and do not consider the novelty of firms' projects. Further, we add to the literature that considers the role of timing, in terms of the developmental stage of projects, and M&A decisions in the pharmaceutical industry. Whereas Siebert & Tian (2020) and Grabowski & Kyle (2008) consider the stage of clinical development when studying the effects of M&As, we assess how timing influences M&A decisions in the first place. Finally, this paper contributes to and extends the M&A literature that explores the identities of targets and acquirers. While prior related research typically categorizes firms based on size alone (see for example Arroyabe, 2021; Bena & Li, 2014; Szücs, 2014; Danzon *et al.*,

[5] A recent antitrust case from Europe shows that enforcers are increasingly paying attention to the dynamics in the innovation space. In May 2021, the European Commission fined the company Merck KGAA EUR 7.5m for a failure to disclose an R&D project while pursuing the acquisition of Sigma-Aldrich in 2015 (European Commission, 2021). Another example of the growing interest of the enforcement agencies is the transatlantic "Multilateral Working Group on pharmaceutical mergers" launched in March 2021 by the US FTC, the European Commission and other competition authorities.

[6] There are only a few tangentially related studies that deal with different research questions. For example, Dranove *et al.* (2020) and Krieger *et al.* (2018) apply a measure of a drug's novelty to investigate how the incentives of firms to develop novel drugs respond to an external demand shock, represented by a policy change.

[7] For example, Marco & Rausser (2011) investigate the issue of "who merges with whom" in the biotech sector and find that matches are more likely when the patents of acquirers and targets are technologically similar. Relatedly, Meder (2016) finds that the probability to observe a merger between two firms is increasing both in product and pipeline portfolio proximity. Cunningham *et al.* (2021) find that acquisitions in the pharmaceutical industry are more likely when the acquirer has drug products and development projects that make use of the same "mechanism of action" and are in the same therapeutic field as the target.

2007), we further categorize firms based on their incumbency in the market of interest. We thus add to the emerging empirical research which places an emphasis on incumbents and the potential for "killer acquisitions"(see for example Cunningham *et al.*, 2021; Argentesi *et al.*, 2021; Gautier & Lamesch, 2021; Kamepalli *et al.*, 2020).

The rest of the paper is organized as follows. Section 2.2 presents the data and construction of variables. Section 2.3 describes the empirical implementation and section 2.4 presents the results. Lastly, section 2.5 concludes.

2.2 Data

2.2.1 Data sources and data construction

Our analysis relies on detailed data at the project level concerning project characteristics, ownership changes, progression through the stages of development, and patents. To create the required dataset we combined information from multiple sources. In the following section, we describe the main data sources and key steps in the data construction process.

2.2.1.1 Projects

The backbone of our dataset is the Pharmaprojects database from Citeline which provides a comprehensive list of global R&D activity in the pharmaceutical industry at the project level. This database draws on information from multiple sources including companies' press releases, media coverage, patent filings, conference proceedings, regulatory bodies' reports, medical literature as well as direct contact with company representatives and researchers. Several other papers have employed this database to investigate different questions, for example Adams & Brantner (2006); Kyle (2007); Blume-Kohout & Sood (2013); Branstetter *et al.* (2014); Cunningham *et al.* (2021). We use Pharmaprojects to identify all projects related to the treatment of diabetes and their MoA during our sample period (1997-2017).[8] A total of 2711 projects related to diabetes were identified.

2.2.1.2 Ownership changes

We identified changes of ownership for each project in our database by carefully unwinding the sequence of each project's consecutive owners. These ownership changes not only include mergers and acquisitions, but also deals involving sales of divisions, product lines, or individual assets. To do so, we used text information provided in the Pharmaprojects database along with text mining, algorithmic disambiguation, fuzzy string matching, and extensive manual checks.[9] To complement and verify that these changes indeed reflected

[8] In 1997 the Food and Drug Administration Modernization Act required firms to publish information on clinical trials in the registry.

[9] The following text represents a typical example of a project's description available in the database: "BVT-933 (PRX-00933) is an oral 5-HT2C agonist, which was under development by Upsher-Smith (Proximagen before the acquisition; Proximagen Neuroscience before the name change) for the treatment

ownership changes as opposed to name changes, we matched relevant firms with the merger databases Zephyr and SDC Platinum. Manual checks and additional desktop searches for every company were performed to ensure correctness and completeness.

2.2.1.3 Progression though development

We complemented the information on progression through (pre)clinical development in the Pharmaprojects database with additional information from the AACT database.[10] This database lists every study registered at ClinicalTrials.gov - a repository of privately (and publicly) funded clinical studies conducted around the world. We matched studies to projects using fuzzy string matching on sponsor and drug names. This allowed us to identify the development phases each project had passed, along with the start and end dates of those phases.[11] In this way, we re-constructed development histories for 2378 projects (88%) of projects.[12]

To determine whether a project has been successfully launched, we restrict our attention to the US market and check whether the drug appears in the FDA Orange Book. The reason for this is that new drugs are typically launched in the US first, given that it is the largest market accounting for 40% of global pharmaceutical sales (IFPMA, 2017).

2.2.1.4 Patents

The most challenging part of the data construction was matching projects to patents. Matching patents to projects is a complex many-to-many matching problem as one project is typically linked to many patents, and one patent may be relevant to multiple projects. To the best of our knowledge, the only existing source that links patents to projects in the pharmaceutical industry is the FDA Orange Book. However, this database only provides information on drug projects that successfully make it to market.

To identify the patents relating to projects that were abandoned during development we drew on information from multiple sources and developed a novel algorithm which uses the (bio)chemical and pharmacological properties of projects to assign patents. This allowed us to consider not only chemical-based (Krieger *et al.*, 2017) or biological-based drugs (Sampat & Williams, 2019), but combine both approaches.

In brief, for each project in development, we searched for patents that were filed at the USPTO by the project's developers between the inception and termination dates of a

of obesity, diabetes and glaucoma. Biovitrum (now Swedish Orphan Biovitrum) divested a 5-HT2C agonist programme, which included BVT-933, to Proximagen Neuroscience."

[10] Available: https://aact.ctti-clinicaltrials.org/

[11] In cases where complete histories could not be established (for example the date of exit from a phase was missing), we imputed the missing dates by estimating the log-normal distribution of durations per phase and randomly drew a project's phase duration from the estimated distribution. For each such imputation, we have manually checked that the sequence of development milestones was not violated.

[12] The remaining 12% projects were not matched due to insufficient information. For some projects, the Pharmaprojects database did not provide sufficient details and links to the trials registry could not have been established. For others, complete development histories could have not been established even after the imputation procedure.

project.[13] Depending on the type of project and information available, we applied various matching techniques. For small-molecule chemical drugs, we employed several crossroads between chemical, patent, and medical databases to establish project-patent links. For large molecule drugs relying on proteins, we followed the approach of Sampat & Williams (2019) and linked gene identifiers from the Pharmaprojects database to a list of protein and nucleotide sequences and then matched these sequences against the census of sequences disclosed in the US patents to establish patent links. To complement these approaches and increase the matching rates, we also use natural language processing methods and data from Arts *et al.* (2021) to connect projects to patents based on keywords relating to their MoA. A more detailed description of the algorithm and its results are given in Section 2.C.

Combined with additional manual checks of the unmatched entries, this algorithm allowed us to match patents to 1877 projects (79%), representing our final sample.

2.2.2 Construction of key variables

Leveraging the database described above, we construct variables that are central to our analysis. In particular, we develop a categorization of firm types, we use the project-patent link to characterize how novel ("high risk/high gain") projects are, and distinguish between the phase in which projects are in. The construction of these key variables is described in greater detail below.

2.2.2.1 Firm types

The database comprises over 900 different pharmaceutical companies. These firms are very heterogeneous in terms of their size, previous experience in R&D (both in diabetes and other therapeutic fields), and their market presence (in terms of whether they have launched drugs on the market). To understand which types of firms are engaging in which kinds of transactions, we cluster firms into bins ("firm types").

Our categorization delineates between firms based on two main criteria: (i) size (market presence in the pharmaceutical industry) and (ii) incumbency (market presence in antidiabetics). Size matters in the context of M&A decisions as larger firms have deeper financial pockets and can benefit from economies of scale and scope when it comes to clinical trials, obtaining regulatory approval, production and commercialization (Arroyabe, 2021; Bena & Li, 2014; Szücs, 2014; Danzon *et al.*, 2007). Incumbency, on the other hand, may impact acquisition motives as incumbents will have different incentives in comparison to non-incumbents related to defending or expanding their existing position (Argentesi *et al.*, 2021; Cunningham *et al.*, 2021; Federico *et al.*, 2020).[14] Our aim is to assign firms

[13] In our analysis we link projects to US patents. Although the scope of our R&D projects is global, the geographic product market of interest is the US. To protect intellectual property rights and prevent loss of exclusivity upon launch, projects must be protected by a valid US patent.

[14] Whereas the size-based categorization has a traditional standing in the literature, the focus on incumbency has recently grown given new theoretical (see for example Bryan & Hovenkamp, 2020; Letina

to bins such that firms in the same bin have similar capabilities, financial resources, and incentives to engage in M&A.

Along the size dimension, we distinguish between big firms, small firms, stars, and young and mature pipeline firms. Big firms have a stable market share in the pharmaceutical industry of more than 1%.[15] This group includes firms typically thought of as "big pharma" such as Johnson & Johnson and Pfizer. Small firms have at least one launched product in any pharmaceutical market, but less than 1% market share over their entire lifetime.[16] A few firms in the pharmaceutical industry have grown rapidly over the last 20 years (e.g. Gilead, Teva). We separate these fast-growing companies and label these as "stars". The last group consists of all firms that do not have any launched drugs (in any pharmaceutical market) and thus are purely engaged in R&D activities: "pipeline" firms. Since this is a very large group, within which we cannot make further cuts based on revenues, we adopt a different approach: We use the filing date of a firm's first pharmaceutical patent to further separate pipeline firms into two more homogeneous groups. Young pipelines are all firms with a first patent filing less than 5 years ago and mature pipelines are all firms with a first patent filing more than 5 years ago.[17] As such, the former group captures all the young ventures that have just entered the market, whereas the latter group captures firms that have already been working on pharmaceutical R&D projects for at least 5 years, and thus managed to collect experience or grow to a size of multiple R&D projects on average.

Concerning incumbency, we distinguish between incumbents, non-incumbents and leaders. Incumbents, as opposed to non-incumbents, have at least one antidiabetic drug on the market. "Leaders" are incumbents with significant market share in antidiabetics, exceeding 10% over the entire sample period. Only four firms traditionally dominating the antidiabetics market meet this definition - Novo Nordisk, Sanofi, Merck & Co., and Eli Lilly. Due to the specific structure of the antidiabetics market, we separate these stable "leaders" from the other 17 incumbents in our sample.

Table 2.1 presents the resulting matrix from combining the two dimensions, populating eight bins: (i) leaders, (ii) big incumbents, (iii) big non-incumbents, (iv) stars, (v)

et al., 2020; Norbäck *et al.*, 2020; Haucap *et al.*, 2019; Gilbert, 2019; Federico *et al.*, 2018, 2017) as well as empirical research identifying potentially anticompetitive outcomes in acquisitions, particularly in the pharmaceutical industry (Cunningham *et al.*, 2021) and the tech industry (Argentesi *et al.*, 2021; Gautier & Lamesch, 2021; Kamepalli *et al.*, 2020). By incorporating incumbency as a dimension, our categorization is more fine-grained than previous studies which focus only on size (typically big vs. small) (see Arroyabe, 2021; Bena & Li, 2014; Szücs, 2014; Danzon *et al.*, 2007).

[15] We compute market shares using the R&D Scoreboard data published by the European Commission. Every year since 2003, the Commission has published the list of the largest firms, categorized by sectors, together with their revenues, R&D spending, and other information. The maximum market share of any given pharmaceutical firm has never exceeded 8%.

[16] We choose the threshold of 1% since it provides us with a consistent means to isolate sufficiently large firms, which are often referred to as "big pharma" - for example, Pfizer, Roche, or Johnson & Johnson.

[17] By construction, after a passage of 5 years, a young pipeline firm can switch to a mature pipeline bin. A total of 159 of our firms do so. Given that this is only based on the passage of time (and eventually the age of the firm) and is unrelated to firms' activities, such switching should be exogenous. We also adopted an alternative threshold of 3 years. Results remain robust to this definition.

small incumbents, (vi) small non-incumbents and (vii) mature pipelines and (viii) young pipelines. Figure 2.B.5 in Appendix 2.B gives examples of firms belonging to every bin.

Table 2.1: Firm types

		Incumbency		
		Non-incumbent	**Incumbent**	**Leader**
Size	**Big**	*Big non-inc* (N=19)	*Big inc* (N=11)	*Leader* (N=7)
	Star	*Stars* (N=8)	-	-
	Small	*Small non-inc* (N=195)	*Small inc* (N=10)	-
	Pipeline (mature)	*Mature pip* (N=300)	-	-
	Pipeline (young)	*Young pip* (N=408)	-	-

Note: At any point in time, each firm belongs to a single bin, but can change bins over time. Counts of the number of firms in each bin are in parentheses. Definitions are as follows; **Leader**: market share in antidiabetics above 10%. **Big incumbents**: at least one launched antidiabetic drug and average market share in pharma above 1%. **Big non-incumbents**: no launched antidiabetic and average market share in pharma above 1%. **Stars**: market share in pharma below 0.75% on entry to sample and above 1% at sample end (no star has ever launched an antidiabetic). **Small incumbents**: none of the above, at least one launched drug in any pharma market and antidiabetics. **Small non-incumbents**: none of the above, no launched antidiabetic, and at least one launched drug in any pharma market. **Mature pipelines**: no launched drugs and filed first pharma patent more than 5 years ago. **Young pipelines**: no launched drugs and filed first pharma patent less than 5 years ago.

2.2.2.2 Novelty of projects

The projects in our sample are likely to differ widely in terms of the involved risks and potential impact - or their "high risk/high gain" nature. Indeed, while the "high risk/high gain" inventions have the potential for high impact and can disrupt established industries (Christensen, 2013), they are also inherently more uncertain with respect to their technological and commercial performance (Fleming, 2001; Hall & Lerner, 2010; Verhoeven *et al.*, 2016).[18]

To understand which types of projects are being acquired and differentiate between the risks and potential of these projects, we characterize projects as "high risk/high gain" based on patent information available for each project and the Novelty in Technological Origins (NTO) indicator developed by Verhoeven *et al.* (2016).[19] This indicator measures the ex-ante technological novelty of patents by assessing the extent to which a patent sources knowledge of previously unconnected fields. In practice, a patent scores on NTO

[18] Krieger *et al.* (2018) show that more novel innovation is riskier and novel drug candidates are less likely to be approved by the FDA, but conditional on approval, novel drugs are more valuable and earn higher revenues. Foster *et al.* (2015) then show that research introducing new combinations of chemicals is more likely to become highly cited but also displays a higher variance in their citations, confirming the high risk and high gain character.

[19] Alternative measures of drugs' novelty were used in the literature. Dranove *et al.* (2020) measure the novelty of drugs using a count of previous deployments of the drugs' mechanism of action. Krieger *et al.* (2018) restrict themselves to small molecule drugs and define a drug as novel if it is molecularly distinct from prior candidates. However, patents provide more comprehensive information and allow us more precisely to measure the "high risk/high gain" potential connected to their underlying technology.

if it makes a combination between its own IPC code (eg. A61K31) and an IPC code from its referenced patents that have not yet occurred in the years previous to the application year of the patent (Verhoeven *et al.*, 2016).[20] Following the advice of industry experts suggesting that all important patents are filed before a project enters clinical testing, we only use patents assigned before entering clinical trials and before an event (for the treated projects) to define novelty. Since in our setting a project can have one or more assigned patents (3.8 patents on average with a median of 1) and NTO is measured on the patent level, we aggregate at the project level and consider a project to be "high risk/high gain" if at least one of project's patent scores on the NTO indicator.[21]

An example of a "high risk/high gain" project in our sample is the first inhalable insulin. As opposed to commonly used injectable insulin, inhalable insulin was thought to be a significant disruptor with substantial revenue potential, allowing painless administration. In total, 1060 projects in our sample are assigned NTO status and thus meet the criteria to be considered "high risk/high gain".

2.2.2.3 Phase of projects

We observe the progression of all projects through the stages of development. To analyze the timing dimension of acquisitions, we aggregate the dataset to the project-phase level. We distinguish between preclinical, early (Phase I), and late (Phase II and Phase III) development phases. We group Phase II and III as both are aimed at testing the safety and efficacy profile of a drug and are frequently run in parallel with each other. A more technical reason for this grouping also relates to the fact that we do not have enough data points for proper econometric analysis on separate phases, especially for Phase III. The launched phase is not considered in this paper since it does not represent an R&D phase.

2.2.3 Final database

Each project adds one observation per phase to our database. The values of variables for each observation are measured at the beginning of the phase, except when a project has experienced an ownership change in a phase. In this case, we measure the values of all variables one semester before the transaction. A project is dropped from our database after the phase in which it was discontinued. Depending on how far a project has progressed, it can thus contribute at most three observations (Preclinical, Early, and Late).[22] In

[20] Verhoeven *et al.* (2016) perform series of analyses to verify that the NTO measure correlates with other existing constructs (eg. "originality" of Trajtenberg *et al.* (1997) or "radicalness" of Shane (2001)) but performs better on characteristics typical for technological novelty. They also analyze the technological impact generated by NTO inventions and find that such inventions have indeed a higher dispersion in terms of forward citations received, and are more likely to end up among the set of highly cited patents, consistent with their "high risk/high gain" nature.

[21] Since the probability to score on novelty mechanically increases with the rising number of assigned patents, we always control for the number of project's assigned patents in the regression analysis.

[22] However, should more than one ownership change occur during a phase, each of these represents a separate observation - in which case the total number of observations per project can exceed three.

Appendix 2.D, we provide an illustrative example of how all dimensions of the data and the variables defined in the previous sections connect.

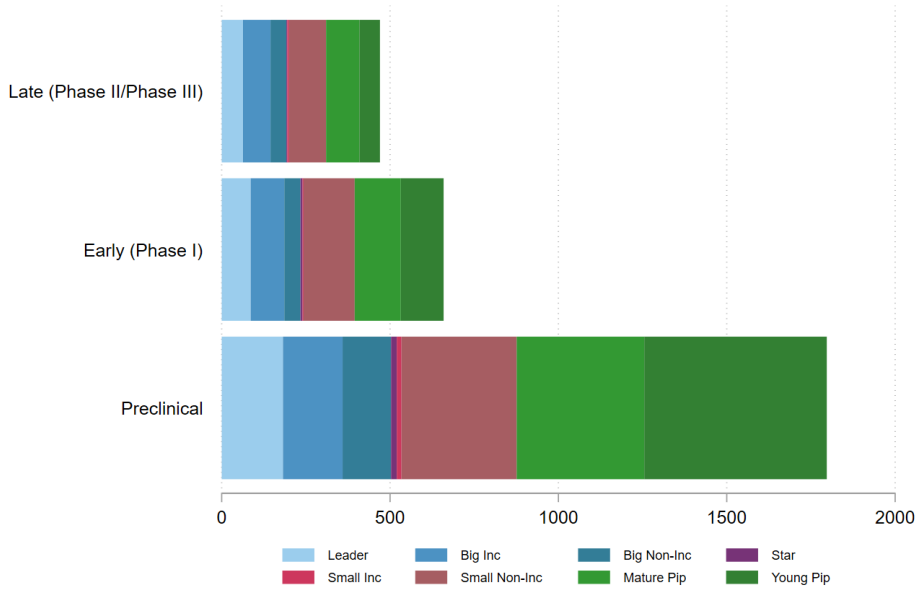
The final database amounts to 2926 project-phase observations relating to 1864 projects. There are 196 ownership changes relating to 181 underlying projects.[23] This represents 7% of observations and 10% of projects. The ownership changes include mergers and acquisitions, which affect all projects of the target company. In addition, ownership changes might only involve a portion of the target's projects, for example when a company acquires a division or specific research programs.[24] Last, ownership changes might involve corporate spin-offs in which a company (the target) separates one or more projects into a new, separate entity (acquirer).

Figure 2.1 first presents the funnel structure of the antidiabetics R&D. The sample contains 1797 preclinical projects. As projects progress, their number gradually reduces. Only 659 projects moved to Phase I and 470 progressed beyond phase I. Table 2.2 then presents key summary statistics by various splits of the sample, separately for the observations with an ownership change ("Event") in a phase and without an event in a phase ("No event"). The last column presents a p-value of the mean difference between the two groups.

[23] Considering the launched project phases would increase the number of observations to 3009. The number of events would increase to 209, relating to 190 projects. These samples are only used for some sensitivity checks in the results section.

[24] A major example of such transaction is AstraZeneca's acquisition of the diabetes business from Bristol-Myers Squibb in 2014 by which BMS has exited the diabetes business and refocused on other therapeutic areas (Mullard, 2014).

Figure 2.1: R&D Funnel



Note: The figure shows the funnel structure of antidiabetics R&D. Each bar shows the number of projects developed in each phase by the type of firms.

Table 2.2: Summary statistics

	No event Count	Mean	Event Count	Mean	Difference p-value
Timing:					
Preclinical	1679	0.62	118	0.54	0.03
Early	607	0.22	52	0.24	0.61
Late	423	0.16	47	0.22	0.04
Acquirer:					
Leader	315	0.12	16	0.07	0.02
Big inc	331	0.12	28	0.13	0.77
Big non-inc	209	0.08	30	0.14	0.01
Star	16	0.01	10	0.05	0.01
Small inc	21	0.01	1	0.00	0.52
Small non-inc	546	0.20	58	0.27	0.04
Mature pip	567	0.21	48	0.22	0.69
Young pip	704	0.26	26	0.12	0.00
Target:					
Leader	315	0.12	10	0.05	0.00
Big inc	331	0.12	4	0.02	0.00
Big non-inc	209	0.08	13	0.06	0.31
Star	16	0.01	1	0.00	0.79
Small inc	21	0.01	3	0.01	0.46
Small non-inc	546	0.20	65	0.30	0.00
Mature pip	567	0.21	73	0.34	0.00
Young pip	704	0.26	48	0.22	0.19
Project characteristics:					
NTO	1060	0.39	117	0.54	0.00
Observations	2709	2709	217	217	2926

2.3 Empirical implementation

Our analysis is centred around descriptive statistics and a simple linear probability model isolating the key drivers of M&As in the antidiabetics market when controlling for various company characteristics, project characteristics, and fixed effects. The generic regression we estimate has the following form:

$$Prob(Dep.Var_{it} = 1) = \alpha + \beta X_{it} + \gamma FE + \epsilon_{it} \tag{2.1}$$

We estimate these regressions on two different samples. The first sample consists of all project-phase observations, encompassing projects with and without ownership changes. This allows us to explore the different drivers of the likelihood that a specific project in a given phase experiences an ownership change. More specifically, we utilize the variables defined above and analyze which companies are likely acquirers or targets (who), which type of projects undergo ownership changes in terms of the technological novelty (what) and in which phase (when). The second sample is restricted to project-phase observations relating to ownership changes only. This allows a deeper exploration of the transaction patterns between acquirers and targets (who) and links them to the *when* and *what* dimensions.

The dependent variable in each analysis is a binary indicator whose definition differs depending on the question at hand. For example, when analyzing the likelihood that a specific project experiences an ownership change (eg. who acquires?, who sells?) the dependent variable is a binary indicator that equals to one for a project (i) and phase (t) affected by an ownership change and zero otherwise. In other cases, the dependent variable depends on the question analyzed in a particular regression analysis. The footnotes accompanying each set of results provide an exact definition of the dependent variable in each regression.

The independent variables included in the vector X depend on the question that is examined. For example, to examine which types of acquirers (targets) are likely to buy (sell) projects, X contains binary indicators for the acquirer (target) bins. To study which projects undergo ownership changes, X contains the binary NTO indicator. In all instances, the β coefficients are then of interest, as these indicate which variables are positively or negatively associated with the likelihood of an ownership change.

The vector FE includes several groups of fixed effects which filter out static differences along several dimensions relevant to the pharmaceutical R&D. First, the cohort fixed effects group together projects initiated around the same time. This controls for time trends and/or technological trends. Ideally, we would control for cohorts depending on the year in which development of a particular drug was initiated. However, due to the sample size, projects were aggregated into 7 cohorts, depending on a 3-year window in which the project was initiated (the first cohort groups project was initiated between 1997 and 1999,

the second cohort groups project was initiated between 2000 and 2002, and so forth).

The second set of fixed effects relates to the mechanism of action of projects. Mechanism of action is a distinctive feature of drugs and determines how a drug produces its effect in the body. MoAs are also closely linked to the types of side effects (Berger & Iyengar, 2011), and suitability for treatment in different patient populations (Association *et al.*, 2019; Chaudhury *et al.*, 2017). From the development perspective, significant heterogeneities exist between various MoAs in the development life cycle, underlying science, success rates and market launch, and the extent of development activity within MoA.[25] The number of distinct MoAs in the dataset amounts to 389 and exceeds the number of treated observations, making it impossible to include separate fixed effects for each MoA. We identified large enough MoAs (with at least 30 projects in development), included separate fixed effects for these and grouped all the remaining MoAs into one category. We also created a separate MoA fixed effect for the drugs with unknown MoA.

The third set of fixed effects relates to fixed differences between the technological areas where projects were developed, as indicated by patents assigned to projects. These fixed effects are defined based on IPC groups of the assigned patents. Each project scores on as many IPC group indicators as the underlying patents refer to. Similarly to the MoA, since the number of IPC groups exceeds the number of treated observations, we created separate fixed effects for the most populated technological areas (more than 100 projects) and aggregated the rest into one category.

The last set of fixed effects relates to a geographical area where companies operate.[26] We introduce fixed effects for four regions depending on whether the company's headquarters is located in South-East Asia (Japan, China, India, Singapore and Taiwan), Europe, Northern America (Canada and the US), or the rest of the world.

Given the substantial number of fixed effects that our analysis requires, we resort to ordinary least squares (OLS) as the primary estimation method. Whilst OLS is by construction better suited to accommodate a large number of fixed effects and can also estimate coefficients of groups where every member of the group has the same value for the dependent variable (Caudill *et al.*, 1988), it is conceptually less well suited for settings with dichotomous outcomes. The main issue is that the model assumes a linear functional form and the range of outcomes predicted by the OLS model is unrestricted, resulting possibly in predicted probabilities outside of the [0,1] probability domain. To test the robustness of the results based on the OLS model, we also report analyses in Appendix 2.A employing

[25] In our sample, only 18 of 389 MoAs were launched. Also the distribution of development activity is highly skewed and the largest 7 MoAs account for almost half of the development. Dranove *et al.* (2020) find similar results, more specifically that the majority of known MoAs are rarely launched (five or fewer times) while a small minority accumulates a large number of launches.

[26] For example, Yeo (2013) show that geographical distance harms takeover flows. Ornaghi (2009b) finds that companies with a smaller cultural and geographic distance have a higher probability of agreeing on a M&A because there is little chance of cultural conflicts. Hsu *et al.* (2021) finds that innovative firms in low-innovation countries are more likely to undertake cross-border deals and select innovative targets.

a logistic regression in cases where the logistic model can be estimated.[27]

A valid statistical inference depends on the treatment of the standard errors. Due to the strong path-dependence of the pharmaceutical R&D process, it is unlikely that individual project-level error terms would be independent. As a baseline, our analysis thus uses standard errors clustered at the level of projects. This allows us to explicitly take into consideration the correlation between errors relating to the same project but different development phases and adjust the standard error accordingly. However, this still assumes that clusters (projects) are independent. One obvious caveat is that, apart from the few partial M&As where assets or divisions of a firm are changing hands, ownership changes affect the whole firm. As such, the treatment (ownership change) is assigned at the level of firms, rather than individual projects. With this empirical design, clustering at the level of firms rather than projects seems plausible - since this allows the error terms of projects within the same firms (clusters) to be correlated but independent across different firms (Abadie *et al.*, 2017). Robustness analyses with clustering at the firm level for all key results are included in Appendix 2.A.

2.4 Results

The sections below present descriptive evidence and the results of our regression analyses. First, we examine which firms are the most active acquirers and which firms are selling projects the most often (2.4.1). Conditioning on a project changing ownership, we continue by investigating the interplay between acquirers and targets - more specifically we test whether some acquirers are more likely to acquire certain targets (2.4.2). Section 2.4.3 looks at the role of timing in the takeovers of projects. Section 2.4.4 investigates the role of project characteristics, more specifically whether the high risk/high gain projects are more attractive for acquisitions and when firms transact over them. Section 2.4.5 analyzes the types of firms which buy or sell high risk/high gain projects.

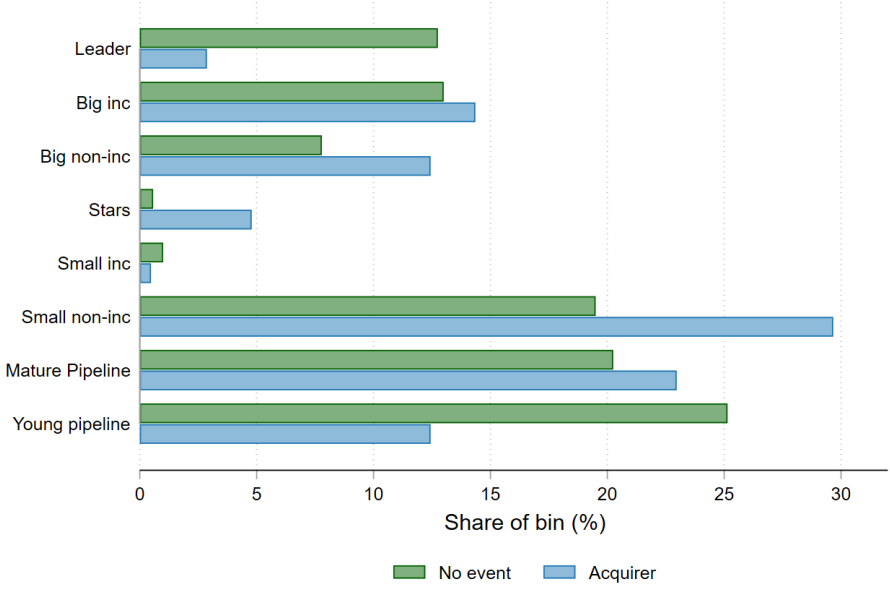
2.4.1 Who acquires and who sells?

Figure 2.2 shows the activity of the different types of acquiring firms. In particular, the figure compares the distribution of various types of firms in the sample where no ownership changes occurred (green) to the sample with ownership changes (blue). The figure indicates that the most frequent acquirers were small and mature pipeline firms - accounting for more than 50% (107) of all ownership changes. Although less in absolute numbers, we can also observe a significant M&A activity by big firms - which is almost equally split

[27] When we analyze the sample restricted to the set of ownership changes only (196 observations), the sample size is too small and the number of fixed effects is too large to estimate the logistic model. In addition, in small samples, logit model coefficients can have a substantial bias away from zero (Rainey & McCaskey, 2021). In these cases, the logit specifications are not reported. Probit results are not reported since they yield results identical to the logit specification.

between big incumbents (14% - 25 ownership changes) and big non-incumbents (13% - 22 ownership changes).

Figure 2.2: Who acquires?



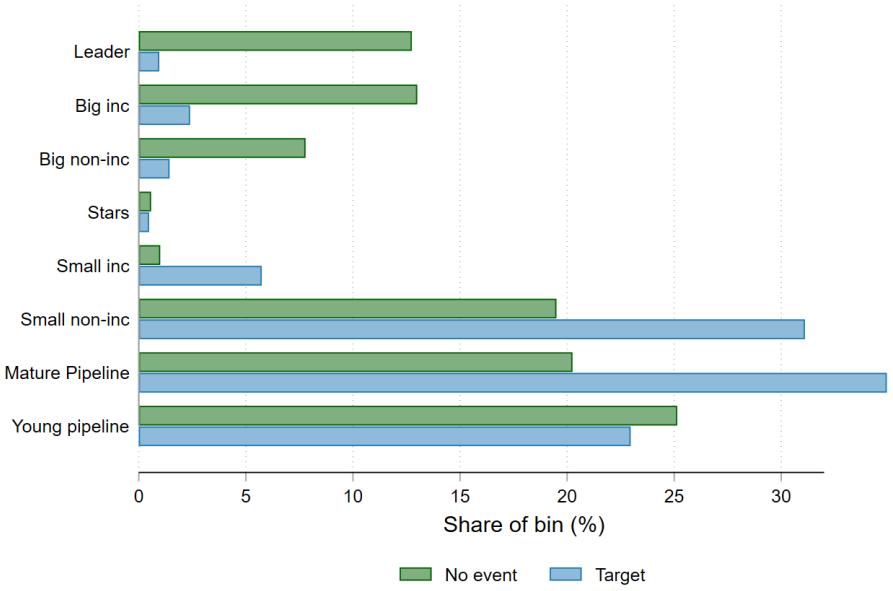
Note: The figure shows the distribution of observations in the "No Event" sample (green) and "Event" sample (blue) split by acquirer bins. Bars of the same colour always sum up to 100%.

While being responsible for a relatively large part of in-house R&D (12% of the "no-event" sample), the four industry leaders - Novo Nordisk, Eli Lilly, Sanofi, and Merck & Co. - accounted only for 3% (6) of acquisitions and seldom offered projects for sale. Hence, interestingly, leaders were under-represented as acquirers and seemed to primarily rely on their own R&D. In contrast, despite being a very small group representing 1% of internal R&D, star firms seemed to be most overrepresented as acquirers as they engaged in almost 5% of all ownership changes (10).

Analogically to Figure 2.2 , Figure 2.3 shows the distribution of the different types of firms among targets. The most typical targets were generally small and pipeline firms, particularly then the small non-incumbent firms and mature pipeline firms - each of them selling projects in more than 30% of the transactions. Compared to the "no-event" sample, these two types of targets seem to be also the most overrepresented among the firms selling projects. In contrast to this, big firms (leaders, big incumbents as well as non-incumbents) were rarely selling their projects - and were also substantially underrepresented as targets.

The table 2.3 test these descriptive patterns in a regression framework. Given the low

Figure 2.3: Who sells?



Note: The figure shows the distribution of observations in the "No Event" sample (green) and "Event" sample (blue) split by acquirer bins. Bars of the same colour always sum up to 100%.

number of observations for some groups (for example small incumbents acquired only 1 project), we aggregate bins together to allow for statistical analysis. The relevant grouping and base group are always indicated in the regressions. In line with the descriptive analysis, the first column shows that compared to big incumbents, leaders are significantly less likely to acquire projects (at 1% significance level), whereas stars appear as most likely and aggressive acquirers (at 1% significance level).[28] Importantly, the regression results also indicate that compared to big incumbents - who are often stylized as prototypes of acquirers (Cunningham *et al.*, 2021; Argentesi *et al.*, 2021; Gautier & Lamesch, 2021) - small and mature pipeline firms are not less likely to acquire projects. As regards the targets, the regression in column 2 reveals that small, mature pipeline and young pipeline firms are significantly more likely (at 1% level) to sell projects in M&A transactions.

Table 2.3: Who acquires and who sells?

	(1) Acquirers	(2) Targets
Leader	-0.071***	-0.007
	(0.018)	(0.009)
Big non-inc	0.017	0.014
	(0.024)	(0.011)
Star	0.283***	0.020
	(0.104)	(0.050)
Small	0.015	0.121***
	(0.020)	(0.015)
Mature pip	-0.014	0.113***
	(0.020)	(0.015)
Young pip	-0.056***	0.067***
	(0.019)	(0.012)
Cohort FE	Yes	Yes
MoA FE	Yes	Yes
Country FE	Yes	Yes
Obs	2926	2926
Adj. R2	0.039	0.048
Base	Big-inc	Big-inc

Note: This table presents the OLS estimates of the likelihood to acquire projects (column 1) or to sell projects (column 2) using the project-phase sample. The dependent variable is a binary indicator equal to one for observations i experiencing an event in phase t (treated) and zero otherwise. The independent variables are binary indicators for membership in the acquirer bin (column 1) or the target bin (column 2). Fixed effects capture membership in cohorts (Cohort FE), grouped mechanism of action FE (MoA FE) and region FE (Country FE). Errors are clustered at the project level and displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Appendix 2.A presents several robustness checks and extensions. More specifically, table 2.A.12 shows that the results hold when considering big mergers, excluding partial

[28] We pick the base group as big incumbents since these firms are generally not overrepresented as acquirers and it is also the most natural group for bringing insights in the context of the current literature.

M&As and spin-offs, and also when considering transactions happening in the product markets once projects have been launched. Table 2.A.13 shows that the findings are also fully robust when considering a logistic regression and clustering of standard errors at the firm level.

2.4.2 Who acquires whom?

After establishing which firms are the most likely acquirers of projects compared to never acquired projects, this section focuses only on the sample of the 196 ownership changes and analyzes the transaction dynamics between the acquirers and targets.[29]

Figure 2.4: Transaction between targets and acquirers

Note: The figure shows a matching between different types of acquirers and targets. The numbers represent counts of ownership changes, the heat map then visualizes the observed frequencies. The total number of ownership changes amounts to 196.

Figure 2.4 shows the matching between acquirers and targets. Small and pipeline acquirers were acquiring projects of firms of similar or smaller types - in particular, small firms acquired projects of other small firms in 50% (=29/58) of their acquisitions, with the

[29] If a project was not affected by an ownership change in a given period, the identity of the acquirer and the target in the dataset coincides. A pairing analysis must be therefore restricted to the treated sample only.

rest targeting mature pipelines (23). Mature pipelines then acquired over 60% (=30/48) of their projects from other mature pipelines and the rest from the young pipelines (13). Thus, a large chunk (33%) of the overall M&A activity in the antidiabetics market takes place between purely innovating companies that do not sell products (yet).

As regards the acquisitions by big firms, the figure reveals that important differences exist between the strategies depending on incumbency - while the incumbents (including leaders) targeted projects of the youngest ventures in almost 50% (=15/31) of their transactions and small non-incumbents in 36% (=9/25) of their transactions, big non-incumbents focused largely, in 75% (=17/22) of cases, on takeovers of projects belonging to more established small firms. Finally, despite being a small acquirer group, stars have targeted primarily pipeline firms (70% (=7/10)), especially the young ones (50% (=5/10)).

Table 2.4: Who acquires whom?

	(1) T is pip	(2) T is young pip	(3) T is young pip
Leader + big inc	0.279**	0.430*	0.264**
	(0.136)	(0.242)	(0.115)
Star	0.494**	0.336	0.337*
	(0.216)	(0.307)	(0.192)
Small	0.344***	-0.205	-0.005
	(0.112)	(0.221)	(0.075)
Mature pip	0.642***	-0.066	0.095
	(0.112)	(0.219)	(0.092)
Young pip	0.570***	0.116	0.230*
	(0.131)	(0.232)	(0.119)
Cohort FE	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes
Country FE A+T	Yes	Yes	Yes
Obs	189	121	196
Adj. R2	0.230	0.225	0.191
Base	A Big-noninc	A Big-noninc	A Big-noninc
Target Sample	Small + pip	Pipeline	All

Note: This table presents the OLS estimates of the likelihood to acquire projects of a particular target type. In column (1), the estimation sample is restricted to the ownership changes of small and pipeline targets. The dependent variable is equal to one if the target was a pipeline firm and zero if the target was a small firm. In column (2), the estimation sample is restricted to the ownership changes of pipeline targets. The dependent variable is equal to one if the target was a young pipeline firm and zero if the target was a mature pipeline firm. In column (3), the estimation sample is unrestricted and encompasses all ownership changes. The dependent variable is equal to one if the target was a young pipeline firm and zero otherwise. Fixed effects capture membership in cohorts (Cohort FE), grouped mechanism of action (MoA FE) and region (Country FE) for both acquirers and targets. Errors are clustered at a project level and displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 2.4 reports the results from regressions where we estimate the likelihood that projects of the most frequent types of targets are taken over by specific types of acquirers. Due to the low number of observations in the leader group (6), we cannot reliably estimate the coefficient for this group alone. Therefore, we aggregate the leaders and big

incumbents together.[30] In column (1), we restrict the sample to small and pipeline targets and compare which acquirers are more likely to acquire projects of pipeline firms as opposed to small firms. The results confirm the above pattern and indicate that only big non-incumbents (base group) are significantly more likely to acquire small firms, whereas other firms target pipeline firms. Column (2) restricts the sample to pipeline firms and tests which acquirers are likely to take over projects of young pipelines as opposed to mature pipelines. The results confirm the above matching between incumbent acquirers and young pipeline targets. Lastly, Column (3) does not restrict the sample and test which acquirers are likely to acquire young pipelines compared to all other targets. The results reveal that in addition to incumbents, also stars are likely to acquire young pipeline targets. The robustness check in Appendix 2.A (Table 2.A.14) shows that these results remain robust when allowing for correlation of projects within the acquiring firm and clustering the standard errors at the firm, rather than at the project level.

2.4.3 When do the transactions happen?

Table 2.5 shows the distribution of acquirers' M&A activity over the phases of development by count and by the share that the given phase represents on the total of each acquirer's transactions. In the preclinical phase, where project outcomes are still extremely uncertain, pipeline companies and small companies accounted for the bulk of the action. At the same time, preclinical transactions were also what these companies engaged in the most often - small firms undertook 54% and mature pipelines even 70% of all their acquisitions in the preclinical phase.

Table 2.5: Who acquires when (summary statistics)?

Acquirer type	Preclinical	Early	Late	Total
Leader + big inc	14 (45%)	12 (39%)	5 (16%)	31 (100%)
Big non-inc	9 (41%)	3 (14%)	10 (45%)	22 (100%)
Stars	8 (80%)	1 (10%)	1 (10%)	10 (100%)
Small	32 (54%)	16 (27%)	11 (19%)	59 (100%)
Mature pipeline	33 (69%)	8 (17%)	7 (15%)	48 (100%)
Young pipeline	14 (54%)	7 (27%)	5 (19%)	26 (100%)

Note: This table presents descriptive evidence on when each type of company acquires projects. The rows give the type of acquirer. The columns give the number of acquired projects in each phase and the total. The percentages in parentheses then give each the share of each phase on the total number of transactions per acquirer type.

Consistently with the funnel structure of the R&D process in the pharmaceutical industry (Schuhmacher *et al.*, 2016; Paul *et al.*, 2010), the absolute number of transactions drops with more advanced phases. Along with that, also the structure of the activity changes. In

[30] Results remain robust if we exclude leaders as acquirers from the sample instead.

later phases where project outcomes are much more certain, big non-incumbents become relatively most important. Therefore, unlike others, the group of big non-incumbents seems to be engaging in a wait-and-see strategy and acquiring projects once they progressed at least to Phase II/III trials. Indeed, acquiring late was the most typical action of big non-incumbents as acquirers (45% of cases).

Table 2.6: Who acquires when?

	(1) Preclinical	(2) Late	(3) Preclinical	(4) Late
Leader + big inc			-0.104	-0.183
			(0.161)	(0.133)
Big non-inc	0.104	0.183		
	(0.161)	(0.133)		
Star	0.388**	-0.119	0.284	-0.302**
	(0.157)	(0.109)	(0.173)	(0.125)
Small	0.298**	-0.137	0.194	-0.321***
	(0.122)	(0.095)	(0.124)	(0.105)
Mature pip	0.247**	-0.105	0.143	-0.288**
	(0.114)	(0.093)	(0.144)	(0.119)
Young pip	0.095	-0.062	-0.009	-0.246*
	(0.141)	(0.111)	(0.160)	(0.128)
Cohort FE	Yes	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Obs	196	196	196	196
Adj. R2	0.188	0.171	0.188	0.171
Base	Leader + Big-inc	Leader + Big-inc	Big-noninc	Big-noninc
Phase sample	All	All	All	All

Note: This table presents the OLS estimates of the likelihood to acquire projects in a particular development phase. In columns (1) and (3), the dependent variable is equal to one if the project was taken over in the preclinical phase and zero if it was taken over in the clinical phase. In columns (2) and (4), the dependent variable is equal to one if the project was taken over in the late phase and zero if it was taken over in the preclinical or early phase. The independent variables are binary indicators for membership in the acquirer bin. Fixed effects capture membership in cohorts (Cohort FE), grouped mechanism of action FE (MoA FE) and region FE (Country FE) for acquirers. Errors are clustered at the project level and displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

The regression in table 2.6 confirms the descriptive findings when controlling for the relevant factors. In column (1) we see that compared to big incumbents (incl. leaders), stars, small firms and mature pipelines were significantly more likely to acquire projects in the preclinical phase. These firms are thus willing to accommodate risk since with all development ahead, preclinical projects are extremely uncertain to become marketed drugs. In contrast, in column (4), we see that stars, small, and pipeline acquirers are significantly less likely to acquire late, while big firms (and especially the non-incumbents) are likely to acquire late. This hints at a particular market entry strategy by big non-incumbent firms to wait and see which of the projects are the most successful. Previous literature has also highlighted this point from a different angle. Big firms may find it disadvantageous to engage in an R&D race with small firms, as they can obtain access to innovation

through acquisition Phillips & Zhdanov (2013). From the small firm's perspective, which are targeted by big non-incumbents, a company can struggle in later phases of innovation, because of its small size, lack of commercialization capabilities, or need for financial support. Therefore, selling out to a big acquirer might be the easiest way to bring its products to the product market (Comanor & Scherer, 2013). These results are robust when allowing for correlation of projects within the acquiring firm and clustering the standard errors at the firm level (please refer to Table 2.A.15 in Appendix 2.A).

2.4.4 Which projects are more likely to be taken over?

The summary statistics have already indicated that "high risk/high gain" projects changed ownership relatively more often (54%) compared to the non-treated projects (39%, see table 2.2). Table 2.7 then presents an analysis, testing this in a regression framework, when controlling for the relevant factors.

Table 2.7: What is acquired and when?

	(1)	(2)
NTO	0.035***	-0.015
	(0.013)	(0.029)
NTO × Preclinical		0.064**
		(0.030)
NTO × Early		0.055
		(0.034)
Preclinical		-0.047**
		(0.022)
Early		-0.038
		(0.024)
Cohort FE	Yes	Yes
MoA FE	Yes	Yes
Country FE	Yes	Yes
IPC	Yes	Yes
Patent nb.	Yes	Yes
No patents	Yes	Yes
Obs	2926	2926
Adj. R2	0.045	0.044
Base		Late

Note: This table presents the OLS estimates of the likelihood to acquire high risk/high gain projects using the project-phase sample. The dependent variable is a binary indicator equal to one for observations i experiencing an event in phase t (treated) and zero otherwise. In column (1), the independent variable is a binary indicator for NTO. In column (2), the independent variables are binary indicators for phases and NTO. Fixed effects capture membership in cohorts (Cohort FE), grouped mechanism of action (MoA FE) and region (Country FE). In every regression, We also control for the fixed effect for each IPC group and the number of patents assigned to a particular project. A dummy variable is included to control for the projects with no assigned patents, for which the novelty indicators are zero by construction. Errors are clustered at the project level and displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Column (1) of the table suggests that "high risk/high gain" status of a project is indeed a strong predictor of an ownership change. An NTO status is associated with 3.5 percentage point increase in the probability of being acquired. Given the baseline

probability of the ownership change of 7 percentage points, an increase by 3.5 percentage points represents an increase of 50% in the probability to be taken over. Thus, our analysis reveals that the potential breakthrough projects are much more likely to change hands. Column (2) then connects the "high risk/high gain" status of a project to the timing dimension and reveals that "high risk/high gain" projects are taken over in the earlier phases, namely in the preclinical phase (p-value of the interaction term is 0.036) and the early phase (p-value of the interaction term is 0.102). The robustness checks in Table 2.A.16 in Appendix 2.A then further confirm these findings. The results are robust when employing a different estimation method (logistic regression) as well as when clustering the standard errors at the firm level.

2.4.5 Who holds, buys and sells high-risk/high gain projects?

Knowing that NTO projects are more likely to change hands, which types of firms are likely sources of these projects? Table 2.8 examines this question. Unlike other regressions, in this case, the dependent variable corresponds to the NTO binary indicator, not to the binary indicator for the ownership change. The explanatory variables then include the binary variables for the membership in the target bins. The results indicate that potentially disruptive "high risk/high gain" projects are held by small non-incumbents and mature pipeline firms compared to big incumbents.[31]

Are specific types of companies likely to participate in the transactions involving these NTO projects? Table 2.9 analyses which types of firms are among the likely targets for their NTO projects. In this analysis, the dependent variable equals again the binary indicator for ownership change, and explanatory variables include the NTO indicators, together with the identities of the most frequent targets (small and pipeline companies) and the interaction terms. For transparency and ease of comparison, the first column of the table coincides with the first column of the 2.7. The results show that when accounting for the identity of the target, the main effect becomes insignificant and shifts entirely to the bin of mature pipeline firms. Thus, it is the mature pipeline firms who not only hold NTO projects but are also the most likely ones to sell them. Again, these results are robust to firm-level clustering (Table 2.A.16 in Appendix 2.A).

From the perspective of acquirers, table 2.10 shows two regressions aimed at analysing which types of firms take over NTO projects from others. Firstly, relative to big firms, only stars and mature pipelines are likely to acquire NTO projects and technological novelty is thus a key determinant in the acquisition strategies of these risk-taking firms. Secondly, relative to mature pipelines, big firms (including incumbents and leaders) are in general less likely to acquire potentially disruptive, "high-risk/high-gain" projects. This is an important finding in the context of the current policy debates. Federico *et al.* (2020)

[31] These findings are fully aligned with previous theoretical insights of Henkel *et al.* (2015) who show that new entrants to a market tend to be superior to incumbents in originating radical innovations.

Table 2.8: Who holds which projects?

	(1) NTO
Leader	0.070
	(0.047)
Big non-inc	0.066
	(0.046)
Star	0.142
	(0.170)
Small inc	0.159
	(0.111)
Small non-inc	0.084**
	(0.042)
Mature pip	0.086**
	(0.041)
Young pip	0.036
	(0.039)
Cohort FE	Yes
MoA FE	Yes
Country FE	Yes
IPC	Yes
Patent nb.	Yes
No patents	Yes
Obs	2926
Adj. R2	0.346
Base	Big-inc

Note: This table presents the OLS estimates of the likelihood to own various projects using the project-phase sample. The dependent variable is a binary indicator equal to one for observations i that score on the NTO indicator and zero otherwise. The independent variables are binary indicators for the membership in the target bin. Fixed effects capture membership in cohorts (Cohort FE), grouped mechanism of action (MoA FE) and region (Country FE). We also control for the fixed effect for each IPC group and the number of patents assigned to a particular project. A dummy variable is included to control for the projects with no assigned patents, for which the novelty indicators are zero by construction. Errors are clustered at the project level and displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 2.9: Who is likely to sell which projects?

	(1)	(2)
NTO	0.035***	0.002
	(0.013)	(0.012)
NTO × Small		0.024
		(0.027)
NTO × Mature pip		0.055*
		(0.029)
NTO × Young pip		0.034
		(0.028)
Small		0.101***
		(0.017)
Mature pip		0.089***
		(0.016)
Young pip		0.070***
		(0.015)
Cohort FE	Yes	Yes
MoA FE	Yes	Yes
Country FE	Yes	Yes
IPC	Yes	Yes
Patent nb.	Yes	Yes
No patents	Yes	Yes
Obs	2926	2926
Adj. R2	0.045	0.073
Base		Big + star

Note: This table presents the OLS estimates of the likelihood that an NTO project of a particular target type is acquired, using the project-phase sample. The dependent variable is a binary indicator equal to one for observations i experiencing an event in phase t (treated) and zero otherwise. The independent variables are binary indicators for NTO and indicator variables for membership in the target bin and interaction terms. Fixed effects capture membership in cohorts (Cohort FE), grouped mechanism of action (MoA FE) and region (Country FE). We also control for the fixed effect for each IPC group and the number of patents assigned to a particular project. A dummy variable is included to control for the projects with no assigned patents, for which the novelty indicators are zero by construction. Errors are clustered at the project level and displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

notes that a successful incumbent firm that is profiting greatly from the status quo has a powerful incentive to preserve those profits, and this can mean slowing down or blocking disruptive threats. However, our findings indicate that those firms are less likely to acquire such disruptive projects in the first place - limiting the potential scope of such pre-emptive discontinuations.

Table 2.10: Who acquires NTO?

	(1)	(2)
Big		-0.249**
		(0.116)
Star	0.452*	0.203
	(0.240)	(0.245)
Small	0.135	-0.114
	(0.116)	(0.111)
Mature pip	0.249**	
	(0.116)	
Young pip	-0.025	-0.274*
	(0.166)	(0.154)
Cohort FE	Yes	Yes
MoA FE	Yes	Yes
Country FE	Yes	Yes
IPC	Yes	Yes
Patent nb.	Yes	Yes
Obs	196	196
Adj. R2	0.320	0.320
Base	Big	Mature pip

Note: This table presents the OLS estimates of the likelihood to acquire NTO projects using the sample of all ownership changes. The dependent variable equals one if project i was an NTO project and zero otherwise. The independent variables are binary indicators for acquirer bins. Fixed effects capture membership in cohorts (Cohort FE), grouped mechanism of action (MoA FE) and region (Country FE). We also control for the fixed effect for each IPC group and the number of patents assigned to a particular project. A dummy variable is included to control for the projects with no assigned patents, for which the novelty indicators are zero by construction. Errors are clustered at the project level and displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Given that mature pipelines firms are more likely to hold, sell, and acquire NTO projects, the last table 2.11 analyses whether the NTO acquisitions are driven by mature pipeline - mature pipeline transaction, or other transactions in which the mature pipeline firms were targets. The results do indeed confirm that NTO plays a central role only in occasions where a mature pipeline firm acquired projects of another mature pipeline target. In other transactions, NTO is not an important predictor of a pair-wise match between firms. Thus, mature pipeline - mature pipeline transactions are likely to involve "high-risk/high-gain" projects. These results hold irrespective of the level of error clustering (Table 2.A.17 in Appendix 2.A).

To conclude,"high-risk/high-gain" profile plays an important role in acquisitions of R&D projects. These projects are likely to be originated by pipeline firms and acquired in the early and uncertain stages of their development. While playing an important role

Table 2.11: Which matches are more likely for NTO projects?

	(1)
A mature pip + T mature pip	0.203*
	(0.120)
A not mature pip + T mature pip	0.156
	(0.101)
Cohort FE	Yes
MoA FE	Yes
Country FE	Yes
IPC	Yes
Patent nb.	Yes
Obs	196
Adj. R2	0.297
Base	Other

Note: This table presents the OLS estimates of how likely a particular match of acquirer and target is for the NTO projects, using the sample of all ownership changes. The dependent variable equals one if the acquired project was an NTO project and zero otherwise. The independent variables are binary indicators for combinations of acquirer and target bins. Fixed effects capture membership in cohorts (Cohort FE), and regions (Country FE). We also control for the fixed effect for each IPC group and the number of patents assigned to a particular project. A dummy variable is included to control for the projects with no assigned patents, for which the novelty indicators are zero by construction. Errors are clustered at the project level and displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

in the acquisition strategies of star firms that acquire targets owning such projects, these projects seem to have a particularly prominent role in the transactions involving mature pipeline acquirers and mature pipeline targets.

2.5 Conclusion

Mergers and acquisitions are pervasive in the pharmaceutical industry and closely linked to firms' innovation activities. Making use of a detailed dataset which tracks ownership changes for all corporate R&D related to the treatment of diabetes in the period 1997-2017, this paper analyzes and characterizes M&A involving drugs under development. We provide new and detailed insights into which companies are likely acquirers and targets, at what stage of the drug development cycle acquisitions take place, and which type of projects undergo ownership changes in terms of their technological "high risk/high gain" potential.

We find substantial and so far largely unexplored heterogeneity in the types of M&A transactions, going beyond the narrative of large firms taking over small ones. While recent debates focused mainly on the role of market power in the M&A deals, particularly as regards incentives to buy and "kill" (Cunningham *et al.*, 2021), we show that the firms with the largest market power pursue M&A to a limited extent and avoid takeovers of the potentially most valuable novel projects. Instead, we show that a substantial amount of the total M&A activity is ongoing between small and research-focused firms in early development phases and concerns novel projects. Taken together, our findings highlight that the technology uncertainty and high-risk/high-gain technology profiles of projects are an important part of the changing-ownership-stories. Policymakers and academic researchers should consider these dimensions in any discussions focused on M&As and innovation.

Given the lack of research on this topic, in future work, it would be interesting to investigate even deeper the transaction dynamics between the purely innovating pipeline firms. For example, it would be interesting to see whether the acquisitions of projects of other pipeline firms are motivated by the removal of potential competition (consolidation), technological synergies, or relate to the target's financial distress. Another array of future research relates to the exploration of the role of other technology-based characteristics in the decision of firms to engage in M&As, for example, the closeness of acquirer's and target's projects in technology space. While beyond the scope of this paper, a natural question for follow-on research is also how robust our results are beyond the antidiabetics markets. With sufficient resources, the framework we developed can be scaled to other therapeutic areas as well.

2.A Robustness checks and extensions

Table 2.A.12: Extensions: Who acquires and who sells?

	Acquirers			Targets		
	(1)	(2)	(3)	(4)	(5)	(6)
	Big M&As	No spinoffs/divest	Launched included	Big M&As	No spinoffs/divest	Launched included
Leader	-0.045**	-0.072***	-0.083***	0.012	-0.008	-0.022**
	(0.022)	(0.018)	(0.019)	(0.014)	(0.007)	(0.011)
Big non-inc	0.039	0.015	0.024	0.046***	0.004	0.004
	(0.026)	(0.023)	(0.024)	(0.017)	(0.006)	(0.012)
Star	0.274***	0.270***	0.275***	0.005	0.039	0.005
	(0.105)	(0.103)	(0.104)	(0.052)	(0.053)	(0.050)
Small	0.007	0.005	0.013	0.108***	0.105***	0.121***
	(0.021)	(0.020)	(0.021)	(0.016)	(0.014)	(0.016)
Mature pip	-0.018	-0.025	-0.019	0.104***	0.103***	0.103***
	(0.021)	(0.020)	(0.021)	(0.016)	(0.014)	(0.016)
Young pip	-0.060***	-0.071***	-0.060***	0.058***	0.069***	0.058***
	(0.019)	(0.018)	(0.019)	(0.013)	(0.012)	(0.013)
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	2926	2898	3009	2926	2898	3009
Adj. R2	0.039	0.042	0.043	0.038	0.043	0.053
Base	Big-inc	Big-inc	Big-inc	Big-inc	Big-inc	Big-inc

Note: This table presents the OLS estimates of the likelihood to acquire projects using the project-phase sample. The dependent variable is a binary indicator equal to one for observations i experiencing an event in phase t (treated) and zero otherwise. The independent variables are binary indicators for membership in the acquirer bin (columns 1 to 3) or the target bin (columns 4 to 6) . In columns (1) and (4), ownership changes relating to big mergers are considered as well, resulting in 217 treated observations. Columns (2) and (5) drop all ownership changes relating to spin-offs or partial M&As (divestitures). Column (3) and (6) include observations for the launched phase. Fixed effects capture membership in cohorts (Cohort FE), grouped mechanism of action FE (MoA FE) and region FE (Country FE). Errors are clustered at project level and displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 2.A.13: Robustness: Who acquires and who sells?

	Logit		Cluster SE: Firm	
	(1)	(2)	(3)	(4)
	Acquirers	Targets	Acquirers	Targets
Leader	-1.631***	0.644	-0.071***	-0.007
	(0.507)	(1.230)	(0.015)	(0.011)
Big non-inc	0.222	1.569	0.017	0.014
	(0.325)	(1.154)	(0.029)	(0.018)
Star	1.799***	2.624*	0.283**	0.020
	(0.580)	(1.371)	(0.130)	(0.055)
Small	0.173	3.953***	0.015	0.121***
	(0.290)	(1.009)	(0.021)	(0.021)
Mature pip	-0.192	3.844***	-0.014	0.113***
	(0.308)	(1.009)	(0.020)	(0.020)
Young pip	-1.021***	3.234***	-0.056***	0.067***
	(0.329)	(1.006)	(0.016)	(0.017)
Cohort FE	Yes	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Obs	2926	2926	2926	2926
Pseudo R2	0.097	0.143		
Adj. R2			0.039	0.048
Base	Big-inc	Big-inc	Big-inc	Big-inc

Note: This table presents several robustness checks of the baseline specification regarding the likelihood to acquire projects (odd columns) or sell projects (even columns) using the project-phase sample. Columns (1) and (2) use a logit estimation instead of OLS. Columns (3) and (4) use OLS estimation and cluster standard errors at the firm level (instead of the project level), allowing for correlation of the error term within firms. The dependent variable in each column is a binary indicator equal to one for observations i experiencing an event in phase t (treated) and zero otherwise. The independent variables are binary indicators for membership in the acquirer bin (odd columns) or the target bin (even columns). Fixed effects capture membership in cohorts (Cohort FE), grouped mechanism of action FE (MoA FE) and region FE (Country FE). * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 2.A.14: Robustness: Who acquires whom?

	Cluster SE: Firm		
	(1)	(2)	(3)
	T is pip	T is young pip	T is young pip
Leader + big inc	0.279	0.430**	0.264**
	(0.173)	(0.176)	(0.130)
Star	0.494**	0.336	0.337*
	(0.206)	(0.300)	(0.202)
Small	0.344**	-0.205	-0.005
	(0.138)	(0.181)	(0.074)
Mature pip	0.642***	-0.066	0.095
	(0.140)	(0.179)	(0.089)
Young pip	0.570***	0.116	0.230*
	(0.164)	(0.206)	(0.129)
Cohort FE	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes
Country FE A+T	Yes	Yes	Yes
Obs	189	121	196
Adj. R2	0.230	0.225	0.191
Base	A Big-noninc	A Big-noninc	A Big-noninc
Target Sample	Small + pip	Pipeline	All

Note: This table presents the OLS estimates of the likelihood to acquire projects of a particular target type using the sample of ownership changes only. In column (1), the estimation sample is restricted to the ownership changes of small and pipeline targets. The dependent variable is equal to one if the target was a pipeline firm and zero if the target was a small firm. In column (2), the estimation sample is restricted to the ownership changes of pipeline targets. The dependent variable is equal to one if the target was a young pipeline firm and zero if the target was mature pipeline firm. In column (3), the estimation sample is unrestricted and encompasses all ownership changes. The dependent variable is equal to one if the target was a young pipeline firm and zero otherwise. Fixed effects capture membership in cohorts (Cohort FE), grouped mechanism of action (MoA FE) and region (Country FE) for both acquirers and targets. Errors are clustered at firm level and displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 2.A.15: Robustness: Who acquires when?

	Cluster SE: Firm	
	(1)	(2)
	Preclinical	Late
Leader + big inc		-0.183
		(0.128)
Big non-inc	0.104	
	(0.204)	
Star	0.388**	-0.302***
	(0.165)	(0.115)
Small	0.298**	-0.321***
	(0.143)	(0.097)
Mature pip	0.247*	-0.288***
	(0.130)	(0.104)
Young pip	0.095	-0.246**
	(0.167)	(0.118)
Cohort FE	Yes	Yes
MoA FE	Yes	Yes
Country FE	Yes	Yes
Obs	196	196
Adj. R2	0.188	0.171
Base	Leader + Big-inc	Big-noninc
Phase sample	All	All

Note: This table presents the OLS estimates of the likelihood to acquire projects in a particular development phase using the sample of ownership changes. In column (1), the dependent variable is equal to one if the project was taken over in preclinical phase and zero if it was taken over in clinical phase. In column (2), the dependent variable is equal to one if the project was taken over in the late phase and zero if it was taken over in preclinical or early phase. The independent variables are binary indicators for membership in the acquirer bin. Fixed effects capture membership in cohorts (Cohort FE), grouped mechanism of action FE (MoA FE) and region FE (Country FE) for acquirers. Errors are clustered at firm level and displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 2.A.16: Robustness: What, when, and from whom?

	Logit			Cluster SE: Firm		
	(1)	(2)	(3)	(4)	(5)	(6)
	What	What and when	What from whom	What	What and when	What from whom
NTO	0.434**	-0.375	1.493	0.035***	-0.015	0.002
	(0.191)	(0.396)	(1.141)	(0.013)	(0.026)	(0.013)
NTO × Preclinical		1.070**			0.064**	
		(0.435)			(0.030)	
NTO × Early		0.845*			0.055	
		(0.501)			(0.034)	
Preclinical		-0.824***			-0.047**	
		(0.299)			(0.023)	
Early		-0.623*			-0.038	
		(0.356)			(0.024)	
NTO × Small			-1.481			0.024
			(1.171)			(0.028)
NTO × Mature pip			-1.000			0.055*
			(1.153)			(0.031)
NTO × Young pip			-1.211			0.034
			(1.171)			(0.032)
Small			4.182***			0.101***
			(1.037)			(0.022)
Mature pip			4.003***			0.089***
			(1.024)			(0.019)
Young pip			3.702***			0.070***
			(1.044)			(0.019)
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
IPC	Yes	Yes	Yes	Yes	Yes	Yes
Patent nb.	Yes	Yes	Yes	Yes	Yes	Yes
No patents	Yes	Yes	Yes	Yes	Yes	Yes
Obs	2926	2926	2926	2926	2926	2926
Psuedo R2	0.138	0.139	0.215			
Adj. R2				0.045	0.044	0.073
Base		Late	Big + star		Late	Big + star

Note: This table presents several robustness checks related to the high-risk/high-gain (NTO) status of projects using the sample of all observations. Columns (1) and (4) estimate the likelihood that NTO is a predictor of an ownership change. Columns (2) and (5) estimate in which development phases are NTO projects likely to undergo ownership changes. Columns (3) and (6) then estimate the likelihood that certain targets sells NTO projects. Columns (1) to (3) use use logistic regression for the estimation instead of OLS. Columns (4) to (6) use OLS estimation and cluster standard errors at the firm level (instead of the project level), allowing for correlation of the error term within firms. The dependent variable in each column is a binary indicator equal to one for observations i experiencing an event in phase t (treated) and zero otherwise. The independent variables are binary indicators. Fixed effects capture membership in cohorts (Cohort FE), grouped mechanism of action FE (MoA FE), region FE (Country FE) and technology groups (IPC FE). Additional unreported controls include the number of patents and a binary indicator for projects with no assigned patents. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 2.A.17: Robustness: Who acquires NTO and from whom?

	Cluster SE: Firm	
	(1)	(2)
	NTO by whom	NTO match
Star	0.452*	
	(0.239)	
Small	0.135	
	(0.123)	
Mature pip	0.249**	
	(0.110)	
Young pip	-0.025	
	(0.162)	
A mature pip + T mature pip		0.203*
		(0.112)
A not mature pip + T mature pip		0.156
		(0.103)
Cohort FE	Yes	Yes
MoA FE	Yes	Yes
Country FE	Yes	Yes
IPC	Yes	Yes
Patent nb.	Yes	Yes
Obs	196	196
Adj. R2	0.320	0.297
Base	Big	Other

Note: This table presents the OLS estimates of how likely specific acquirers are to take over high-risk/high-gain (NTO) projects (column 1) and likely a particular acquirer-target match is for the NTO projects (column 2). The estimation sample is restricted to the set of ownership changes only. The dependent variable is equal to one if the project was an NTO project. The independent variables include binary indicators for acquirer types (column 1) or acquirer-target matches (column 2). Fixed effects capture membership in cohorts (Cohort FE), grouped mechanism of action (MoA FE), region (Country FE), and technology groups (IPC FE). Additional unreported controls include the number of patents and a binary indicator for projects with no assigned patents. Errors are clustered at firm level and displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

2.B Examples of companies in the bins

Figure 2.B.5: Examples of firms in bins

2.C Patent matching

This appendix is a guide to matching patents to antidiabetic projects under development. According to our knowledge, there is no publicly available database providing project-patent links beyond launched drugs. The private databases, on the other hand, do not provide sufficient coverage needed for this paper and/or are prohibitively costly. In addition, these databases (e.g. Cortellis) are in-transparent as to how they assign patents to projects.

2.C.1 Patent databases

When establishing project-patent, we focus on US patents exclusively. Three sources of patent data are used: the USPTO patent database, PATSTAT database, and the LENS patent database.[32] In these databases, we particularly utilize the information regarding patent application dates, patent grant dates, information on patent extensions, priority dates, IPC patent classes, patent families, patent assignees, and backward patent references.

The matching algorithm starts with establishing so called 'candidate patent sets' for each project in the sample. The candidate patent set contains all US granted patents of firms that were involved in the development of a project and that were filed for between the initiation date and the termination date of a project (which are retrieved from the progression thought the development phases). The relevant companies for each project include the originator, the final owner, and all owners in the chain in between. We create a list of 1314 unique companies for the set of 2387 projects left in our sample after matching to clinical trial information and apply a fuzzy string-matching routine to establish links between these firms and the patent assignees.[33] To narrow down the patent universe to technologies which plausibly relate to pharmaceutical markets and drug development, we follow Schmoch (2008) and own analysis based on the FDA's Orange book and consider only patents with at least one IPC subclass in A61K, A61M, A61P, C07C, C07D, C07F, C07K, C07H, C08F, C08G, C12N or C12P. After this step, each project has a set of patents filed for during its lifetime and belonging to the relevant firms - so called candidate set - in which we look for the patents relevant for a specific project.

There are two major issues when searching for patents belonging to a specific project. First, there is significant heterogeneity between the projects in our sample, ranging from

[32] Available at https://lens.org and https://patentsview.org/download/data-download-tables

[33] The improve precision of the matching routine, we first perform a standardization procedure to unify fuzzy names of the patents assignees. We remove the legal forms of companies, clean the names from non-alphanumerical characters, remove generic words and combinations of words that do only distinguish separate legal entities, but refer to the same underlying company (eg 'pharmaceutical products', 'intellectual properties', 'healthcare systems'). After this, we standardize the company names by taking the first word of the company's name, keeping the most numerous one and using it to substitute the other names referring to that company. Following the fuzzy string-matching routine, we manually checked the correctness on a random sub-sample of firms. The procedure yielded a minimum of false matches.

small molecule drugs to large molecule drugs. Second, substantial number of projects in the Pharmaprojects database miss information along the relevant dimensions. Our matching procedure therefore consists of several complementary approaches that try to overcome these issues.

2.C.2 Chemistry matching

Patents for 'small molecule' or chemical drugs are matched based on their chemical properties. Using the Surechem[34] database and various cross-roads (CAS numbers, SMILES Chemical structures, and UNII identifiers), we link project's underlying chemical compounds to PubChem.[35] PubChem contains information on the patents protecting specific molecules for some chemical entities. If we find an overlap between the 'candidate set' and the set of patents retrieved from PubChem, we consider these patents as assigned to a project.[36] This procedure results in 411 matches which are excluded from further matching.

2.C.3 Gene matching

Patents for 'large molecule' or biological drugs are matched based on gene sequences. Since the early 1990s, US patent applications claiming genes as intellectual property must disclose the exact DNA or protein sequences claimed in the text of the patent. The sequences are listed in USPTO patent applications in a standard format, labelled with the text 'SEQ ID NO'. Standard bio-informatics methods can be used to compare these sequences against the census of human genome gene sequences to annotate each sequence with standard gene identifiers. In turn, these can be linked to outside databases, eg. the Pharmaprojects database.

We broadly follow the methodology suggested in Sampat & Williams (2019). First, we extracting standard gene identifiers (known under Entrez gene ID) from the Pharmaprojects database. Using the GeneBank generated crosswalks,[37] we connected each of the gene IDs to a list of mRNA, RNA and protein RefSeq accession/version numbers and extracted the corresponding nucleotide sequences from the GeneBank's webpage, including start and end positions of the chain in the sequence, if applicable. To capture the full universe of known sequences relating to a particular gene ID, we utilize both the sequences relating to the annotated human genome as well as the sequences maintained independently of the annotated genome.

[34] SureChEMBL provides free access to chemical data extracted from the patent literature. Available here: ftp://ftp.ebi.ac.uk/pub/databases/chembl/SureChEMBL/

[35] PubChem is a publicly available, open chemistry database at the National Institutes of Health (NIH). Available here: https://pubchemdocs.ncbi.nlm.nih.gov/downloads

[36] If match on company level is found but the patent lies outside of the development window of a project, we check whether other member of the patent's family lies in the development window of a project. If so, this is considered a match as well.

[37] Available here: ftp://ftp.ncbi.nih.gov/refseq/release/release-catalog/release97.accession2geneid.gz

Following the methodology pioneered by Jensen & Murray (2005), we use the BLAST (Basic Local Alignment Search Tool) search engine to compare the above sequences to census of sequences disclosed in the US patents. To arrive at the true set of matches, we only consider blast matches with an E-value of less than 1e-50.[38] This yields a final mapping between projects and patents referring to the respective gene via the disclosed sequences. Again, we only consider patents as assigned to a project when they fall within the development window and belong to one of the relevant firms. This procedure results in 222 matches which are excluded from further matching.

2.C.4 MoA keyword matching

We complement the two above approached by utilizing text analysis and information on mechanism of action (MoA) - the underlying mechanism determining how the drug produces the required target action affecting processes in the body. For projects with known MoA, we first perform cleaning to standardize MoA names,[39] obtaining a set of tokens ("keywords"). Using a combination of TF-IDF algorithm and a manual check, we also retrieve all relevant synonyms relating to a particular MoA[40] and add those to the relevant "keywords". To find the counterparts of these "keywords" in patents and establish matches, we utilize the database of Arts *et al.* (2021) who pre-process the text in the patents by concatenating the title and abstract and claims text, lowercasing the text, tokenizing all words, and eliminating stop words based on a manually compiled list, removing words with only one character, numbers, and words that appear only once across all patents. We pair a patent to a project if all MoA "keywords" or an abbreviation are found in a patent document. We then check that only patents are kept that belong to the relevant firms and which were applied for during the development window of a drug project. This procedure results in 487 matches which are excluded from further matching.

2.C.5 Remaining matching

If no match for a project has been found so far, we proceed by various plausible exclusion restrictions. For example, projects with a single patent in the candidate set is considered matched. Similarly, all patents are assigned to a project if the firm had only one project under development. We also matched projects to 0 patents if a firm had no US granted patent (we checked all these instances manually to verify the absence of the US patents).

[38] Sampat & Williams (2019) use an E-value of exactly 0. However, we apply less strict threshold as it was confirmed by bio-informatics specialist that our threshold level is commonly applied in the field and a threshold of strict zero might be too restrictive.

[39] This includes tokenization, removal of special characters or words comprised of single letter only, and stemming using Porter's stemmer.

[40] For example, the MoA "glucagon-like peptide" is often only mentioned using its abbreviation glp1 or glp-1. The fibroblast growth factor 21 is known under fgf21 or fgf-21. The DPP-IV mentioned above is sometimes referred to as DPP-4, DPP4 or dipeptidyl peptidase 4 inhibitor.

Lastly, 93 launched projects were matched to patents based on information in the FDA's Orange Book. In total, this yielded an additional 660 matches.

The remaining set of 570 unmatched projects were checked manually. For each of these projects, patent text was compared to the above project's properties (where available). We hired a chemistry student to then read these text fields, compare those, and decide which patents from the set of candidate patents should be relevant for a particular project. This resulted in another 161 matched project. We drop the final unmatched projects from our sample. This leaves us with a final sample of 1941 matched projects and 437 unmatched projects, representing a match rate of 81%. A total of 4999 patents were assigned to the projects.

2.C.6 Summary statistic on matching

Although the Orange Book only considers drugs that were eventually launched on the market, and thus only captures the most selective subset of successful projects not representative of the entire pipeline, it is currently the only source of information on patent-project links. Below, we present basic descriptive statistics to put the results of the matching procedure into a perspective.

Table 2.C.18: Orange Book and the matched sample: summary statistics

Matched sample			
Patents		Projects	
Mean - projects per patent	1.8	Mean - patents per project	4.32
Max - projects per patent	25	Max - patents per project	113
Share of patents with single assigned project	67.3	Share of projects with single assigned patent	38.14
		Share of projects with no assigned patent	9.64
Orange Book			
Patents		Projects	
Mean - projects per patent	1.95	Mean - patents per project	10.91
Max - projects per patent	12	Max - patents per project	46
Share of patents with single assigned project	66.16	Share of projects with single assigned patent	9.76
		Share of projects with no assigned patent	0

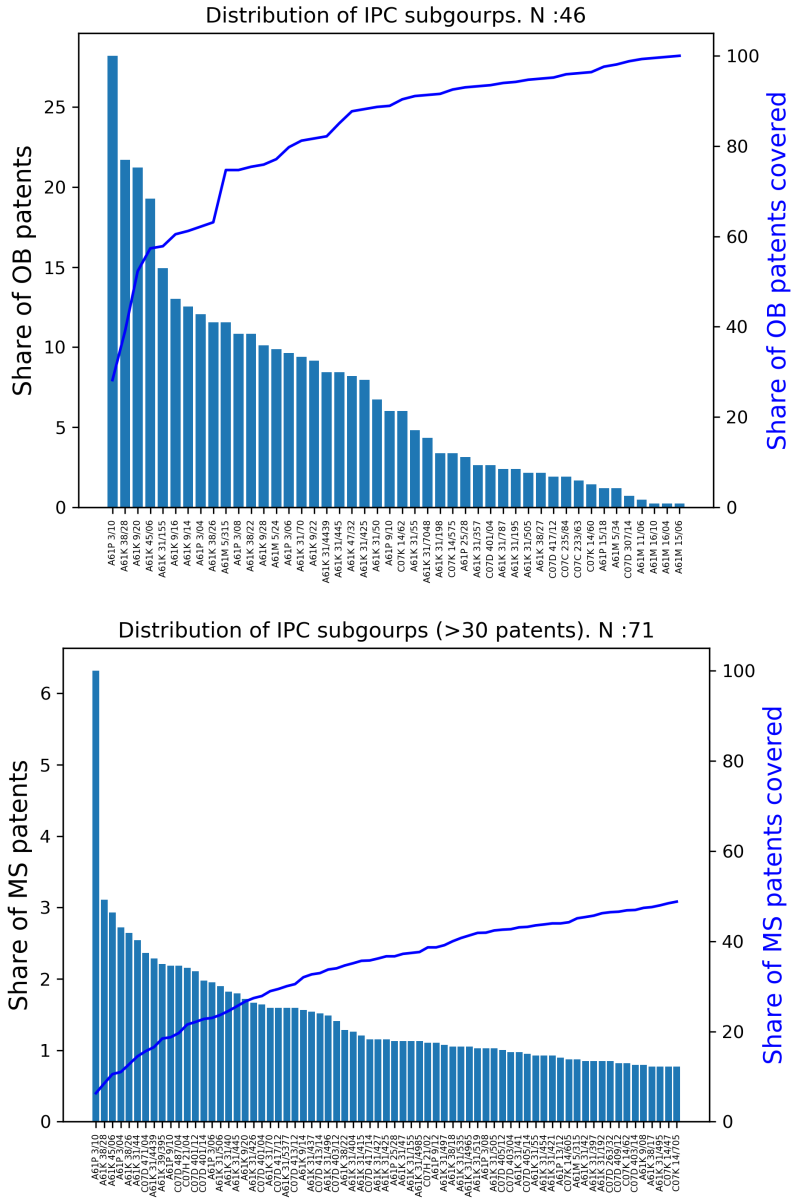
Since a project can have more than one patent assigned, we are presenting a project level and a patent perspective in Table 2.C.18. Overall, the presented figures lend credibility to the outcome of the patent matching. Form the patent level perspective, the average number of projects assigned per patent amounts to 2 across both samples and in both cases with around 67% of patents assigned to a single project. From the project level perspective, a project has on average 4.32 patents in the matched sample and almost 11 patent in the Orange book. Considering that the matched sample includes much broader set in earlier development phases, a much higher fraction of the projects have only a single patent assigned compared to the Orange Book (38% in the matches sample vs. 10% in the

Orange Book). This result should be expected since it indicates that successful projects that are launched to the market have substantial patent protection, with none of the launched drugs being unpatented. On average, in line with the findings of Argente *et al.* (2019), average number of patents per project in the matched sample is lower than in the Orange Book sample.

The Figure 2.C.6 provides a comparison between the types of patents that were assigned to the projects (lower part) and how these compare to the Orange book (upper part). Using IPC classification subgroups level (the most granular classification available in the IPC classification), we plot the distributions of the occurrence of the IPC classes in the two sets. We should expect that whilst the matched sample will include many more IPC categories compared to the Orange Book due to the dispersion of the pipeline R&D activities, as least the top classes in both sets should be similar if properly assigned.

The distribution of the matched sample is much broader compared to the Orange Book, and each IPC subgroup occurs less frequently. For example, the most frequent subgroup A61P 3/10 occurs in more than 25% of patents in the Orange Book and in slightly more than 6% of cases in the matched sample. However, comparing the ordering of the different IPC subgroups, we can find that among the most 5 frequent groups in each sample, 3 subgroups are shared and have high relevance to diabetes drugs: A61P 3/10 - drugs for hyperglycaemia, e.g. antidiabetics, A61K 38/28 - insulins, and A61K 38/28 - Mixtures of active ingredients without chemical characterization, e.g. antiphlogistics and cardiaca. In addition, third most frequent subgroups in the matched sample - A61K 38/26 - refers to Glucagons. This shows that the matched patents indeed relate closely to diabetes and at least the most frequent technological subgroups closely mirror the sample of drugs launched on the market.

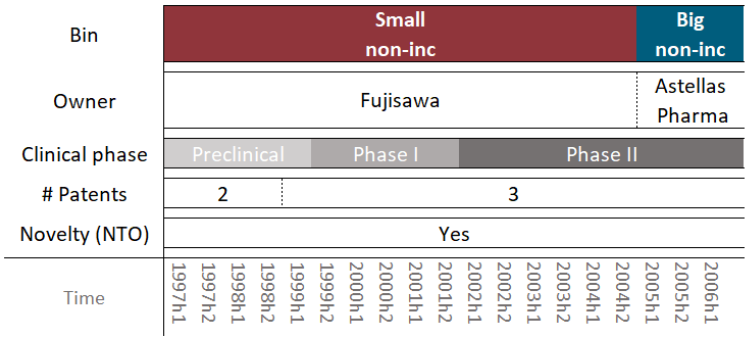
Figure 2.C.6: Distribution of IPC subgroups in Orange Book and matched samples



2.D Illustrative example

Figure 2.D.7 uses the drug project FK-614 (also known as ATx-08-001) as an example to illustrate how all dimensions of the data and the variables connect together. This project's mechanism of action is a PPAR gamma agonist - a popular MoA to treated diabetes. The development of FK-614 was initiated by Fujisawa, a small non-incumbent and a Japanese company. The projects is in the sample as of the starting date (first semester of 1997) since Fujisawa initiated the preclinical development in 1996. Based on the 2 US patent families protecting the compound at the beginning of its development and before entering the clinical trials, the project was identified as novel. While in Phase II testing, the project was acquired by Astellas Pharma, a big non-incumbent company at that time. The project was never successfully launched, as its development was abandoned in Phase II testing.

Figure 2.D.7: Project example and relevant variables



As illustrated by Table 2.D.19, the drug FK-614 adds three observations to the database. The first measurement takes place before at the beginning of the preclinical phase (1997h1). The second measurement takes place at the beginning of the early phase (Phase I) in 1999h2. The third and the last measurement takes place before the acquisition of the drug by Astellas Pharma (2004h2).

Table 2.D.19: Database extract

Drug ID	Phase	Date	type	Event	Acquirer	A bin	Target	T bin	NTO	Patents
24056	Preclin	1/01/1997		0	Fujisawa	Small non-inc	Fujisawa	Small non-inc	1	2
24056	Early	1/07/1999		0	Fujisawa	Small non-inc	Fujisawa	Small non-inc	1	3
24056	Late	1/07/2004	M	1	Astellas Pharma	Big non-inc	Fujisawa	Small non-inc	1	3

Chapter 3

Are M&As spurring or stifling innovation? [1]

Abstract. The innovation impact of acquisitions of small nascent competitors by large product market incumbents is hotly debated, where incumbents might preemptively terminate or "kill" the projects of these small targets. This paper provides empirical evidence on which M&A deals spur and which stifle innovation. We do not only look at the product market position of the acquiring firm but additionally at the position of both parties in the technology markets. Using the setting of antidiabetics, our granular dataset tracks life cycles and patenting for all antidiabetic drugs under development ("projects") between 1997 and 2017. We show that most terminations of the acquired projects occur when projects are still far from launch into the product market, indicating that early R&D stages might be more important to focus on than late stages. Furthermore, transactions have a positive impact on innovation when large product market incumbents acquire projects close to their own in both technology and product markets. In these deals, both the target's and acquirer's projects enjoy innovation benefits through increased patenting, consistent with the exploitation of technology synergies. Our results hint at the crucial role of technology positions when assessing pharma M&As.

[1] This chapter is co-authored with Jo Seldeslachts (KU Leuven & DIW Berlin), and Reinhilde Veugelers (KU Leuven). We thank seminar participants at KU Leuven and DIW Berlin for their valuable comments and suggestions. Jan Malek acknowledges support from FWO through project 3H190094.

3.1 Introduction

Acquisitions of small nascent competitors by large incumbents in highly innovative sectors like pharmaceuticals have recently drawn the attention of researchers and policy makers alike. These deals typically involve target firms with not yet marketed products and thus largely escape antitrust scrutiny (See e.g Malek *et al.*, 2021; Wollmann, 2020). However, they can still have a sizeable effect on innovation and, therefore, competition. Of particular concern are those transactions where large product incumbent firms buy startups working on innovative projects, intended to preemptively discontinue startup's projects to avoid erosion of the incumbent's market power. Such "killer acquisitions" might stifle innovation (Cunningham *et al.*, 2021).

However, while the acquirer's product market incumbency has been rightly on the radar, M&As also evolve around strategies how to obtain new technological resources or reap technological benefits from combining technological know-how. Caffarra *et al.* (2020), for example, present the concept of "reverse killer acquisitions" when established incumbents engage in "buy instead of build" type of acquisitions. Rather than killing the projects of the target, acquirers integrate those into their innovation ecosystem. What the impact of such situations will be on innovation does not depend only on the positions of the parties in product markets but also on technological dimensions (Cassiman *et al.*, 2005). Whilst missing the technology side runs the risk of erroneous assessments, empirical studies remain thin on assessing product market and technology characteristics together (Veugelers, 2012).

This paper studies whether acquisitions of small targets spur or stifle innovation, depending on the incumbency and closeness of the acquirer and the target in both the technology and product markets. We focus on the antidiabetics industry, a therapeutic area within the pharmaceutical industry where innovation is well-defined and well-documented at the level of individual projects. We trace 1345 projects developed between 1997 and 2017 and connect them to their R&D milestones and ownership changes. We further match each project to its patents, which allows us to zoom in on the technology characteristics of the projects. With this granular dataset at hand, we estimate how different outcomes - a project's termination, progression, and new patenting - change if projects were affected by ownership changes, compared to similar projects that were not affected by ownership changes. The structure of our data allows us to analyze not only the impact on the projects that were taken over, but also on antidiabetics projects in the acquirer's portfolio. We can thus assess the impact on each transacting party separately and also the overall combined effect of a transaction.[2]

To further explore heterogeneity in the effects and separate M&A deals with positive effects from the harmful ones, our analysis considers the acquirer's product market in-

[2]Previous literature has stressed the need to investigate the effects of acquisitions on both the target firms as well as the acquiring firms since the effects on these firms can substantially differ (Szücs, 2014).

cumbency, technology incumbency, as well as the closeness of the acquirer's and target's R&D and product markets positions. We define product market incumbents as firms with launched antidiabetic drugs, while R&D incumbents are firms with their own pre-existing antidiabetic R&D portfolio. The closeness between projects in product markets is characterized by the overlap in "mechanism of action" (MoA).[3] The closeness between projects in the technology space is based on text similarity contained in the patents associated with each R&D project.[4] We further look at the early or late stages of development in which projects were subject to a M&A transaction.

We find that, on average, M&As harm the target projects. Target projects are more likely to be terminated and less likely to progress, compared to similar non-acquired projects. They also experience, on average, less patenting after a transaction, further suggesting a negative impact on innovation.[5] However, this overall pattern differs markedly along the product market and technology dimensions, and the stages of development of the projects. Acquisitions that take place in the later phases of clinical development, where the costs of development are high but the risks of failure low, are limited in number but not likely to generate negative effects.

For the majority of transactions involving earlier R&D projects, the technology incumbency of the acquiring firm is the key driver of the innovation effects. Early acquisitions by firms with no prior activity in antidiabetics R&D projects result in negative outcomes for the target projects, as such projects are significantly more likely to get stuck (more likely termination and less likely progress). Their underlying technology is also less likely to be further developed (less follow-on patenting). In contrast, for firms with prior activity in antidiabetics R&D, we see substantial differences depending on the technology closeness and product market overlap between the acquirer's and target's antidiabetics projects. We find positive innovation effects for a subset of transactions. These occur when acquirers are both technology and product market incumbents, and where target and acquirer projects are close both in technology and product markets. Not only target projects but also acquirer projects in these deals are significantly less likely to be terminated compared to a no-deal scenario. Furthermore, both the acquirer's and target's projects are experiencing more follow-on patenting compared to all other types of deals. These findings suggest that the positive innovation effects are likely to arise from synergies and two-way spillovers between the target and the acquirer. The closeness in technology space and product markets thus seem to create room for positive innovation effects, rather than for negative killing

[3] A mechanism of action determines for each drug the biochemical process through which it produces the desired effect in the body. MoA is central to R&D development, as clinical trials are centred around the drug's mechanism of action, combined with the therapeutic area. MoA also plays important role in the definition of the relevant antitrust markets, as drugs with the same MoA are typically seen as substitutable.

[4] Our robustness checks also consider alternative measures of closeness in the technology space, namely backward patent citations and Jaffe similarity based on the patent classification.

[5] This general pattern is consistent with the findings of Szücs (2014) who finds that whilst acquirers still pursue their R&D, they prefer to exploit rather than explore the targets' R&D stock.

motives.[6]

The rest of this paper is organized as follows. Section 3.2 provides more details on the research questions and related literature. Section 3.3 presents the data sources, including the explanation of the construction of the relevant sample. Section 3.4 presents the empirical implementation, motivation, and definition of all relevant variables. Section 3.5 presents the results and section 3.6 discusses the findings, contributions and concludes.

3.2 Research questions and related literature

On the one hand, a growing body of research in industrial organization has theoretically (e.g. Motta & Peitz, 2020; Motta & Tarantino, 2021; Norbäck *et al.*, 2020; Gilbert, 2018; Federico *et al.*, 2017, 2018, 2020) or empirically (e.g. Affeldt & Kesler, 2021a,b; Argentesi *et al.*, 2021; Cunningham *et al.*, 2021; Gautier & Lamesch, 2021) studied how product market incumbency and potential cannibalization concerns affect innovation outcomes from acquisitions, particularly in the pharma and the tech sectors.[7]

On the other hand, the literature in innovation, management, and finance has highlighted the need to account for technology motives to assess the effects of acquisitions. For instance, Cloodt *et al.* (2006) and Ahuja & Katila (2001) empirically analyze the role of technological vs. non-technological acquisitions in innovative industries. While both find positive effects of technological acquisitions on innovative output, Ahuja & Katila (2001) find no significant effect of non-technological acquisitions and Cloodt *et al.* (2006) find a negative impact on innovative performance after the merger for the non-technological acquisitions.[8] Apart from technology motives, this literature has also highlighted the need to account for relatedness between the target and the acquirer. Ornaghi (2009a) finds for the pharmaceutical sector that firms' product market relatedness has a positive effect on the post-merger outcomes while technology relatedness seems to have a detrimental impact. Bena & Li (2014) find that acquirers with prior technological linkage to their target firms produce more patents afterwards and conclude that synergies obtained from combining innovation capabilities are important drivers of acquisitions. Meder (2015) finds a positive effect on the number of development projects in postmerger periods in submarkets where both merging firms overlap in terms of their knowledge and development activities.

In line with Cassiman *et al.* (2005), we argue that a comprehensive assessment of the effects of acquisitions of small targets by large incumbents on innovation and long-run competition requires combining insights from both streams of literature and considering

[6] In our sample, none of the product market incumbents behind these effects is diabetes market leader with the largest market power (Sanofi, Eli Lilly, Novo Nordisk and Merck & Co.). These effects are thus fully driven by other incumbent firms. This result fully aligns with the findings of Malek *et al.* (2021) who find that market leaders are on average less likely to be engaged in acquisitions in the first place.

[7] For a broader discussion of factors affecting innovation post-M&A, see the review by Jullien & Lefouili (2018).

[8] Examples of other work on the relevance of technology in M&As include Andersson & Xiao (2016); Wagner (2011); Frey & Hussinger (2011).

technology as well as product market dimensions. Hence, in this paper, we study whether acquisitions of small targets *spur* or *stifle* innovation, depending on the acquirer's R&D and market incumbency as well as on the closeness between acquirer's and target's R&D projects in the technology space and in the future product markets.

We expect "killer" motivations for acquisitions to more likely occur not only from market incumbent acquirers but also from technology incumbents, particularly when the target would hold technologies of relevance for future markets which are in the scope of interest of the incumbent. At the same time, we expect more positive innovation effects for target projects when they can benefit from the acquirer's expertise and resources acquired from their experience. These effects are more likely when there is a higher overlap in technology portfolios between a target and an incumbent acquirer. Technology relatedness creates more room for synergies and/or technology spillovers from and to other projects within the acquirer's portfolio, generating not only benefits for the target projects, but also the twin acquirer's projects. A lack of market or technology incumbency or a large technology distance between the target and acquirer would suggest accessing new markets or technologies as motives for acquisition. For these types of transactions, we expect fewer killer motives, but also less scope for positive synergies and spillover effects.

We look at both early and late (Phase III) stages of technology development in which projects were subject to a M&A transaction.[9] This allows us to control for the exogenous risk of failure, which varies across technologies and MoAs but is particularly high in the early stages of research and development. Although research remains silent regarding the effects of earlier transactions where uncertainty to reach a market is still very high, these early transactions may impact the technology and product market landscape substantially. We expect that outcomes in these stages are driven particularly by the technology dimensions of projects and not by product portfolio cannibalization concerns, since reaching the product market is still very far off and uncertain.

3.3 Data sources and sample construction

3.3.1 Data sources

Our analysis requires rich data about each project that has been in development to treat diabetes. We require information on project's changes of ownership, progress through development (including information on launches, if any), and finally information on technologies covering each diabetes R&D project. A multitude of sources was connected to bring such data together. The backbone of our dataset is the Pharmaprojects database from Citeline, providing a list of global R&D project activity in the pharmaceutical in-

[9] We choose this cut in the data since the difference in the uncertainty of projects at this stage is the largest. While the probability to launch a Phase III project reaches almost 50%, the probability to launch Phase II project reaches only 15%. The probability that a drug does not transition to the next phase is the highest when in Phase II clinical trials (Veugelers *et al.*, 2020).

dustry and several basic variables, such as a brief description of the project, information on the project's therapeutic area and MoA.[10] After collecting all relevant information, our sample includes 1345 projects relating to diabetes, developed between 1997 and 2017.[11]

To identify which projects changed ownership and when we unwound the histories of ownership for each project in our database. We first performed a fuzzy string matching with private merger databases Zephyr and SDC Platinum to search for mergers and acquisitions (M&A) and sales of individual assets or product lines. This was complemented by text analysis of projects' descriptions and eventually verified manually. After identifying all M&A deals affecting projects in our sample, we excluded conglomerate mergers between big pharmaceutical companies as these covered a multitude of therapeutic areas and were unlikely to be specifically related to antidiabetics R&D. We also excluded deals where the targets are large firms, representing divestitures or spinouts. These deals are typically driven by different motives orthogonal to the motives for acquisition, such as killing motives or search for synergies or access to know-how. Finally, we only kept targets which are small or pipeline companies.[12]

To see a project's progression through the various development stages and whether and when termination occurred, we mapped every project to the relevant clinical studies and extracted start and end dates for every clinical trial phase.[13] To identify the technological content of each project, we developed an algorithm assigning patents to projects based on their underlying biochemical properties. For each project in development, we searched for patents that were filed at the USPTO by its developers between the inception and termination dates of a project.[14] Depending on the information available, we complementarily used various techniques to establish patent-project links. For chemical drugs, we employed several cross-roads between chemical, patent, and medical databases. For drugs relying on therapeutic proteins, we followed the approach of Sampat & Williams (2019) and linked gene identifiers from the Pharmaprojects database to a list of protein and nu-

[10]This database is considered the most comprehensive on the market as it collects information from multiple sources including company's press releases, media coverage, patent filings, conference proceedings, regulatory bodies' reports, medical literature as well as direct contact with company representatives and researchers.

[11]The Food and Drug Administration Modernization Act required firms to publish information on clinical trials in the registry from 1997 onwards. We are thus able to track the progress in development from this year onward.

[12]Pipeline companies have no launched drugs in any therapeutic area in the entire pharmaceutical industry. Small firms can have launched drugs but the share of their revenues on the revenues of the entire pharmaceutical industry has not exceeded 1% during the sample period. The revenue information is based on the R&D Scoreboard data published by the European Commission.

[13]This is based on the AACT database (accessible here https://aact.ctti-clinicaltrials.org/) using public data from the US clinical trials registry ClinicalTrials.gov. Our algorithm assigns studies to projects based on matching the sponsor to the primary developer(s) and based on matching the drug names published both in the clinical studies and the Pharmaprojects database. In cases where we were not able to establish full project histories, e.g. due to missing data, we imputed the missing dates by estimating the log-normal distribution of durations per phase and drew randomly a project's phase duration from the estimated distribution. For each such imputation, we manually checked that the sequence of development milestones was not violated and dropped the observation if it was.

[14]The US market for pharmaceutics is the most important geographical market worldwide by volume. A patent at USPTO is therefore critical for appropriating the commercial returns for developed drugs.

cleotide sequences and then established patent links by matching these against the census of human genome sequences disclosed in the US patents. To complement these approaches and increase the matching rates, we also used natural language processing methods and data from Arts *et al.* (2021) to connect a project to patents based on keywords relating to their mechanism of action. Combined with additional manual checks of the unmatched entries, this algorithm allowed us to match patents to 80% of the projects in our sample. A more detailed description of the algorithm and its results is given the Appendix 2.C.

3.3.2 Sample construction

To identify projects relevant to our analysis, we look at all 120 M&A transactions in our sample. These cover 157 antidiabetics R&D projects which changed hands ("*target*" perspective). For each of these projects, we know the development stage in which the transaction happened (Preclinical, Phase I, Phase II, or Phase III), yielding project-phase observations. For these 120 transactions, we also extract information on the antidiabetics R&D portfolio that the acquiring companies owned at the time of the transaction, if any ("*acquirer's*" perspective).

A concrete example illustrating this approach is the acquisition of Kos Pharmaceuticals by Abbott in the first half of 2007. At the time of the acquisition, Abbott had three antidiabetic in-house projects in its portfolio (all in the preclinical phase) and acquired one antidiabetic project from Kos Pharmaceuticals (in Phase II). Hence, this transaction brings 4 observations into our database - 3 from the acquirer and one from the target. Overall, our sample contains 157 antidiabetics target projects and 185 acquirer's antidiabetics projects. In total, we thus have 342 treated observations for our analysis.

3.4 Assessing innovation effects of M&As

To assess the innovation effects of M&As, we employ a simple regression framework measuring the innovation outcomes of the projects from the respective treatment groups (targets/acquirer's/all) compared to their relevant control groups. The canonical regression equation has the following form:

$$Outcome_{ip} = \beta_p \cdot f(Treated_{ip}) + \gamma_p FE_{ip} + \epsilon_{ip}, p \in \{A, T, A+T\}. \tag{3.1}$$

The observation i corresponds to a project-phase and p represents the acquirer (A), target (T), or all (A+T) project sets, respectively. The linear function $f(.)$ and the corresponding vector of coefficients β is the treatment effect of the main interest. As will be further detailed below in section 3.4.1, $f(.)$ encompasses variables for various cuts of the treated sample, allowing us to study the heterogeneity in the treatment along various product and technology market positions of the involved parties. By default, robust

standard errors are used for statistical inference.[15]

We measure how M&As affect three different project-level outcomes, all aimed at capturing various aspects of innovation. *Termination* captures occasions when a firm has abandoned the development of an antidiabetics project. Since firms do not publish information on intentional terminations of specific projects (which would mirror killing), we proxy such situations by capturing cases which hint at a lack of development activity. In particular, we classify a project as terminated if it has not continued in development to the next phase and stayed in a phase exceptionally long – i.e. project's phase duration is beyond the mean plus one standard deviation of a typical phase duration.[16]

The variable *Progress* signals whether a particular project has been actively and successfully pursued. Using information retrieved from the clinical trials databases, this binary indicator equals one if a project has successfully transitioned into the next development phase from a phase in which it was acquired (for the target) or from the phase in which it was affected by a transaction (acquirer) and equals zero if a project stayed in the same phase.

As a further measure of innovation success, beyond progress in clinical trials, we also look at whether a project leads to further technological developments. To assess the impact of M&As on additional innovation, we use all patents assigned to each project along all phases of development. For the treated projects with a M&A transaction, the variable *New Patents* is a binary indicator that equals one if a project received new patent(s) matched to that project and granted to the company after the date of the transaction.[17] For the non-treated projects in our sample, we cannot distinguish periods before and after a transaction. Therefore, for all control projects, the *New Patents* equals one if a project received new patent(s) matched to that project and granted to the developer in the phases following the phase of the transaction of the project to which it is matched.

The nature of the binary outcomes would warrant estimation of the equation 3.1 using logit or probit models. However, due to our sample size and the need to accommodate a substantial number of fixed effects as outlined below, we often face issues with the convergence of these models. In addition, the heterogeneity in M&A treatment captured by interactions of the M&A indicator with independent variables is of the main interest of our analysis. It is though difficult to understand and interpret non-linear models with interaction terms.[18] For all these reasons, we resort to estimation using ordinary least squares and use this consistently across all specifications to report results.

[15] Two robustness checks performed in the results section indicate that our findings remain robust when various forms of correlation are allowed between errors and clustering at various levels is used.

[16] These values are based on the estimated phase distributions of the progressing projects. The phase-specific distributions and cutoff values are presented in the Appendix 3.A in figure 3.A.11.

[17] For the acquirers, the post-acquisition entity corresponds to the acquiring firm. For the target, the post-acquisition entity corresponds to the combined entity (either the acquiring company or the target company)

[18] For example, (Ai & Norton, 2003, p.129) mention that "the interaction effect ... cannot be evaluated simply by looking at the sign, magnitude, or statistical significance of the coefficient on the interaction term when the model is nonlinear".

3.4.1 Heterogeneity in innovation effects

As pharmaceutical M&A deals are very diverse in terms of parties involved,[19] we expect considerable heterogeneity in innovation outcomes as outlined in section 3.2. This section formalizes the product market as well as the technology market positions of the involved parties.

Product market incumbency. Targets acquired by firms that are already active in the product market can benefit from the complementary assets that the incumbents may hold to bring their R&D to commercial success, most notably their (financial) resources, production and commercialization capabilities, experience, or brand name (Arora *et al.*, 2009; Andersson & Xiao, 2016; Grabowski & Kyle, 2008). On the other hand, incumbents can be driven by protecting their existing market positions - preemptively acquiring targets with disruptive R&D projects which could eat into their existing markets to prevent them from being successful (Federico *et al.*, 2020) Indeed, Cunningham *et al.* (2021) empirically document the latter concern that incumbents acquire a potential competitor with an innovative project under development and discontinue the development of the target's innovation, the "killer" acquisition case. To quantify incumbency, we consider the acquiring firm to be a *market incumbent* if it had at least one launched antidiabetic project at the time of the transaction. In the opposite case, it is considered a *market non-incumbent*. Only a selection of big pharmaceutical companies in our sample is classified as market incumbents, namely Bristol-Myers Squibb, GlaxoSmithKline, Merck KGAA, Takeda, Pfizer, Roche, Merck & Co., Sanofi, and Novo Nordisk, with the last three being the "leaders" in the diabetes market.[20]

R&D incumbency. The ability of a firm to pursue any innovation in the pharmaceutical industry depends on its existing experience with a given technology - its R&D incumbency (Arora *et al.*, 2009; Abrantes-Metz *et al.*, 2004; Adams & Brantner, 2003). After an acquisition, a firm might simply fail to develop an innovation into a successful product because it has not worked in such an area before, rather than due to strategic motives. The industrial organization literature studying the effects of M&A seems to be silent about this important driver.[21] Given the clear delineation of R&D activities in our dataset, we can directly measure R&D incumbency depending on the R&D position of the acquiring firm in the antidiabetic market before a particular transaction. To study the effects, we thus split the treated projects into two exclusive groups and define two treatment variables - one for the projects acquired by *R&D incumbents* (firms with an existing portfolio of antidiabetic R&D projects before a transaction) - and one for the projects acquired by *R&D non-incumbents* (firms with an empty portfolio of antidiabetic

[19] Please see Malek *et al.* (2021) for a detailed overview of the literature on which firms are more likely to be engaged in M&A deals.

[20] The leader's market shares in antidiabetics exceeded 10% and they jointly accounted for 13% of all R&D (Malek *et al.*, 2021).

[21] One of the few related studies differentiating between incumbents and non-incumbents depending on their innovation activities is Czarnitzki & Kraft (2010) which focuses on firm's licensing decisions.

R&D projects).[22]

Product market and technology market incumbency correlate but are not the same. All product market incumbents are R&D incumbents, but not all R&D incumbents are product market incumbents. Accordingly, we define three groups: R&D incumbents but not (yet) product market incumbents, R&D incumbents who are product market incumbents, and R&D non-incumbents.

Closeness in future product markets. The industrial organization literature has so far predominantly analyzed the role of the closeness of competition on innovation by looking at the closeness of portfolios between acquirers and targets at the firm level (See for example Meder, 2016; Yu *et al.*, 2016; Bena & Li, 2014; Hoberg & Phillips, 2010). Apart from Cunningham *et al.* (2021), it has not accounted for closeness at the individual product/project level as is standard in antitrust proceedings. Our dataset allows overcoming this issue. We measure the closeness between the antidiabetic R&D projects of the acquirers and targets depending on whether they had an overlap in the "mechanism of action" (MoA). The mechanism of action determines the biochemical process for each drug through which it produces the desired effect in the body. As such, MoA is central for the R&D development (Nat. Med, 2010) as an overlap implies that firms can build on their experience to run clinical trials for familiar MoAs. MoA is also a central feature in the product market space, as launched drugs with the same MoA are often regarded as substitutes by patients and physicians, thereby delineating the boundaries of the (future) relevant product markets.[23] Thus, MoA overlap also potentially results in competition between the projects in the (future) product market space. To study the role of MoA overlap in our analysis, we further split the treated projects acquired by R&D incumbents into two exclusive groups, depending on whether a particular acquirer (target) project shared an MoA with the target (acquirer) projects, or not. Since the R&D non-incumbent acquirers have an empty antidiabetics portfolio before a transaction, an MoA overlap is undefined for these.

Technological closeness. To fully understand the effects of M&A on innovation in highly innovative industries requires looking beyond closeness in the product market space and considering the role of technology closeness. MoA may imply technology closeness to some extent but still leaves heterogeneity in terms of technologies which share the same MoA. Our paper extends this literature significantly by using a detailed measure of technological closeness at the project level.

We measure the technology closeness between acquirer's and target's projects using a measure based on the text of the patents we have assigned to projects. Specifically, based

[22]The portfolio of antidiabetic R&D non-incumbents can be empty since projects in other therapeutic areas the company may hold are not considered in our analysis.

[23]Drugs with different MoAs have distinct efficacy and safety profile, which are the key factors for physicians when prescribing drugs (M.9461 - Abbvie/Allergan, para 51). Examples of antitrust cases where relevant markets were defined using the combination of the therapeutic area and project's mechanism of action include M.9274 - GSK/Pfizer Consumer Healthcare Business, M.7275 - Novartis/GSK Oncology, M.8955 - Takeda/Shire.

on the database of Arts *et al.* (2021), we obtain precleaned and standardized keywords from abstract, title and claims of US-granted patents. This captures the most relevant parts of each patent. Pooling keywords from all patents relating to a particular project, we can compute pair-wise cosine similarity between all projects in the database according to the following formula:[24]

$$Cosine(A, B) = \frac{A \cdot B}{\|A\| \cdot \|B\|}. \tag{3.2}$$

This results in more than 2 million project pairs, with each having a corresponding similarity value. We use this measure to identify whether a particular target (acquirer) treated project was technologically close to an acquirer (target) project (in which case it is called a *Twin*), or not (in which case it is called a *Sibling*). An acquirer-target pair is technologically close (a *Twin*) if it scores for technological similarity higher than 0.16, representing the 95th percentile of the distribution of technological similarity between all possible observation pairs in our dataset.[25] We verify the sensitivity of our results using different cutoff values.

Alternative measures of technological closeness. Our similarity measure, being based on the text of the patents associated with the projects, closely reflects the projects' technological nature. Such level of detail is an obvious advantage of this relatively new measure (Arts *et al.*, 2018, 2021). In this section, we present two more aggregated but also more commonly used closeness measures based on patent citations and patent classification (Aharonson & Schilling, 2016). This serves to check the robustness of our results and to illustrate whether more commonly used measures with fewer data-intensive requirements can be used in practice to replicate the most important findings.

As a first alternative measure of technological closeness, we employed backward patent citations. It is well established that patent citations are only a noisy proxy of the relevance of the knowledge disclosed since the citation of prior art might be a result of strategic behaviour by firm's lawyers or by patent examiners (Alcácer *et al.*, 2009; Cotropia *et al.*, 2013). The citation process was never designed to represent a taxonomy (Aharonson & Schilling, 2016). However, patent citations are readily available and this has made them popular in the literature (See e.g. Bena & Li, 2014; Ornaghi, 2009b). As an alternative to our twin binary variable, we create a binary indicator equal to one if the acquirer's (target) projects made at least one backward reference to the patents attached to the target (acquirer's) projects.

Another alternative measure we employ relies on the well-established classification system of patent offices. The literature has already shown that this aggregated classification

[24] Cosine similarity is a measure between zero and one, with zero meaning that A and B are orthogonal, and one meaning that A and B are identical.

[25] R&D non-incumbent acquirers have an empty antidiabetics portfolio before a transaction and by definition have neither twin nor sibling projects. A detailed explanation with an example of the construction of the technology variables is given in figure 3.A.11 and in Appendix 3.A.

system might not capture all the technological characteristics of an invention. Moreover, different classes and subclasses might contain significant overlap so that technologically similar patents can have a different classification (Arts *et al.*, 2018).[26] We use the IPC patent classifications and compute a cosine similarity between any two projects using the patent counts in the various technology classes (IPC) they cover. This is a standard measure in the innovation literature (See e.g. Jaffe, 1986; Ahuja & Katila, 2001; Ornaghi, 2009b; Bena & Li, 2014). In line with the text-based measure, we chose a 95th percentile threshold to distinguish between technologically close and distant projects.

3.4.2 Control groups

To establish innovation effects from M&As, observations exposed to an ownership change need to be compared to a no-deal counterfactual (control group). In an ideal setting, we would have an experiment with firms acquiring projects at random - leaving the probability of the treatment to be exogenous. This would allow us to isolate the effects of acquisition cleanly - holding all other factors constant, all observed differences would capture the effect of the treatment (M&A) only. With observational data, such an approach is not feasible. In addition, we are not aware of a quasi-experiment that would exogenously shift the transaction propensity at the level of individual projects, allowing an alternative way of isolating the effects.[27] Given these limitations, we resort to the approaches most used in other empirical studies investigating the effects of mergers at the level of firms.[28]

In all specifications, we employ a set of fixed effects FE to capture invariant characteristics of projects that need to be controlled for as they may confound the analysis of effects from M&A analysis. The first set of fixed effects controls for the differences across the various development phases of the projects, most notably their differences in technology uncertainty and risks of failure, as discussed above. The second set of fixed effects is cohort fixed effects which group projects initiated around the same time. This controls for time trends and technological trends. The third set of fixed effects controls the number of patents of each project.[29] This is partly a technical control, as a higher number

[26] Arts *et al.* (2018) show that text-matched patents are more likely to cite each other, belong to the same patent family, or have a common inventor or assignee compared to patents that are matched based on rougher patent classification.

[27] A reform that has been extensively used as a source of a quasi-random variation providing a demand shock in the pharmaceutical industry is Medicare Part D. However, this reform is not suitable for our purposes. It is not clear whether it affected the merger propensity and did so at the level of individual projects. Moreover, it only provides a static change, whereas our M&A events are scattered over 20 years.

[28] For example, to find an appropriate control group, Gugler *et al.* (2003) use forced-matching and only consider the non-treated firms belonging to the same industry(ies) as that (those) of the two merging firms as controls. In cases more observables are used to find a control group, propensity score matching has been frequently employed. Szücs (2014) match each treated firm to a control firm based on accounting data (total assets, income, total sales, total debt, number of employees, firm age and R&D expenditures). Also Ornaghi (2009a) uses propensity scores - computed based on patent expiration, R&D performance, and other variables. Stiebale & Szücs (2019) and Bena & Li (2014) use propensity score with the combination of forced matching.

[29] To limit the potential influence of outliers, we have discretized the number of project's patents.

of patents will affect the likelihood of scoring on technology closeness. The breadth and depth of technological know-how of a project may also affect the stewardship of projects towards success. It allows us to better sort out the effect of incumbency and relatedness from size. The last set of fixed effects relates to the projects' Mechanism-of-Action (MoA). This allows us to control for the time-invariant, distinctive technology and market features specific to each group of drugs within the same MoA.[30]

In the pharmaceutical setting, the assumption that technologically similar projects follow a similar innovation trajectory is natural. Minimizing the observable technology differences between the treatment and control groups is thus critical to ensure that absent the acquisition, treated projects would have followed a development trajectory of the counterfactual group. Although this is to some extent controlled for by the above-fixed effects, particularly the MoA fixed effects, there can still be differences in the technology spectrum within MoA. Thus, we also look at controls that are closer in technology space than being within the same MoA. To this end, we use the technology information we have available for the projects through patent documents. Similarly to our procedure for identifying technology twins, we build on patent keywords as developed by Arts *et al.* (2021) to identify similarity in the technological content of an invention. For each treated project, we search the technologically closest, never-treated counterpart not belonging to the same company and within the same development phase, giving us a set of 342 control observations.

Section 3.C presents results from a placebo exercise to see whether our matching strategy yields valid controls (ie projects in the original and matched samples follow the same development trajectory). We use the sample of never treated observations and for each observation find one which is in the same phase and is technologically closest. We then compare the difference in the outcomes between the two pooled samples. The results show that without our set of other FE controls, the matches are not significantly different on *new patents*, but they are on *progress* and *termination*. However, together with our set of other FE controls, differences in *progress* are no longer significant and on *termination* drastically reduced, the latter nevertheless remaining statistically significantly different. We, therefore, use this matched sample in two robustness exercises. First, we restrict our sample to the 342 treated observations and their 342 matched control observations, leaving a pooled sample of 684 observations to run our econometric specifications, including the standard set of fixed effects. Second, we in addition include a set of 342 treated-control pair dummies, which allows us to compare each treated project to its matched counterfactual directly (this approach corresponds to one-to-one forced matching).

We created four binary variables, grouping projects with 1, 2-4, 5 -10, and more than 10 patents. These cutoffs were chosen to guarantee enough observations per group. Results are fully robust to alternative cuts.

[30] We also included fixed effects for technology classes (IPC) of the projects' patents. However, these did not give additional effects over MoA fixed effects, while consuming many degrees of freedom. The reported results, therefore, do not include IPC fixed effects.

3.5 Results

3.5.1 Descriptive evidence

Before presenting our econometric results, we first provide some descriptive evidence. To do so, we compare our 342 acquirer's or target projects subject to M&A transactions with their counterfactuals, i.e. the set of never treated, most technologically similar projects not from the same firm and in the same phase, as described in the previous section. Table 3.1 below summarizes for our full sample the means of the outcome variables for both target and acquirers' projects, compared to their control projects.

These descriptives already suggest that significant differences in innovation outcomes exist between treated projects and their controls, suggesting negative innovation effects from acquisitions: the progression rate for the treated projects is lower, and their termination rate is higher. The likelihood to receive new patents post-acquisition is lower for treated projects compared to their counterfactuals. These negative results are mostly driven by target projects. There are few significant patterns on the side of the acquirers' projects, except for the latter to be also more likely to be terminated.

Only a small minority of our cases are in Phase III observations (46), the last stage before an application for launch can be filed, and the stage which is most on the radar of competition authorities. Table 3.1 shows that these cases display different innovation effects from the overall sample. The negative innovation effects results found overall relate to transactions taking place in the early phases of development. In contrast to these negative effects, Phase III target projects (10 observations) are not significantly different from their non-treated controls in terms of termination and new patents, but they are significantly more likely than their controls to progress to the next stage, applying for market launch. As negative innovation effects seem mostly to occur in the early stages before Phase III, ignoring these projects may risk missing out on cases with important long-term effects on antidiabetic therapeutics.

In what follows, we focus our main analysis on the sample of earlier development stages (ie preclinical, Phase I and Phase II), dropping phase III projects (and their corresponding counterfactuals).[31] Table 3.D.3 in Appendix 3.D presents splits of the treated sample (together with the corresponding frequencies) that we consider in our econometric effects analysis. It shows that product and technology market characteristics are correlated, but that there is nevertheless sufficient off-diagonal variation, advocating for considering both dimensions in the assessment. For about 40% of the treated projects, the R&D incumbent acquirers were not product market incumbents. For R&D incumbent acquirers, the closeness between target and acquirer is sufficiently distinct in the product and technology dimensions. About 38% of the treated projects with R&D incumbents involve cases with

[31] The econometric analysis based on the equation 3.1 confirms that phase III target cases, albeit displaying a high variance in results, have a significantly higher probability to progress. Full regression results can be found in table 3.E.4 in Appendix 3.E.

Table 3.1: Summary statistics - M&A vs. non-M&A for Acquirer's, Targets, and Combined (within subgroups)

	Acquirer			Target			All		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Mean No M&A	Mean M&A	p-value	Mean No M&A	Mean M&A	p-value	Mean No M&A	Mean M&A	p-value
Panel A: Termination									
Full sample	0.06	0.14	0.01	0.16	0.39	0.00	0.11	0.26	0.00
Late	0.00	0.00	.	0.00	0.40	0.14	0.00	0.09	0.15
Early	0.07	0.16	0.01	0.16	0.39	0.00	0.11	0.27	0.00
Early: R&D non-Inc				0.08	0.45	0.00	0.08	0.45	0.00
Early: R&D Inc - Twin	0.02	0.13	0.03	0.33	0.19	0.22	0.13	0.15	0.65
Early: R&D Inc - Sib	0.09	0.17	0.07	0.24	0.42	0.09	0.13	0.23	0.02
Early: R&D Inc - MoA	0.00	0.07	0.33	0.36	0.14	0.20	0.17	0.10	0.46
Early: R&D Inc - no MoA	0.07	0.16	0.01	0.25	0.37	0.20	0.12	0.22	0.01
Early: R&D Inc - Twin + MoA	0.00	0.08	0.33	0.45	0.09	0.06	0.22	0.09	0.23
Early: R&D Inc - non Twin+MoA	0.07	0.16	0.01	0.24	0.37	0.15	0.11	0.22	0.01
Panel B: Progression									
Full sample	0.44	0.44	1.00	0.38	0.22	0.00	0.41	0.34	0.04
Late	0.33	0.72	0.02	0.00	0.60	0.04	0.26	0.70	0.00
Early	0.45	0.41	0.44	0.39	0.20	0.00	0.42	0.31	0.00
Early: R&D non-Inc				0.39	0.17	0.00	0.39	0.17	0.00
Early: R&D Inc - Twin	0.45	0.34	0.24	0.44	0.33	0.41	0.45	0.34	0.15
Early: R&D Inc - Sib	0.45	0.44	0.89	0.37	0.18	0.07	0.43	0.38	0.35
Early: R&D Inc - MoA	0.33	0.27	0.70	0.43	0.50	0.72	0.38	0.38	1.00
Early: R&D Inc - no MoA	0.46	0.42	0.49	0.39	0.18	0.02	0.44	0.36	0.09
Early: R&D Inc - Twin + MoA	0.42	0.25	0.41	0.45	0.55	0.69	0.43	0.39	0.77
Early: R&D Inc - non Twin+MoA	0.45	0.42	0.57	0.39	0.19	0.02	0.44	0.36	0.11
Panel C: Patenting									
Full sample	0.25	0.25	1.00	0.38	0.18	0.00	0.31	0.22	0.01
Late	0.28	0.22	0.71	0.60	0.20	0.24	0.35	0.22	0.34
Early	0.25	0.26	0.90	0.38	0.18	0.00	0.31	0.22	0.01
Early: R&D non-Inc				0.37	0.14	0.00	0.37	0.14	0.00
Early: R&D Inc - Twin	0.25	0.32	0.39	0.41	0.37	0.79	0.30	0.34	0.61
Early: R&D Inc - Sib	0.25	0.23	0.64	0.37	0.13	0.02	0.28	0.20	0.11
Early: R&D Inc - MoA	0.13	0.40	0.11	0.50	0.43	0.72	0.31	0.41	0.42
Early: R&D Inc - no MoA	0.26	0.24	0.69	0.35	0.18	0.04	0.29	0.23	0.17
Early: R&D Inc - Twin + MoA	0.17	0.42	0.19	0.55	0.55	1.00	0.35	0.48	0.38
Early: R&D Inc - non Twin+MoA	0.26	0.25	0.79	0.35	0.17	0.03	0.28	0.22	0.18
Observations	370	370	370	314	314	314	684	684	684

This table presents summary statistics for the Progression, Termination, and Patenting outcome variables depending on the M&A status (columns) and project's characteristics (rows). Each of the panels presents decsriptive evidence for one outcome variable. In each panel, mean differences (and the corresponding p-value) for the full sample are presented in the first row. The second and third rows use the subsample of Late or Early treated projects, respectively, and present mean differences between the treated and the counterfactual projects (and corresponding p-value). The rows four to ten in each panel then focus on the heterogeneity in M&A treatment and present descriptives for various types of exclusive M&A treatment categories, comparing them always to the respective matched control group (ie. non-M&A, same-phase, technologically closest matches). This corresponds to the sample of 342 treated and 342 control observations (684 observations in total). Columns (1)-(3) use the subsample of the acquiror's portfolios, columns (4)-(6) show the the targets, and columns (7)-(9) combine both perspectives.

technology twin but only 13% involve cases with MoA overlap. Of the technology twin cases, only one out of four also has an MoA overlap, leaving most of the twin cases to have no MoA overlap. Unfortunately, the number of observations in subcategories becomes very small, which limits the possibilities for segmentation to examine the heterogeneity of effects. For instance, we have only 24 observations of projects involving R&D incumbent acquirers which have both technology twins and MoA overlap, and only 7 observations which have no technology twins, but MoA overlap.

Nevertheless, splitting the sample in table 3.1 into subgroups using technology and product market dimensions shows that the negative innovation effects on target projects are most likely to occur with R&D non-incumbent acquirers. The group of target projects acquired by an R&D incumbent is more likely to generate positive innovation effects, (lower termination rates, higher progression rates, more new patents), but only when the acquirer also has projects which are technology close to the target firm and with an MoA overlap. In these cases, even the acquirers' projects are more likely to have follow-on patents, consistent with synergy effects driving these transactions, rather than killing motives. All this descriptive evidence is already highly suggestive of the heterogeneity in innovation effects along the technology dimensions of the involved parties.

3.5.2 Econometric results

This section presents the econometric results of the paper in several sections, testing equation 3.1, including our full set of controls and using all diabetics projects in our sample (excl Phase III). We first present the overall results, followed by several sections focusing on the heterogeneity behind the main effect. As a robustness check, we then present the same analysis using only the sample of treated projects with their technology closest matches.

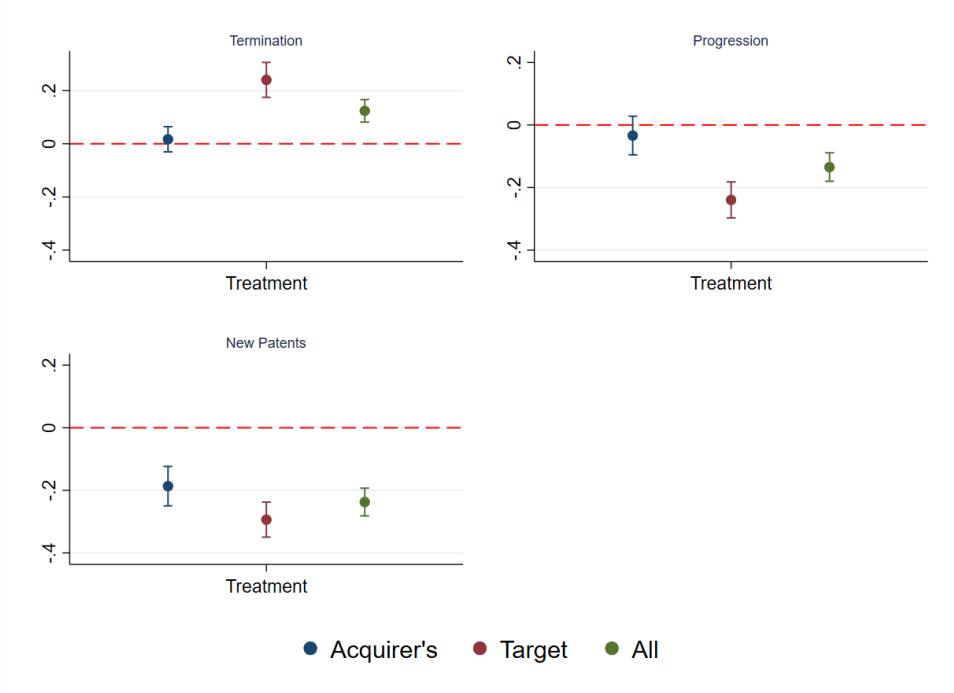
The estimated effects of the treatment (M&A) on R&D antidiabetics projects are presented graphically. The three panels of each figure present impacts on the different project-level outcomes - termination, progression rate to the next development phase, and the likelihood to receive new patents post-transaction. The bars in the figure represent the mean effect with 90% confidence intervals as regards the impact on the acquirer's antidiabetics portfolio (blue), target projects (red) and overall cumulative impact of an acquisition or merger (green). The regression equations underlying the graphs are provided in Appendix 3.E.

Baseline effect of M&A

In line with the descriptive results, figure 3.1 shows that M&As have a negative impact on target projects' innovation outcomes, as they have a significantly lower progression rate, significantly higher termination rate and a significantly lower probability to generate

new technology inventions post acquisition compared to the controls. Also projects of the acquirer are less likely to generate new technology inventions post-acquisition. Taken together, M&As results in unfavourable effects for the projects that are taken over and the transaction overall results in less patenting activity for the antidiabetics R&D.[32]

Figure 3.1: Effects of M&As: Baseline effect



Note: This figure visualizes the estimated effects of M&As compared to all never treated projects (no matching). The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated_i + \epsilon_i.$$

The bars give the mean of the total effect with 90% confidence intervals for the treatment group. The effects on acquirer's projects are plotted in blue, on target projects in red, and all projects in green. The red dashed line highlights a zero effect. Full regression results are presented in table 3.E.5 in Appendix 3.E.

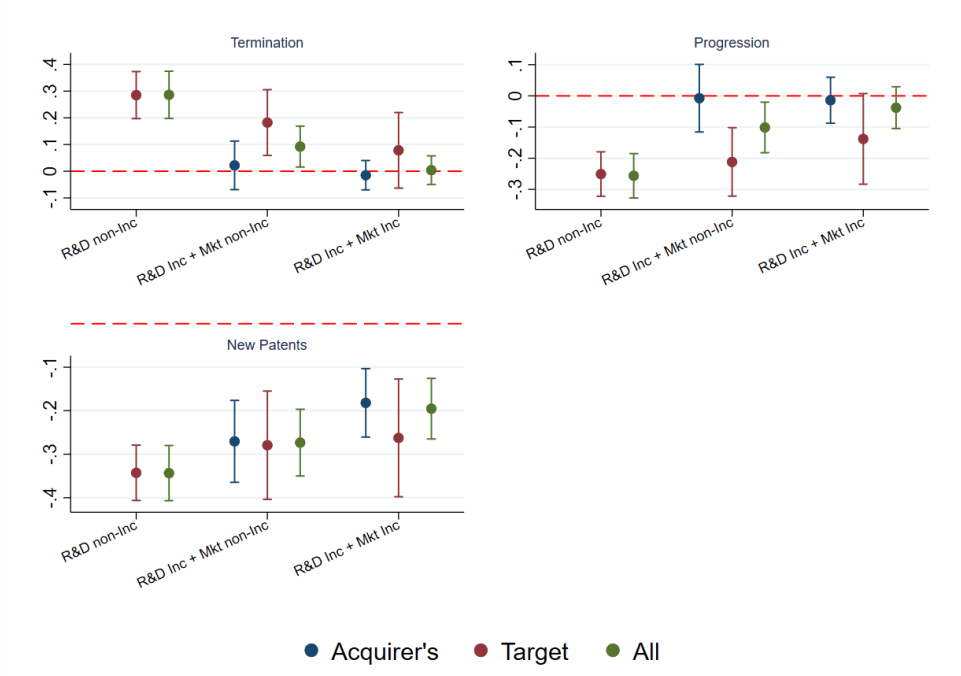
R&D incumbency and Market incumbency

Figure 3.2 presents the results of our analysis when we consider the R&D and product market incumbency of the acquiring firm. As explained in section 3.4.1, R&D incumbency

[32] These baseline results are also fully robust when estimating the effects using logit model as reported in Table 3.E.6 in Appendix 3.E. The magnitudes of marginal effects are also very similar across models. Beyond these three main outcome variables, we have also analyzed the impact on the speed of progression. These results are not reported as we have not identified a pattern beyond the treated projects being significantly less likely to experience fast development in all specifications. This finding remains robust to various specifications of "fast development".

captures whether the acquiring firm has had an active antidiabetics R&D portfolio before it decided to take over a project from another firm, which allows us to study the effects on both acquirer's and target projects. In contrast, R&D non-incumbents obtain their first antidiabetics projects via acquisition, and their preexisting antidiabetics portfolio is empty. This implies that we can only study the effect of such transactions on the target projects that non-incumbents acquired and that the impact of relatedness/overlap between target and acquirer projects cannot be analyzed for this group, but only for the group of R&D incumbents.

Figure 3.2: Effects of M&As: R&D incumbency and market incumbency



Note: This figure visualizes the estimated effects of M&As depending on both R&D incumbency and market incumbency status, compared to all never treated projects (no matching). The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 Treated : R\&DInc_i + \beta_3 MktInc_i + \beta_4 MktInc_i \times Treated : R\&DInc_i + \epsilon_i.$$

The bars give the mean of the total effect with 90% confidence intervals for the respective exclusive treatment groups (β_1 for *R&D non-incumbents*, β_2 for *R&D incumbents and market non-incumbents*, and $\beta_2 + \beta_3 + \beta_4$ for *R&D incumbents and market incumbents*). The effects on acquirer's projects are plotted in blue, on target projects in red, and on all projects in green. Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed. The red dashed line highlights a zero effect. Full regression results are presented in table 3.E.7 in Appendix 3.E.

Figure 3.2 shows that projects acquired by R&D non-incumbents have significantly negative effects on innovation. Their acquired projects are less likely to progress and more likely to be terminated. Also their technologies are less likely to be developed. In contrast, acquisitions by R&D incumbents show less negative effects on target projects and with higher variance, leaving more scope for possible positive cases in this subgroup. Figure 3.2 also shows that it is the R&D incumbency rather than the product market incumbency that makes the clearest cut. Product market non-incumbents are not as likely to generate negative innovation effects as R&D non-incumbents.

Technology closeness and MoA overlap

We next look at the role of closeness between the target and acquirer's antidiabetic projects. Cunningham *et al.* (2021) show that product market incumbents are more likely to discontinue target projects, particularly when the target project overlaps in its MoA with the acquirer's portfolio. In this spirit, this section looks at the role of closeness in the product market space (MoA overlap) but also complements it by closeness in the technology space (technology "twin" measure). This segmentation can only be done for acquirers which are R&D incumbents with their own pre-transaction antidiabetics portfolio.

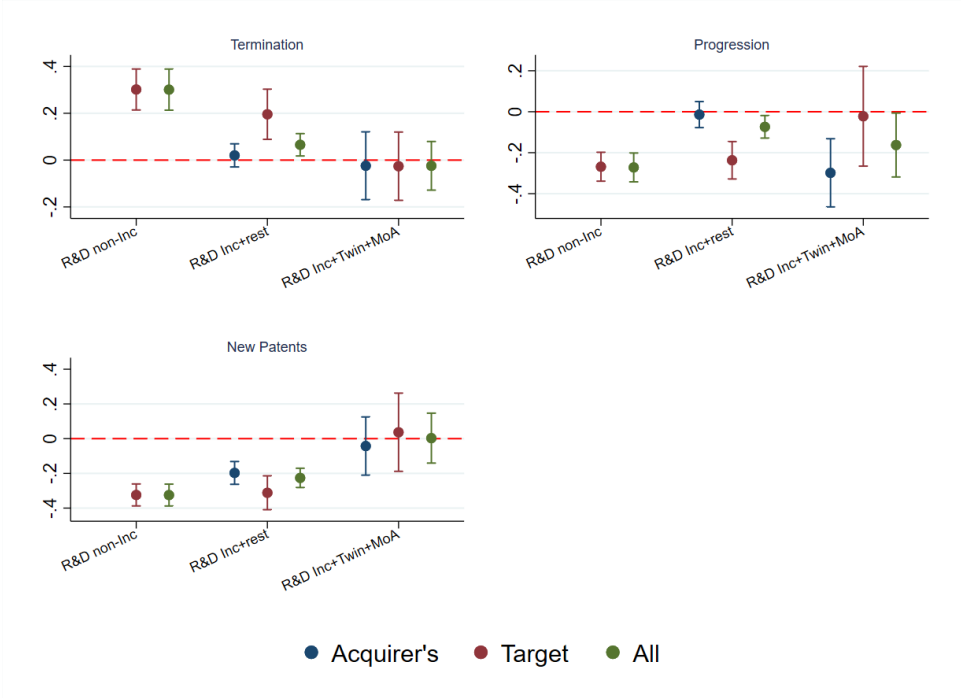
Comparing the econometric results for R&D incumbent acquirers (see tables 3.E.8 and 3.E.9 in Appendix 3.E) shows that the negative innovation effects of M&As on target projects hold most clearly when acquirer and target are technologically distant. In terms of magnitudes, these outcomes are comparable to the effects on target projects acquired by R&D non-incumbents. This holds both if we measure technology distance by whether the parties have technology twin projects, or whether their projects share an MoA. As figure 3.3 shows, only when target projects with R&D incumbent acquirers are sufficiently technologically similar (Twins) and in addition also share MoA, the target progression rates, termination rates and the likelihood to receive new patents are less negative than for R&D non-incumbents or R&D incumbent acquisitions with distant projects. These target projects no longer display significantly negative innovation effects. This suggests that when parties are closer in (technology) market space, the scope for exploiting complementarities/synergies/spillovers to generate positive innovation effects kicks in stronger than the incentive to "kill" the target projects.

Interactions with product market incumbency

To investigate further the trade-off between negative versus positive innovation effects from M&As, figure 3.4 presents the results when the joint measures for closeness are interacted with the product market incumbency status of the acquiring firm.

The figure shows target projects are significantly more likely to progress, less likely to be terminated and more likely to have follow-on developments when the acquirer is an R&D and product market and the acquirer and the target are sufficiently close, i.e. have

Figure 3.3: Effects of M&As: Technological closeness and MoA overlap

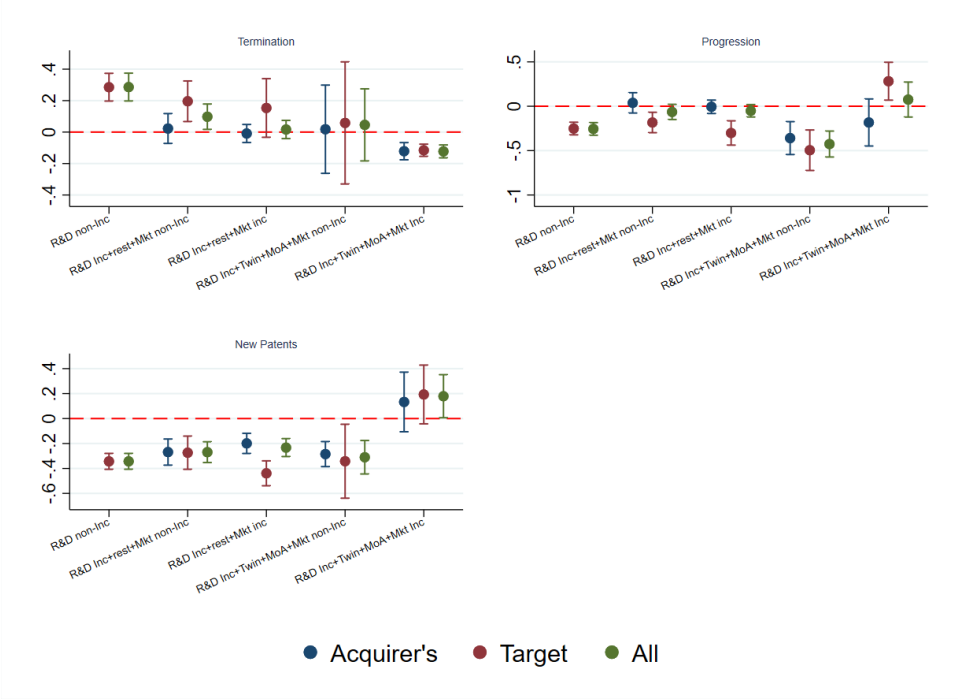


Note: This figure visualizes the estimated effects of M&As depending on the combination of technological closeness and MoA overlap, compared to all never treated projects (no matching). The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \\ + \beta_2 Treated : R\&DIncRest_i + \beta_3 Treated : R\&DIncTwinMoA_i + \epsilon_i.$$

The bars give the mean of the total effect with 90% confidence intervals for the respective exclusive treatment groups (β_1 for *R&D non-incumbents*, β_3 for *R&D incumbents with technologically close projects with MoA overlap*, and β_2 for all other projects of the R&D incumbents, namely *technologically distant projects with or without MoA overlap and technologically close projects without an MoA overlap*). The effects on acquirer's projects are plotted in blue, on target projects in red, and all projects in green. Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed and MoA/Tech closeness measures are undefined (no connection exists between the target and acquirer's antidiabetics projects). The red dashed line highlights a zero effect. Full regression results are presented in table 3.E.10 in Appendix 3.E.

Figure 3.4: Effects of M&As: Technological closeness, MoA overlap, and market incumbency



Note: This figure visualizes the estimated effects of M&A depending on the combination of technological closeness, MoA overlap, and market incumbency compared to all never treated projects (no matching). The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 MktInc_i + \\ + \beta_3 Treated : R\&DIncRest_i + \beta_4 Treated : R\&DIncRest_i \times MktInc_i + \\ + \beta_5 Treated : R\&DIncTwinMoA_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

The bars give the mean of the total effect with 90% confidence intervals for the respective exclusive treatment groups (β_1 for *R&D non-incumbents*, $\beta_2 + \beta_5 + \beta_6$ for *R&D inc and market inc with technologically close projects with an MoA overlap*, β_5 for *R&D inc and market non-inc with technologically close projects with an MoA overlap*, β_3 for all *other projects of the R&D inc and market non-inc*, and $\beta_2 + \beta_3 + \beta_4$ for all *other projects of the R&D inc and market inc*). The effects on acquirer's projects are plotted in blue, on target projects in red, and on all projects in green. Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed and MoA/Tech closeness measures are undefined. The red dashed line highlights a zero effect. Full regression results are presented in table 3.E.11 in Appendix 3.E.

technology twin projects and share an MoA. In addition, these positive innovation effects not only hold for the target projects in these deals, but also for the acquirer's projects. Acquirer's projects are significantly less likely to be terminated and significantly more likely to have follow-on patents compared to all other types of deals. Thus, the positive effects we find are bidirectional between the target and acquirer's projects. These bidirectional effects create the scope for synergies/spillovers when uncertainty in outcomes is still high, dominating any "killing" motives.

Interestingly, these positive effects are not driven by the "leader" acquirers with the largest market power who are less likely to engage in M&A in the first place (Malek *et al.*, 2021). Dropping leading firm acquirers from our analysis does not change our results, confirming that non-leading product market incumbents (eg. Pfizer, Roche, or Takeda) drive these results. This further indicates that for product market incumbents, cannibalization concerns may not be the leading motive behind their acquiring projects which are (technologically) close to their projects and in the same future relevant markets. On the contrary, the technology closeness of acquired projects enables them to leverage technology benefits for the involved projects. These results thus show that it is the four-wise combination of incumbency and closeness, in both product market and technology markets that results in significant positive innovation effects.

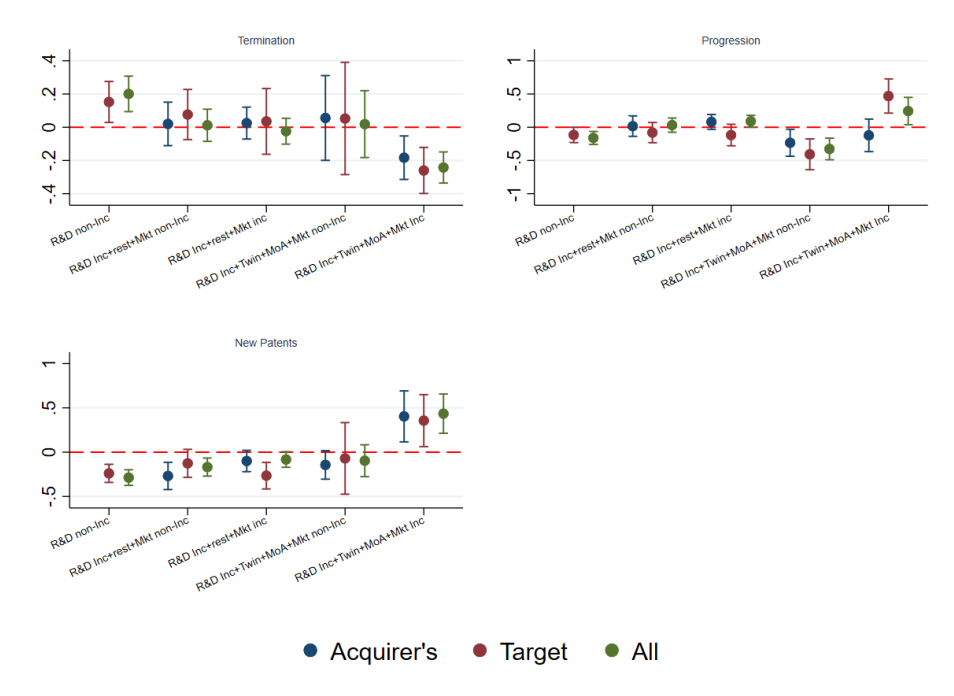
Robustness: Matching with technology similar counterfactuals

In this section we present results when we restrict our sample to the 342 treated observations matched to their technologically closest, never-treated counterpart not belonging to the same company and within the same development phase, leaving a sample of 684 observations to run our econometric specifications.

Without forced matching, but controlling for our set of FEs, our main results remain highly robust. Overall, figure 3.5 shows that target projects are significantly more likely to experience negative innovation effects. They progress less, are terminated more and see a lower probability of follow-on patenting. The negative effects are from projects where the acquirer is an R&D non-incumbent. For projects involved in deals with an R&D incumbent, there is more variance in outcomes and more likelihood for positive effects. Our main result, namely that only cases where R&D incumbent acquirers are also product market incumbents and when target and acquirer's projects are sufficiently technology close and share MoA, remains robust: target projects are more likely to progress, less likely to get terminated and more likely to get follow-on patents. Also the acquirer's projects in these deals are more likely to enjoy positive innovation effects. In sum, all main results are robust with this tighter matching.

Even if we enforce exact matching, i.e. when in addition to our FE set, we include a set of 342 treated-control pair dummies, directly comparing each treated observation with its closest match, our positive innovation results for the cases where R&D incumbent

Figure 3.5: Robustness: Effects of M&As: Technological closeness, MoA overlap, and market incumbency with pooled matching



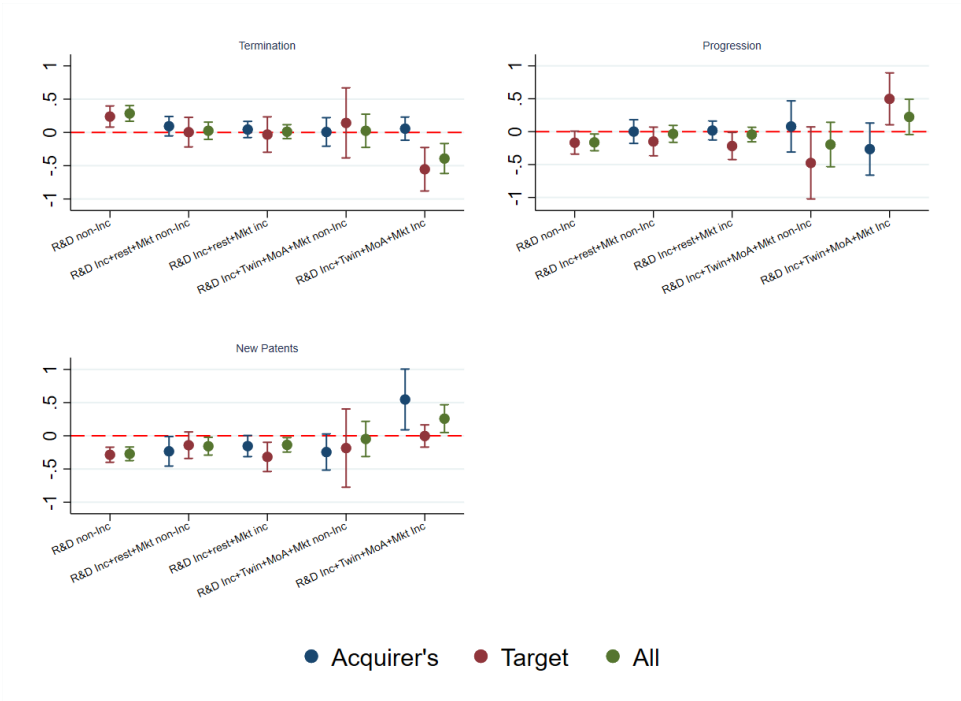
Note: This figure visualizes the estimated effects of M&A depending on the combination of technological closeness, MoA overlap, and market incumbency compared to the pooled group of never treated counterfactuals (not the same firm but the same phase and technologically closest). The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 MktInc_i + \\ + \beta_3 Treated : R\&DIncRest_i + \beta_4 Treated : R\&DIncRest_i \times MktInc_i + \\ + \beta_5 Treated : R\&DIncTwinMoA_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

The bars give the mean of the total effect with 90% confidence intervals for the respective exclusive treatment groups (β_1 for *R&D non-incumbents*, $\beta_2 + \beta_5 + \beta_6$ for *R&D inc and market inc with technologically close projects with an MoA overlap*, β_5 for *R&D inc and market non-inc with technologically close projects with an MoA overlap*, β_3 for all *other projects of the R&D inc and market non-inc*, and $\beta_2 + \beta_3 + \beta_4$ for all *other projects of the R&D inc and market inc*). The effects on acquirer's projects are plotted in blue, on target projects in red, and on all projects in green. Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed and MoA/Tech closeness measures are undefined. The red dashed line highlights a zero effect. Full regression results are presented in table 3.E.12 in Appendix 3.E.

acquirers are also product market incumbents and when target and acquirer's projects are sufficiently technology close and share MoA remain robust (Figure 3.6). The positive effect on the likelihood to generate follow-on patents continues to hold significantly only for acquirers' projects. For the target projects, these results are however no longer significantly positive but the combined effect on follow-on patents still remains significantly positive.

Figure 3.6: Robustness: Effects of M&As: Technological closeness, MoA overlap, and market incumbency with exact matching



Note: This figure visualizes the estimated effects of M&A depending on the combination of technological closeness, MoA overlap, and market incumbency compared to the never treated counterfactual counterpart (not the same firm but the same phase and technologically closest). The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 MktInc_i + \\ + \beta_3 Treated : R\&DIncRest_i + \beta_4 Treated : R\&DIncRest_i \times MktInc_i + \\ + \beta_5 Treated : R\&DIncTwinMoA_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

The bars give the mean of the total effect with 90% confidence intervals for the respective exclusive treatment groups (β_1 for *R&D non-incumbents*, $\beta_2 + \beta_5 + \beta_6$ for *R&D inc and market inc with technologically close projects with an MoA overlap*, β_5 for *R&D inc and market non-inc with technologically close projects with an MoA overlap*, β_3 for all *other projects of the R&D inc and market non-inc*, and $\beta_2 + \beta_3 + \beta_4$ for all *other projects of the R&D inc and market inc*). The effects on acquirer's projects are plotted in blue, target projects in red, and all projects in green. Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed and MoA/Tech closeness measures are undefined. The red dashed line highlights a zero effect. Full regression results are presented in table 3.E.13 in Appendix 3.E.

Robustness: Including back Phase III projects

When including Phase III projects, the noisiness of the results slightly increases in line with the different effects pattern of the Phase III projects (Figure 3.7). However, we can still see that the most important results remain robust. Both target and acquirer projects are still significantly less likely to be terminated, target projects are more likely to progress, and patenting for deals involving product market incumbents and technologically close projects in overlapping MoAs are more likely to result in more patenting, compared to all other deals. The overall positive effect on patenting is not significant anymore, caused by the negative patenting effect for Phase III projects, rendering also the full effect insignificant.

Robustness: Conditioning on not being terminated

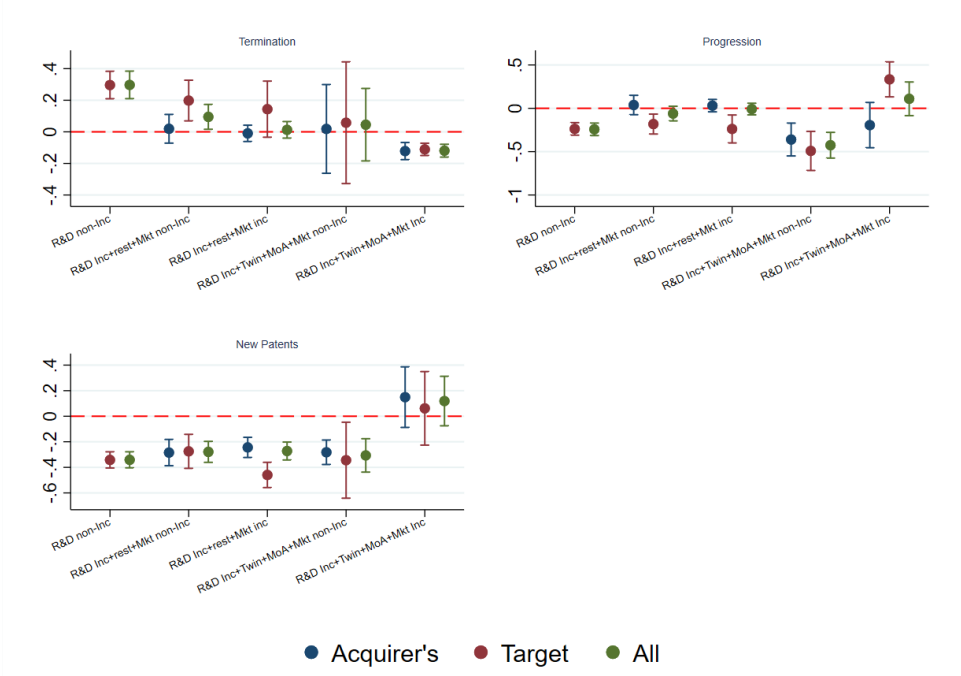
In this section, we perform a robustness check to see how the progression and patenting outcomes change when only considering projects that were not terminated. The main results on progress and new patents still hold. Figure 3.8 shows that conditional on not being terminated, the target projects are more likely to continue compared to the no-deal scenario, and are more likely to receive new patents compared to all other deals, indicating that the R&D activity we observe is driven by pro-active development of projects and their underlying technologies.

Robustness: Alternative measures of technological closeness

In this section, we present results for three alternative measures of technological closeness. Figure 3.9 compares projects' termination rates (top), progression rates (middle) and patenting rates (bottom) for our main empirical specification when taking technological closeness, MoA overlap, and R&D incumbency jointly into consideration. Varying definitions for technological closeness are employed: the default text-based measure (left panels), citation relationships (middle panels), and Jaffe proximity based on patent classification hierarchy (right panels).

The results for R&D non-incumbents and R&D incumbent's projects that are not jointly Twins and overlapping in MoA are replicated well by all measures. However, some differences can be found in the category of technologically close projects with MoA overlap. The citations measure as well as the Jaffe measure contain more noise. This suggests that our detailed, text-based measure has a substantial added value and can uncover important differences that cannot be replicated with more aggregate technology measures.

Figure 3.7: Robustness: Effects of M&As: Technological closeness, MoA overlap, and market incumbency with Phase III projects

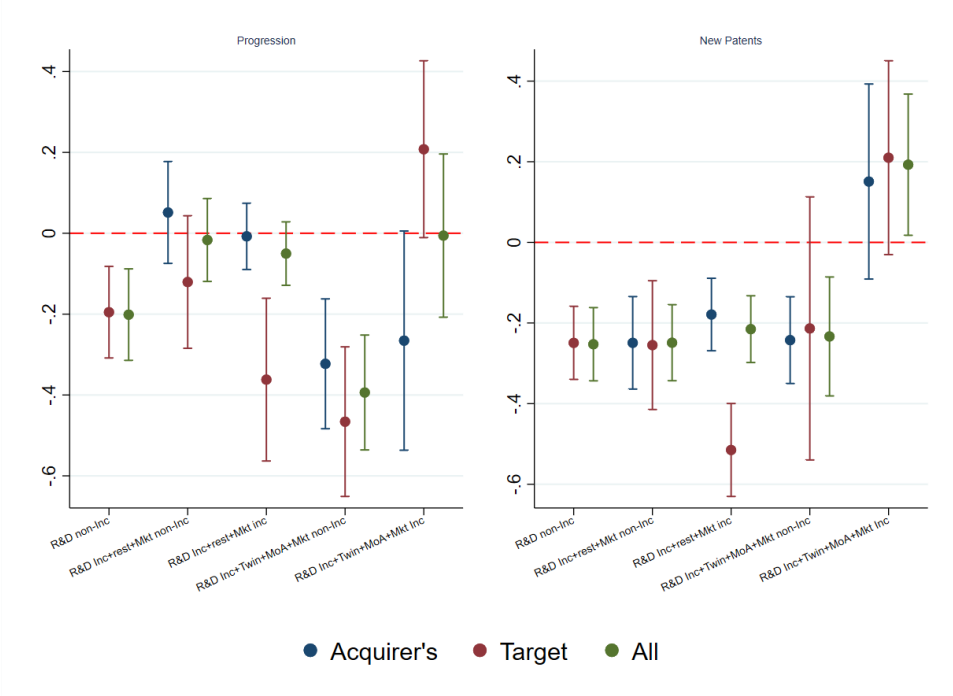


Note: This figure visualizes the estimated effects of M&A depending on the combination of technological closeness, MoA overlap, and market incumbency compared to all never treated projects (no matching). The underlying regression has the following form:

$$\begin{aligned} Pr(DV_i = 1) = \beta_0 &+ \beta_1 Treated : R\&DNonInc_i + \beta_2 MktInc_i + \\ &+ \beta_3 Treated : R\&DIncRest_i + \beta_4 Treated : R\&DIncRest_i \times MktInc_i + \\ &+ \beta_5 Treated : R\&DIncTwinMoA_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i. \end{aligned}$$

The bars give the mean of the total effect with 90% confidence intervals for the respective exclusive treatment groups (β_1 for *R&D non-incumbents*, $\beta_2 + \beta_5 + \beta_6$ for *R&D inc and market inc with technologically close projects with an MoA overlap*, β_5 for *R&D inc and market non-inc with technologically close projects with an MoA overlap*, β_3 for all *other projects of the R&D inc and market non-inc*, and $\beta_2 + \beta_3 + \beta_4$ for all *other projects of the R&D inc and market inc*). The effects on the acquirer's projects are plotted in blue, target projects in red, and all projects in green. Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed and MoA/Tech closeness measures are undefined. The red dashed line highlights a zero effect. Full regression results are presented in table 3.E.14 in Appendix 3.E.

Figure 3.8: Robustness: Effects of M&As: Technological closeness, MoA overlap, and market incumbency for non-terminated projects

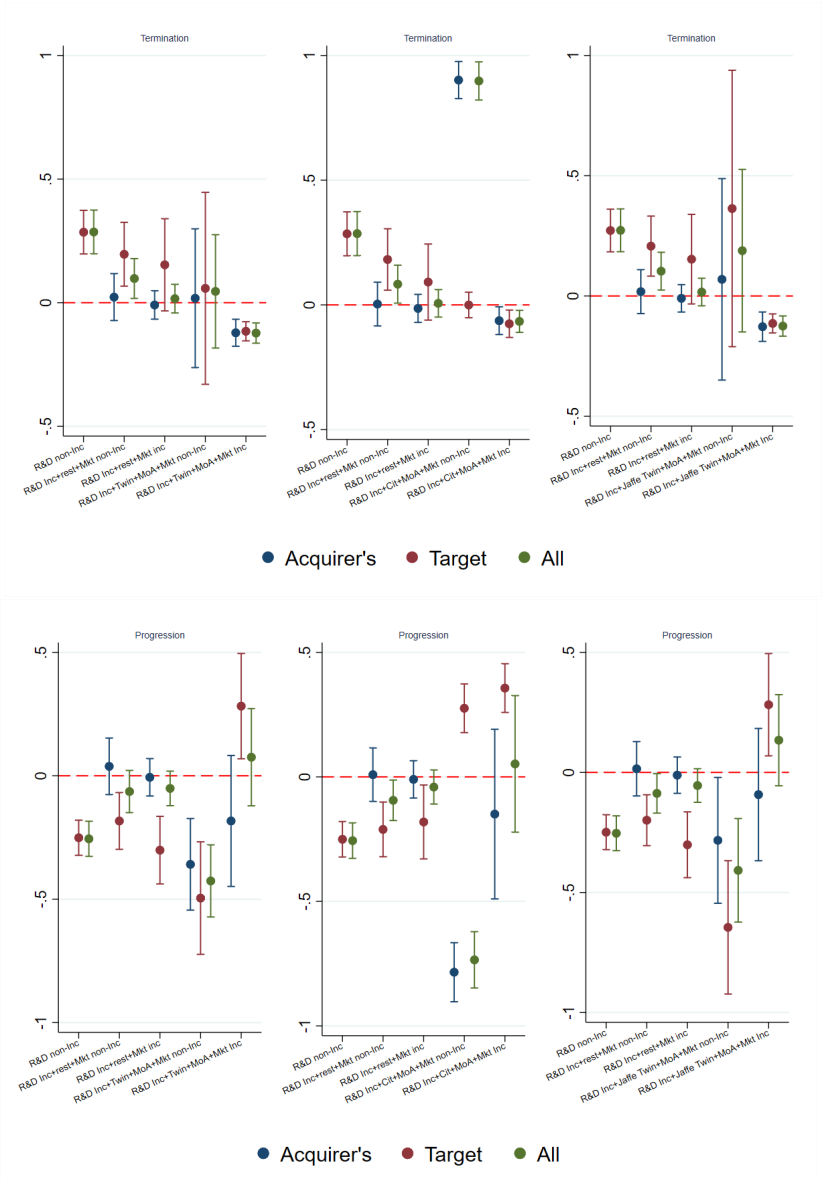


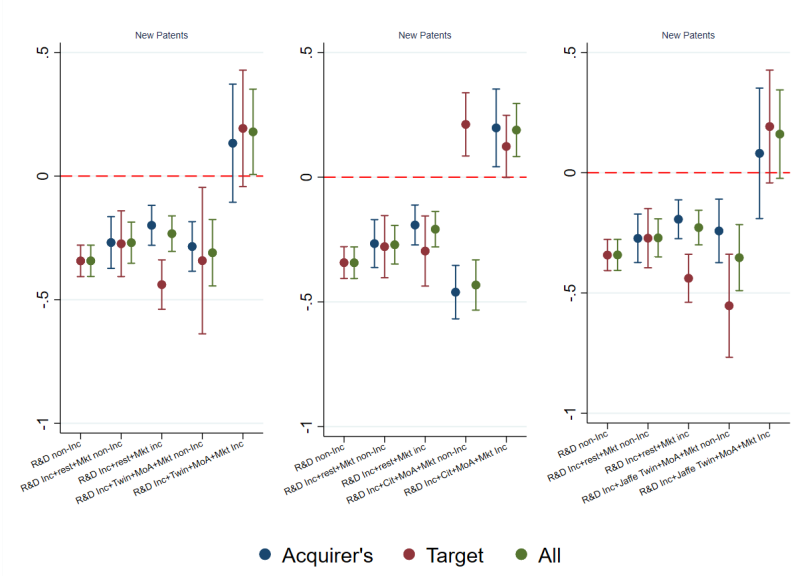
Note: This figure visualizes the estimated effects of M&A depending on the combination of technological closeness, MoA overlap, and market incumbency compared to all never treated projects (no matching). The sample is restricted to non-terminated projects only. The underlying regression has the following form:

$$\begin{aligned} Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 MktInc_i + \\ + \beta_3 Treated : R\&DIncRest_i + \beta_4 Treated : R\&DIncRest_i \times MktInc_i + \\ + \beta_5 Treated : R\&DIncTwinMoA_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i. \end{aligned}$$

The bars give the mean of the total effect with 90% confidence intervals for the respective exclusive treatment groups (β_1 for *R&D non-incumbents*, $\beta_2 + \beta_5 + \beta_6$ for *R&D inc and market inc with technologically close projects with an MoA overlap*, β_5 for *R&D inc and market non-inc with technologically close projects with an MoA overlap*, β_3 for all *other projects of the R&D inc and market non-inc*, and $\beta_2 + \beta_3 + \beta_4$ for all *other projects of the R&D inc and market inc*). The effects on the acquirer's projects are plotted in blue, target projects in red, and all projects in green. Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed and MoA/Tech closeness measures are undefined. The red dashed line highlights a zero effect. Full regression results are presented in table 3.E.15 in Appendix 3.E.

Figure 3.9: Robustness: Effects of M&As: Technological closeness, MoA overlap, and market incumbency with alternative closeness measures





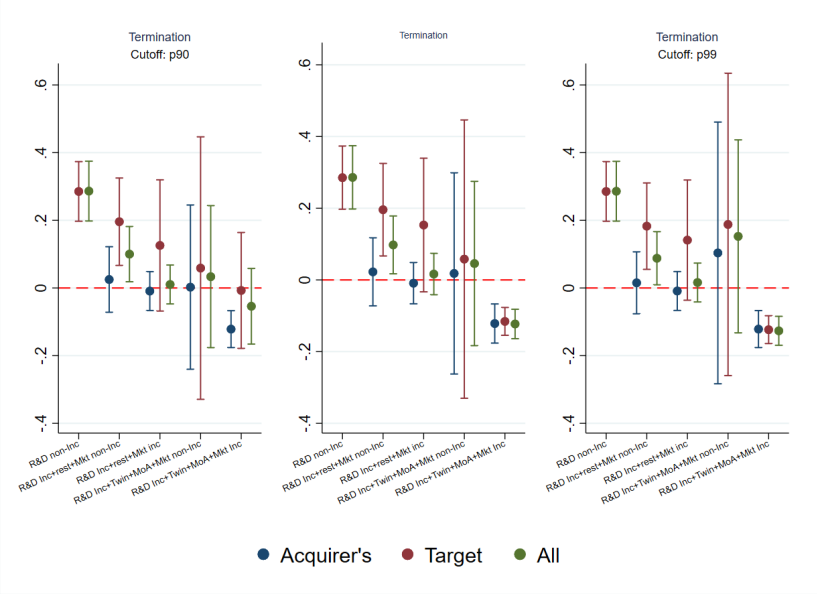
Note: This figure visualizes the estimated effects of M&A on termination rates (top figure), progression rates (middle figure), and patenting (bottom figure), depending on various measures of technological closeness. The first column uses the text-based default definition of technological closeness. The second column uses backward patent citations to proxy technological closeness. The third column uses Jaffe similarity based on co-occurrences of IPC patent groups (e.g. A61P3 - Drugs for disorders of the metabolism) as a measure of technological closeness. These measures are combined with the MoA overlap and market incumbency status of the acquirer. The bars give the mean of the total effect with 90% confidence intervals on the acquirer's projects (plotted in blue), target projects (in red), and all projects (in green). Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed and technological closeness measures are undefined (no connection exists between the target and acquirer's antidiabetics projects). The red dashed line highlights a zero effect. Full regression results are presented in tables 3.E.11,3.E.16, and 3.E.17 in Appendix 3.E.

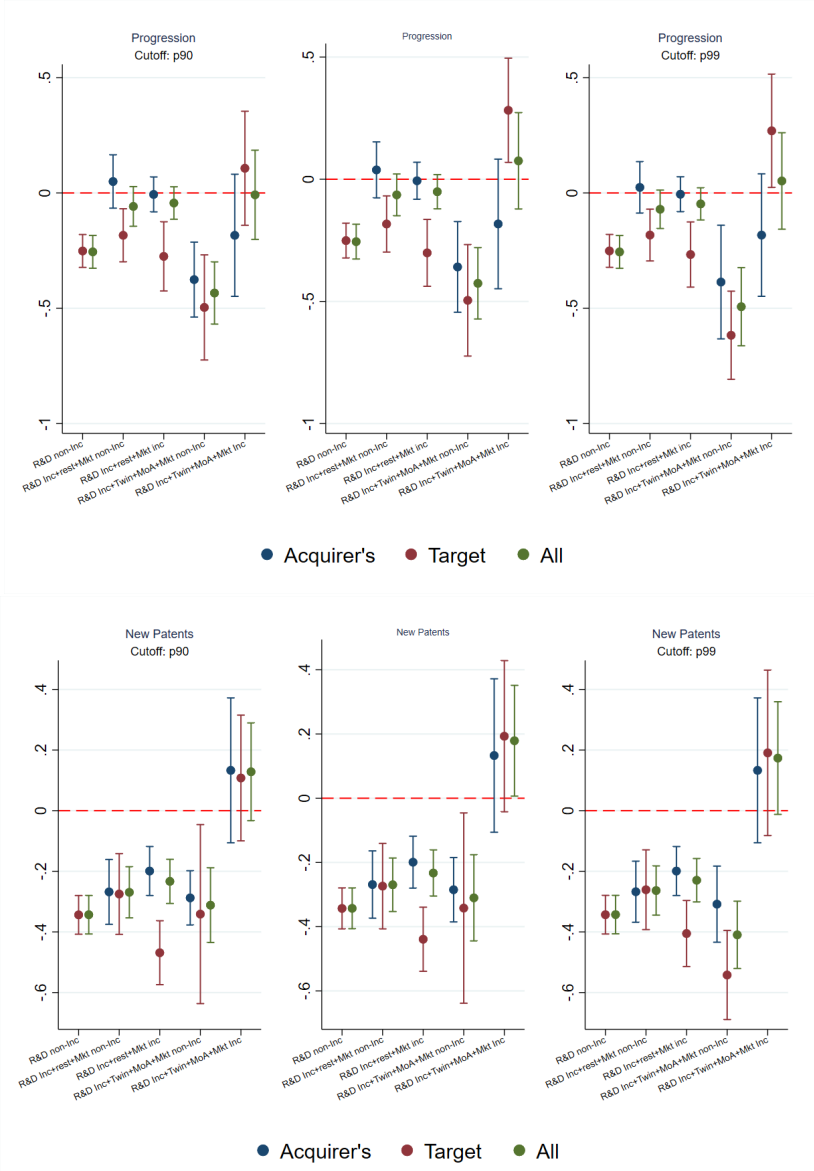
Robustness: Alternative cutoffs for Twins

This section performs a robustness check verifying whether the extent of technological and market closeness affects our main result. Since the MoA definition is fixed, technological closeness is the only parameter we can vary in this robustness analysis. If technological closeness in combination with an MoA overlap is indeed the main driver, the effects should persist with a tighter definition of the technological Twins (i.e. requiring a higher degree of similarity to be classified as a Twin relative to the default specification) and disappear with a looser definition of the technological Twins (ie. requiring a lower degree of similarity to be classified as a Twin relative to the default specification). Figure 3.9 compares the termination rates when the cutoff values are varied.

Focusing on Twin projects with an MoA overlap, the right panel of the figure shows that when tightening the definition for Twins, we can still observe a lower termination rate for the target as well as acquirer's projects. In contrast, when using a lower cutoff value, this effect is rendered insignificant. Interestingly, this already happens when moving from 95th percentile of the similarity distribution to the 90th percentile, indicating that the positive synergy effects are driven by projects in the extreme tail of the similarity distribution. This result also highlights the importance of having a very fine-grained text measure of project similarity. To conclude, the high level of technological similarity is a key driver behind the technology-related synergies.

Figure 3.9: Robustness: Effects of M&As: Technological closeness, MoA overlap, and market incumbency with different cutoffs for Twins





Note: This figure visualizes the estimated effects of M&A on termination, progression, and patenting rates when varying the cutoff value for the definition of Twin projects (based on the baseline text similarity measure). The first column uses a looser definition (90^{th} percentile = 0.11), the second column uses the default value (95^{th} percentile = 0.16), and the last column uses a tighter definition (99^{th} percentile = 0.29). These measures are then combined with the MoA overlap and market incumbency status of the acquirer and projects split into 5 exclusive treatment categories. The bars give the mean of the total effect with 90% confidence intervals on the acquirer's projects (plotted in blue), target projects (in red), and all projects (in green). Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analyzed and technological closeness measures are undefined (no connection exists between the target and acquirer's antidiabetics projects). The red dashed line highlights a zero effect. Full regression results are presented in tables 3.E.18,3.E.11, and 3.E.19 in Appendix 3.E.

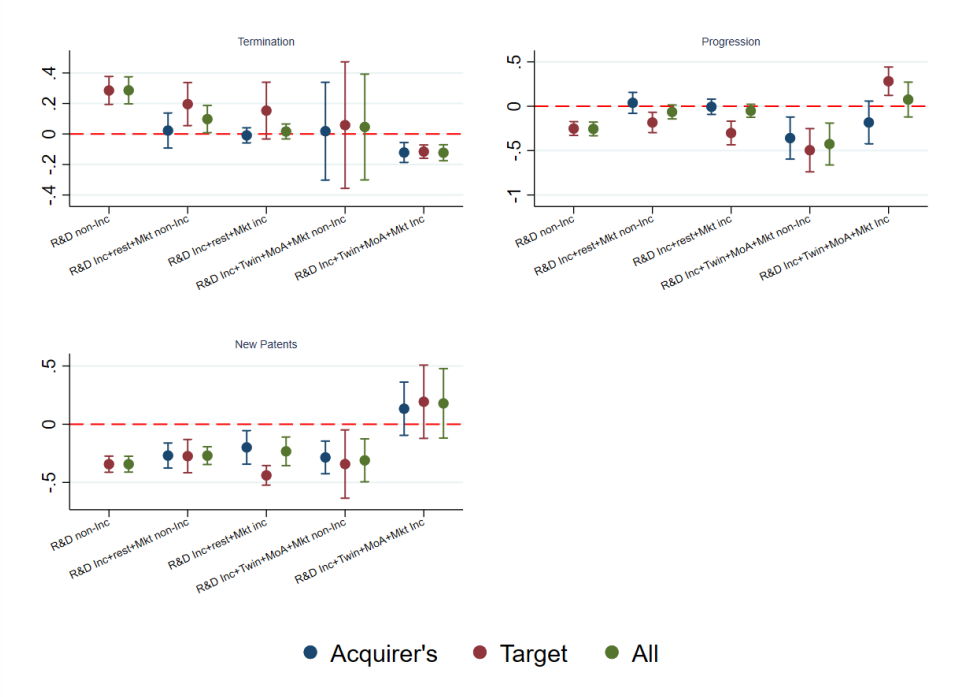
Robustness: Statistical inference and clustering of errors

This section presents two robustness checks addressing the potential concern that the error term might not be independent across individual observations. Whilst this does change the estimated mean coefficients, it impacts the estimated standard errors, might alter their magnitude and, as a result, impact the validity of the statistical inference.

The first plausible source of correlation of the error terms relates to the acquirer and target firms. Projects belonging to the same firm might experience common shocks, since decisions are often made at a firm level, thus having implications for all firm's projects. For example, a cost shock to a particular firm would affect all projects of that firm and make the error terms of the projects correlated with one another. The Figure 3.9 below presents a regression analysis where standard errors are clustered at the firm level. The main result almost does not change for the *progression* and *termination* outcome variables. As regards *patenting*, the confidence intervals are slightly widened. Although this renders the overall *patenting* impact to be neutral rather than positive, this result still implies that only the four-wise combination of incumbency and closeness in product and technology markets is generating better outcomes compared to all other transactions.

The second plausible source of correlation relates to the projects involved in the same transaction. Unlike the previous, this is not necessarily restricted to individual firms but rather extends to projects of the acquirer and target involved in each transaction. The unobserved translation level characteristics subsumed in the error term might involve anything specific to the transaction, for example, the financing of the deal, the rationale for the acquisition, or the integration plans of the acquirer. Such empirical design results in cross-correlation between projects involved in a particular deal (Abadie *et al.*, 2017). The analysis presented in Figure 3.10 employs the sample with matched counterfactuals (684 observations) where both the treated and the control projects can be assigned to a specific transaction. The analysis re-estimates the model with standard errors clustered at the transaction level. Compared to the baseline case reported in the section "Matching with technology similar counterfactuals", the results remain fully robust.

Figure 3.9: Robustness: Effects of M&As: Technological closeness, MoA overlap, and market incumbency (Clustering at firm level)

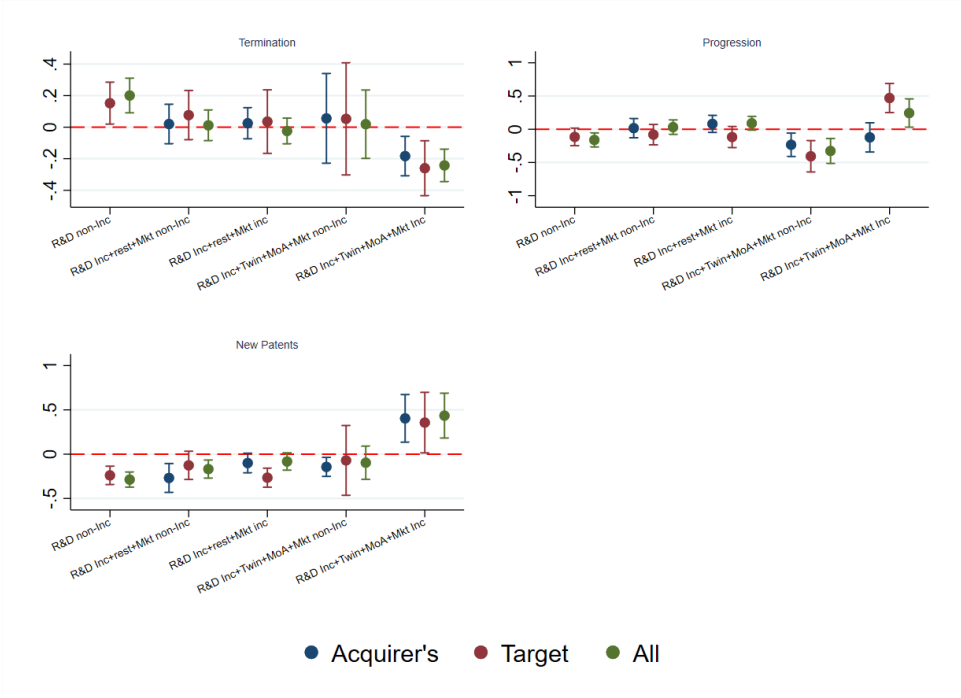


Note: This figure visualizes the estimated effects of M&A depending on the combination of technological closeness, MoA overlap, and market incumbency compared to all never treated projects (no matching). The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 MktInc_i + \\ + \beta_3 Treated : R\&DIncRest_i + \beta_4 Treated : R\&DIncRest_i \times MktInc_i + \\ + \beta_5 Treated : R\&DIncTwinMoA_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

The bars give the mean of the total effect with 90% confidence intervals for the respective exclusive treatment groups (β_1 for *R&D non-incumbents*, $\beta_2 + \beta_5 + \beta_6$ for *R&D inc and market inc with technologically close projects with an MoA overlap*, β_5 for *R&D inc and market non-inc with technologically close projects with an MoA overlap*, β_3 for all *other projects of the R&D inc and market non-inc*, and $\beta_2 + \beta_3 + \beta_4$ for all *other projects of the R&D inc and market inc*). The effects on the acquirer's projects are plotted in blue, target projects in red, and all projects in green. Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed and MoA/Tech closeness measures are undefined. The red dashed line highlights a zero effect. Full regression results are presented in table 3.E.20 in Appendix 3.E.

Figure 3.10: Robustness: Effects of M&As: Technological closeness, MoA overlap, and market incumbency (Clustering at transaction level)



Note: This figure visualizes the estimated effects of M&A depending on the combination of technological closeness, MoA overlap, and market incumbency compared to the pooled group of never treated counterfactuals (not the same firm but the same phase and technologically closest). The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 MktInc_i + \\ + \beta_3 Treated : R\&DIncRest_i + \beta_4 Treated : R\&DIncRest_i \times MktInc_i + \\ + \beta_5 Treated : R\&DIncTwinMoA_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

The bars give the mean of the total effect with 90% confidence intervals for the respective exclusive treatment groups (β_1 for *R&D non-incumbents*, $\beta_2 + \beta_5 + \beta_6$ for *R&D inc and market inc with technologically close projects with an MoA overlap*, β_5 for *R&D inc and market non-inc with technologically close projects with an MoA overlap*, β_3 for all *other projects of the R&D inc and market non-inc*, and $\beta_2 + \beta_3 + \beta_4$ for all *other projects of the R&D inc and market inc*). The effects on the acquirer's projects are plotted in blue, target projects in red, and all projects in green. Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed and MoA/Tech closeness measures are undefined. The red dashed line highlights a zero effect. Full regression results are presented in table 3.E.21 in Appendix 3.E.

3.6 Discussion of results and contributions

Although mergers and acquisitions can have a significant impact on innovation and welfare, the evaluation of innovation effects from M&As in competition enforcement and research alike has so far been unsatisfactory (Veugelers, 2012). This detailed study highlights several key dimensions beyond interactions in the relevant product market space that must be taken into consideration for the assessment of implications that M&As have for innovation.

In line with Cassiman *et al.* (2005), we demonstrate that interactions between merging parties in both product markets and technology markets are essential in the assessment of the M&As' impact on innovation. Our linking of patents to projects and the text of patents allows us to measure technology closeness in a more precise way compared to the existing literature (See e.g. Haucap *et al.*, 2019; Bena & Li, 2014; Ornaghi, 2009b) and uncover product and technology market constellations under which M&As have positive or negative effects on innovation.

We do not find negative innovation effects on acquired projects in advanced stages with high development costs and low uncertainty. Innovation concerns (if any) are occurring primarily in early development phases, when still far from product markets. The most negative effects of these early-stage acquisitions materialize when acquirers are product and technology non-incumbents. Thus, the reasons for negative innovation effects are different than market power-driven "killing" stressed by Cunningham *et al.* (2021). In turn, we find positive effects for early-stage acquisitions by large product market incumbents and projects close in both technology and product markets. This hints at the exploitation of technology synergies and/or the acquirer's experience. In this way, our work connects to the rich stream of literature on the role of synergies in M&As (Maksimovic & Phillips, 2001; Bena & Li, 2014; Grabowski & Kyle, 2008; Hoberg & Phillips, 2010) and the role of technological closeness between assets of acquirers and targets (Ornaghi, 2009b; Bena & Li, 2014; Meder, 2015; Colombo & Rabbiosi, 2014; Desyllas & Hughes, 2010; Yu *et al.*, 2016). From a policy perspective, the merger guidelines both in Europe and the US explicitly acknowledge the need to assess the impact on innovation.[33] Our paper demonstrates that the enforcement framework should be broadened in scope to routinely capture deals along the full range of firms' R&D pipelines, including cases when firms have not yet launched projects.[34] The framework should also consider positions beyond product markets, specifically the acquirer's incumbency in the relevant technology space and the closeness of positions of the merging parties in the technology space. These are essential to separate the beneficial transactions from the harmful ones since not all M&As stifle

[33]Sections 6.4 of the US horizontal merger guidelines and paragraph 38 of the horizontal merger guidelines of the DG COMP.

[34]In this context, the recently announced acquisition of Grail by Illumina scrutinized by both the European Commission and sued by the FTC is an interesting example of the enforcement focus shifting towards targets with no marketed products.

innovation (Denicolò & Polo, 2018).[35]

This paper opens avenues for further work. First, whilst a limitation in resources restricted us to the case of antidiabetics, our methodology is universal and can be scaled to projects covering all therapeutic areas. With such a dataset, the dynamics in the transactions involving non-incumbents would be worth additional exploration. Given that most negative effects on innovation occur here, it would be interesting to see whether these relate to lack of acquirer's experience, lack of resources or simply stem from the fact that antidiabetics was not the area of the acquirer's interest. This would also allow more precise statements regarding the total impact of pharmaceutical M&As on innovation. Second, the richness of the dataset allows us to still explore other technology characteristics of projects involved in the transactions, such as their complexity or disruptiveness which could reveal other critical determining factors for innovation effects. Third, the data can also be further explored to identify in more detail what happens to terminated projects. Do these relate to negative killing scenarios or rather a positive efficiency-enhancing cases of elimination of duplication, or discontinuation of poor-quality projects?

[35] This implication is particularly important in the European context, where the so-called "innovation theory of harm" in the Dow/Du Pont case (M.7932) established a generally undesirable impact of M&As on innovation.

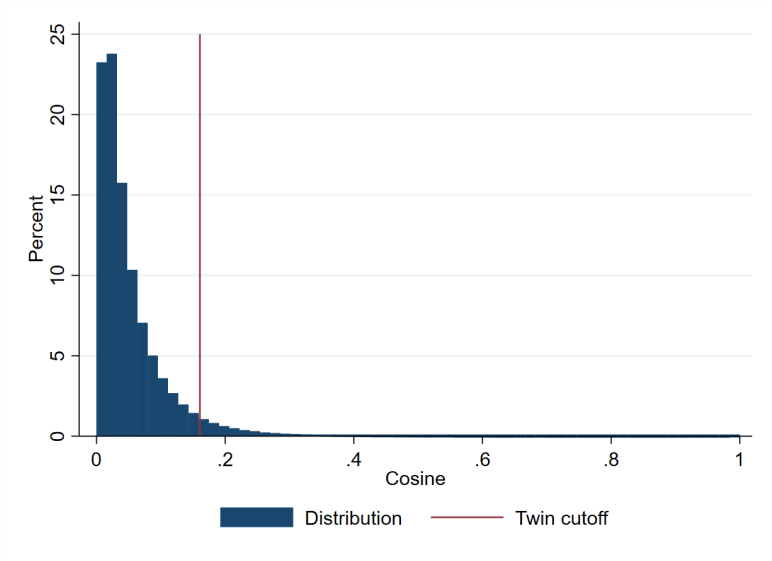
3.A Construction of the technology variables

This section explains how we define the measures of technological closeness between individual projects, and how we split the treated projects between *Twins* and *Siblings* to assess the role of technological closeness in M&As.

We measure the technology closeness by computing cosine similarity between the text of patents we assigned to the acquirer's and target projects (for details on this procedure, please see Appendix 2.C). As a starting point, we match each US patent attached to our projects to the database of Arts *et al.* (2021) who provide patent text in from of standardized and pre-cleaned keywords stemming from patents titles, abstracts, and claims. For each project, we pool keywords together. This then constitutes a vector, characterizing each project by a set of keywords.

We use these vectors to compute cosine similarity between any pair of projects in our sample according to formula $Cosine(A, B) = \frac{A \cdot B}{\|A\| \cdot \|B\|}$, yielding more than 2.4 million observations with the following distribution.

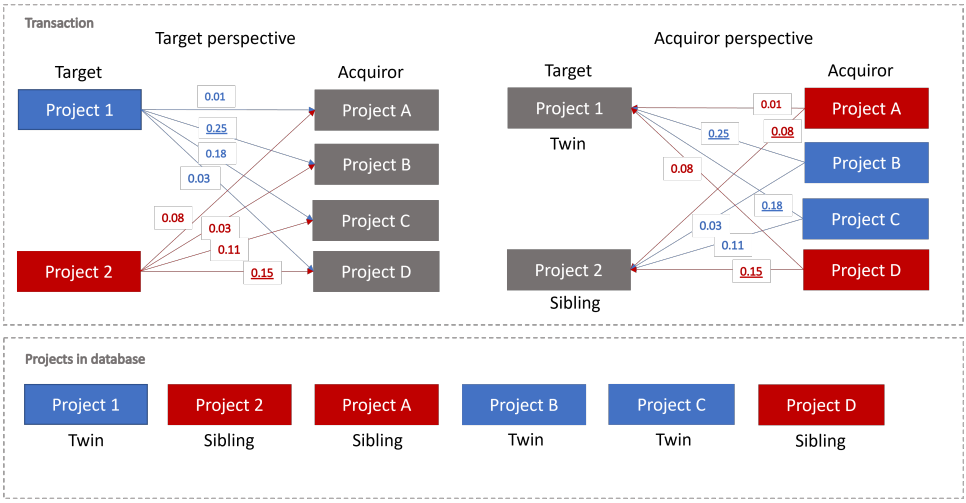
Figure 3.A.11: Distribution of technological similarity



To determine whether a particular project was technologically close (in which case it is called a *Twin*) or technologically distant (in which case it is called a *Sibling*) we use the cut-off value equal to 95th percentile (or 0.16) of the above distribution.

Figure 3.A.12 illustrates how these closeness variables are constructed in practice on an example of a transaction. The transaction involves altogether six treated projects. An acquirer owns a portfolio of four projects - A,B,C,D - and is acquiring a target with two

Figure 3.A.12: Examples: construction of the closeness variables

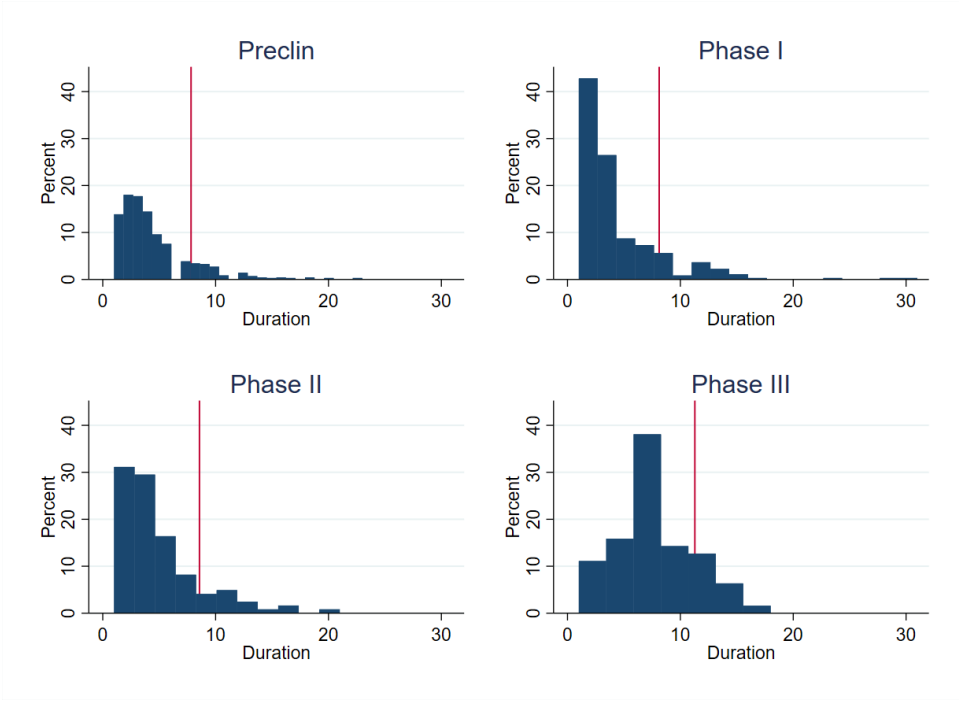


projects - project 1 and 2. The left part of the figure presents the target perspective and shows how target project 1 and project 2 technologically relate to the projects of the acquirer. Project 1 is the most similar to acquirer's project B - with a cosine value of 0.25. This project exceeds our cutoff and would be thus categorized as technologically close - a *Twin*. Project 2 is the most similar to project D - with a cosine value of 0.15. This project does not exceed our cutoff and would be thus categorized as technologically distant - a *Sibling*. Analogically, from the acquirer's perspective, projects A and D are the most similar to project 2 (0.08 and 0.15, respectively) and would be both categorized as *Siblings*, while projects B and C are the most similar to project 1 (0.25 and 0.18, respectively) and would be both categorized as *Twins*. To summarize, the transaction involves 6 treated projects - 1 target twin, 1 target sibling, 2 acquirer twins and 2 acquirer siblings.

3.B Distribution of phase durations

Figure 3.B.13 shows the distribution of projects durations for the different phases of the pharmaceutical R&D process. The blue bars represent the percentage of projects in a particular bin, the red line then plots the critical value of mean plus one standard deviation which is used in our analysis as a cutoff defining *terminations*. Any project exceeding these cutoff values which has not progressed to the next development phase is categorized as terminated.

Figure 3.B.13: Distribution of phase durations and *Termination* cutoff



3.C Placebo analysis

Table 3.C.2: Placebo exercise

	(1) Progress	(2) Termination	(3) New Patents	(4) Progress	(5) Termination	(6) New Patents
Match	-0.031**	0.064***	-0.015	0.003	0.026***	-0.002
	(0.015)	(0.009)	(0.015)	(0.016)	(0.009)	(0.013)
MoA FE	No	No	No	Yes	Yes	Yes
Cohort	No	No	No	Yes	Yes	Yes
Stage	No	No	No	Yes	Yes	Yes
Pat stock FE	No	No	No	Yes	Yes	Yes
Obs	4382	4382	4382	4382	4382	4382
Adj. R2	0.001	0.012	0.000	0.171	0.085	0.373

This table presents results from a placebo exercise to see whether our matching strategy yields valid controls (ie projects in the original and matched samples follow the same development trajectory). We use the sample of never treated observations and for each observation find one which is in the same phase and technologically closest. Then we compare the difference in the outcomes between the two pooled samples. Columns (1)-(3) compare the raw mean differences in outcomes. Columns (4)-(5) compare the mean differences in outcomes when controlling for the full set of fixed effects.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.D Heterogeneity in treated observations

Table 3.D.3: Summary statistics: treatment variable splits

	Acquirer's projects	Target projects	All projects
Phase:			
Before Phase III	167	152	319
Phase III	18	5	23
Acquirer's R&D incumbency:			
R&D Non-Incumbent	0	90	90
R&D Incumbent	185	67	252
Splits conditional on R&D incumbency:			
R&D and Market Incumbency:			
R&D Inc + Mkt Inc	133	27	160
R&D Inc + Mkt non-Inc	52	40	92
Technology closeness:			
Sibling	122	39	161
Twin	63	28	91
MoA overlap:			
No MoA overlap	169	52	221
MoA overlap	16	15	31
Tech + MoA:			
Sibling + no MoA overlap	118	36	154
Sibling + MoA overlap	4	3	7
Twin + no MoA overlap	51	16	67
Twin + MoA overlap	12	12	24
Observations	185	157	342

Note: R&D incumbency of the acqurirer determines the existence of its antidiabetics portfolio before acquiring a project. As such, the analysis of the impacts of M&As on acquirer's portfolio makes only sense for the R&D incumbents (the portfolio of R&D non-incumbents is empty). In addition, this also implies that the technological closeness variables as well as the market overlap variables are only defined for the R&D incumbents (no relationship between projects exists for R&D non-incumbents with empty antidiabetics portfolio). In our dataset, all firms with launched products worked on R&D at the same time. Thus, the category of R&D non-incumbent + market incumbent does not exist and the split between market incumbents and non-incumbents are only defined conditional on R&D incumbency.

3.E Regression results

Table 3.E.4: Results: Phase III

Panel A: All

	All projects		
	(1)	(2)	(3)
	Termination	Progress	New Patents
Treated	0.129***	-0.143***	-0.234***
	(0.026)	(0.028)	(0.027)
Phase III	-0.096***	-0.288***	0.110**
	(0.018)	(0.046)	(0.049)
Treated × Phase III	-0.037	0.434***	-0.336***
	(0.064)	(0.103)	(0.122)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	2028	2028	2028
Adj. R2	0.058	0.109	0.221

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1)	(2)	(3)	(4)	(5)	(6)
	Termination	Progress	New Patents	Progress	Termination	New Patents
Treated	0.029	-0.052	-0.181***	0.239***	-0.237***	-0.294***
	(0.029)	(0.037)	(0.037)	(0.040)	(0.035)	(0.034)
Phase III	-0.088***	-0.289***	0.110**	-0.100***	-0.275***	0.097**
	(0.018)	(0.047)	(0.049)	(0.018)	(0.046)	(0.049)
Treated × Phase III	-0.012	0.319***	-0.410***	0.128	0.625***	-0.205
	(0.035)	(0.110)	(0.145)	(0.225)	(0.229)	(0.193)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1871	1871	1871	1843	1843	1843
Adj. R2	0.031	0.105	0.224	0.076	0.105	0.245

Notes: This table shows the results of the OLS regressions measuring the effect of M&A in Phase III/other phases on all projects (panel A) and acquirer's and target projects (panel B). The treated projects are compared to the set of all never treated projects. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated_i + \beta_2 PhaseIII_i + \beta_4 Treated_i \times PhaseIII_i + \epsilon_i$$

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.5: Results: Baseline

Panel A: All

	All projects		
	(1)	(2)	(3)
	Termination	Progress	New Patents
Treated	0.124***	-0.134***	-0.237***
	(0.026)	(0.028)	(0.027)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	2005	2005	2005
Adj. R2	0.078	0.125	0.240

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1)	(2)	(3)	(4)	(5)	(6)
	Termination	Progress	New Patents	Termination	Progress	New Patents
Treated	0.017	-0.034	-0.187***	0.240***	-0.239***	-0.294***
	(0.029)	(0.038)	(0.038)	(0.040)	(0.035)	(0.034)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1853	1853	1853	1838	1838	1838
Adj. R2	0.053	0.122	0.245	0.098	0.128	0.266

Notes: This table shows the results of the OLS regressions measuring the effect of M&A on all projects (panel A) and acquirer's and target projects (panel B). The treated projects are compared to the set of all never treated projects. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated_i + \epsilon_i$$

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.6: Results: Baseline (logit)

Panel A: All

	All projects		
	(1)	(2)	(3)
	Termination	Progress	New Patents
Treated	0.836***	-0.659***	-1.341***
	(0.160)	(0.141)	(0.181)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	1935	2005	1935
Pseudo R2	0.116	0.111	0.204

Marginal effects

	All projects		
	(1)	(2)	(3)
	Termination	Progress	New Patents
Treated	0.092***	-0.137***	-0.237***
	(0.017)	(0.029)	(0.030)

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1)	(2)	(3)	(4)	(5)	(6)
	Termination	Progress	New Patents	Termination	Progress	New Patents
Treated	0.101	-0.170	-1.055***	1.374***	-1.244***	-1.722***
	(0.247)	(0.179)	(0.245)	(0.196)	(0.221)	(0.260)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1784	1853	1784	1768	1838	1768
Pseudo R2	0.098	0.109	0.205	0.136	0.114	0.227

Marginal effects

	Acquirer's projects			Target projects		
	(1)	(2)	(3)	(4)	(5)	(6)
	Termination	Progress	New Patents	Termination	Progress	New Patents
Treated	0.010	-0.036	-0.189***	0.146***	-0.258***	-0.297***
	(0.024)	(0.038)	(0.042)	(0.020)	(0.045)	(0.042)

Notes: This table shows the results of the logit regressions measuring the effect of M&A on all projects (panel A) and acquirer's and target projects (panel B). The treated projects are compared to the set of all never treated projects. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated_i + \epsilon_i$$

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.7: Results: R&D and market incumbency

Panel A: All

	All projects		
	(1)	(2)	(3)
	Termination	Progress	New Patents
Treated: R&D non-Inc	0.286***	-0.256***	-0.344***
	(0.054)	(0.043)	(0.039)
Treated: R&D Inc	0.092**	-0.101**	-0.273***
	(0.047)	(0.049)	(0.047)
Market Inc	-0.073***	0.077***	-0.082***
	(0.015)	(0.028)	(0.025)
Treated: R&D Inc × Market Inc	-0.015	-0.013	0.160**
	(0.056)	(0.066)	(0.065)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	2005	2005	2005
Adj. R2	0.100	0.132	0.246

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1)	(2)	(3)	(4)	(5)	(6)
	Termination	Progress	New Patents	Termination	Progress	New Patents
Treated: R&D non-Inc				0.285***	-0.251***	-0.343***
				(0.053)	(0.043)	(0.039)
Treated: R&D Inc	0.022	-0.007	-0.270***	0.182**	-0.212***	-0.279***
	(0.055)	(0.066)	(0.057)	(0.075)	(0.067)	(0.076)
Market Inc	-0.076***	0.079***	-0.081***	-0.075***	0.080***	-0.089***
	(0.015)	(0.029)	(0.025)	(0.015)	(0.029)	(0.025)
Treated: R&D Inc × Market Inc	0.039	-0.086	0.169**	-0.029	-0.006	0.106
	(0.064)	(0.081)	(0.076)	(0.113)	(0.112)	(0.112)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1853	1853	1853	1838	1838	1838
Adj. R2	0.061	0.125	0.249	0.107	0.131	0.271

Notes: This table shows the results of the OLS regressions measuring the effect of M&A depending on both R&D and market incumbency status on all projects (panel A) and acquirer's and target projects (panel B). The treated projects are split in three exclusive groups and compared to the set of all never treated projects. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 Treated : R\&DInc_i + \\ + \beta_3 MktInc_i + \beta_4 Treated : R\&DInc_i \times MktInc_i + \epsilon_i.$$

Please note that the group of R&D non-incumbents is not defined for the acquirers since their antidiabetics portfolio is empty before a transaction. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.8: Results: Technological closeness

Panel A: All

	All projects		
	(1)	(2)	(3)
	Termination	Progress	New Patents
R&D non-Inc	0.301***	-0.272***	-0.324***
	(0.053)	(0.043)	(0.038)
R&D Inc + Sib	0.076**	-0.064*	-0.238***
	(0.034)	(0.038)	(0.039)
R&D Inc + Twin	0.019	-0.117**	-0.139***
	(0.040)	(0.053)	(0.053)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	2005	2005	2005
Adj. R2	0.093	0.129	0.243

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1)	(2)	(3)	(4)	(5)	(6)
	Termination	Progress	New Patents	Termination	Progress	New Patents
R&D non-Inc				0.301***	-0.268***	-0.323***
				(0.053)	(0.043)	(0.038)
R&D Inc + Sib	0.022	-0.001	-0.213***	0.243***	-0.246***	-0.313***
	(0.035)	(0.044)	(0.046)	(0.080)	(0.066)	(0.068)
R&D Inc + Twin	0.005	-0.105	-0.129**	0.041	-0.137	-0.169*
	(0.047)	(0.065)	(0.063)	(0.070)	(0.087)	(0.093)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1853	1853	1853	1838	1838	1838
Adj. R2	0.053	0.123	0.245	0.103	0.128	0.267

Notes: This table shows the results of the OLS regressions measuring the effect of M&A depending on technological closeness on all projects (panel A) and acquirer's and target projects (panel B). The treated projects are split in three exclusive groups and compared to the set of all never treated projects. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \\ + \beta_2 Treated : R\&DIncSib_i + \beta_3 Treated : R\&DIncTwin_i + \epsilon_i.$$

Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed and technological closeness is undefined (no connection exists between the target and acquirer's antidiabetics projects). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.9: Results: MoA overlap

Panel A: All

	All projects		
	(1)	(2)	(3)
	Termination	Progress	New Patents
R&D non-Inc	0.301***	-0.272***	-0.325***
	(0.053)	(0.043)	(0.038)
R&D Inc + no MoA	0.066**	-0.072**	-0.229***
	(0.029)	(0.034)	(0.034)
R&D Inc + MoA	-0.010	-0.159*	-0.022
	(0.061)	(0.089)	(0.078)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	2005	2005	2005
Adj. R2	0.093	0.129	0.244

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1)	(2)	(3)	(4)	(5)	(6)
	Termination	Progress	New Patents	Termination	Progress	New Patents
R&D non-Inc				0.302***	-0.269***	-0.325***
				(0.053)	(0.043)	(0.038)
R&D Inc + no MoA	0.024	-0.015	-0.202***	0.193***	-0.235***	-0.315***
	(0.030)	(0.039)	(0.040)	(0.067)	(0.057)	(0.063)
R&D Inc + MoA	-0.055	-0.233**	-0.024	0.032	-0.078	-0.028
	(0.075)	(0.112)	(0.099)	(0.095)	(0.130)	(0.115)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1853	1853	1853	1838	1838	1838
Adj. R2	0.053	0.123	0.246	0.101	0.128	0.268

Notes: This table shows the results of the OLS regressions measuring the effect of M&A depending on closeness in future product markets (MoA overlap) on all projects (panel A) and on acquirer's and target projects (panel B). The treated projects are split in three exclusive groups and compared to the set of all never treated projects. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \\ + \beta_2 Treated : R\&DIncNoMoA_i + \beta_3 Treated : R\&DIncMoA_i + \epsilon_i.$$

Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed and MoA overlap is undefined (no connection exists between the target and acquirer's antidiabetics projects). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.10: Results: Technological closeness and MoA overlap

Panel A: All

	All projects		
	(1)	(2)	(3)
	Termination	Progress	New Patents
R&D non-Inc	0.301***	-0.272***	-0.325***
	(0.053)	(0.043)	(0.038)
R&D Inc rest	0.065**	-0.074**	-0.226***
	(0.029)	(0.033)	(0.034)
R&D Inc + Twin + MoA	-0.024	-0.163*	0.003
	(0.063)	(0.095)	(0.087)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	2005	2005	2005
Adj. R2	0.093	0.129	0.244

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1)	(2)	(3)	(4)	(5)	(6)
	Termination	Progress	New Patents	Termination	Progress	New Patents
R&D non-Inc				0.301***	-0.268***	-0.324***
				(0.053)	(0.043)	(0.038)
R&D Inc rest	0.020	-0.014	-0.197***	0.196***	-0.237***	-0.312***
	(0.030)	(0.039)	(0.040)	(0.065)	(0.055)	(0.059)
R&D Inc + Twin + MoA	-0.024	-0.298***	-0.043	-0.026	-0.022	0.037
	(0.088)	(0.101)	(0.102)	(0.089)	(0.148)	(0.137)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1853	1853	1853	1838	1838	1838
Adj. R2	0.053	0.124	0.245	0.102	0.129	0.268

Notes: This table shows the results of the OLS regressions measuring the effect of M&A depending on the combination of technological closeness and MoA overlap on all projects (panel A) and on acquirer's and target projects (panel B). The treated projects are split in three exclusive groups and compared to the set of all never treated projects. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \\ + \beta_2 Treated : R\&DIncRest_i + \beta_3 Treated : R\&DIncTwinMoA_i + \epsilon_i.$$

Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed and MoA overlap is undefined (no connection exists between the target and acquirer's antidiabetics projects). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.11: Results: Tech closeness, MoA overlap, R&D and market incumbency

Panel A: All

	All projects		
	(1)	(2)	(3)
	Termination	Progress	New Patents
R&D non-Inc	0.286***	-0.255***	-0.343***
	(0.054)	(0.043)	(0.039)
R&D Inc + rest	0.098**	-0.064	-0.269***
	(0.049)	(0.052)	(0.051)
R&D Inc + Twin + MoA	0.046	-0.426***	-0.310***
	(0.139)	(0.089)	(0.082)
Mkt Inc	-0.074***	0.076***	-0.080***
	(0.015)	(0.029)	(0.025)
R&D Inc + rest × Mkt Inc	-0.008	-0.064	0.117*
	(0.060)	(0.069)	(0.069)
R&D Inc + Twin + MoA × Mkt Inc	-0.095	0.425***	0.569***
	(0.141)	(0.147)	(0.131)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	2005	2005	2005
Adj. R2	0.100	0.134	0.250

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1)	(2)	(3)	(4)	(5)	(6)
	Termination	Progress	New Patents	Termination	Progress	New Patents
R&D non-Inc				0.285***	-0.251***	-0.343***
				(0.054)	(0.043)	(0.039)
R&D Inc + rest	0.023	0.038	-0.269***	0.196**	-0.183***	-0.274***
	(0.058)	(0.070)	(0.064)	(0.078)	(0.070)	(0.081)
R&D Inc + Twin + MoA	0.018	-0.359***	-0.285***	0.058	-0.496***	-0.342*
	(0.170)	(0.113)	(0.061)	(0.236)	(0.139)	(0.180)
Mkt Inc	-0.076***	0.078***	-0.079***	-0.076***	0.081***	-0.087***
	(0.015)	(0.029)	(0.025)	(0.015)	(0.029)	(0.025)
R&D Inc + rest × Mkt Inc	0.045	-0.123	0.149*	0.033	-0.199*	-0.078
	(0.067)	(0.085)	(0.082)	(0.137)	(0.110)	(0.101)
R&D Inc + Twin + MoA × Mkt Inc	-0.064	0.098	0.497***	-0.098	0.696***	0.622***
	(0.173)	(0.196)	(0.156)	(0.237)	(0.187)	(0.228)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1853	1853	1853	1838	1838	1838
Adj. R2	0.060	0.126	0.250	0.108	0.135	0.275

Notes: This table shows the results of the OLS regressions measuring the effect of M&A depending on the combination of technological closeness, MoA overlap, R&D and product market incumbency on all projects (panel A) and acquirer's and target projects (panel B). The treated projects are split in five exclusive groups and compared to the set of all never treated projects. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 Treated : R\&DIncRest_i + \\ + \beta_3 Treated : R\&DIncTwinMoA_i + \beta_4 MktInc_i + \\ + \beta_5 Treated : R\&DIncRest_i \times MktInc_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed (no connection exists between the target and acquirer's antidiabetics projects). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.12: Robustness: Tech closeness, MoA overlap, R&D and market incumbency with pooled matching

Panel A: All

	All projects		
	(1)	(2)	(3)
	Termination	Progress	New Patents
R&D non-Inc	0.201***	-0.161***	-0.287***
	(0.065)	(0.059)	(0.054)
R&D Inc + rest	0.012	0.032	-0.168***
	(0.059)	(0.065)	(0.062)
R&D Inc + Twin + MoA	0.019	-0.327***	-0.096
	(0.122)	(0.099)	(0.109)
Mkt Inc	-0.122***	0.215***	-0.069
	(0.035)	(0.055)	(0.048)
R&D Inc + rest × Mkt Inc	0.086	-0.156*	0.154**
	(0.068)	(0.084)	(0.078)
R&D Inc + Twin + MoA × Mkt Inc	-0.139	0.356**	0.600***
	(0.134)	(0.155)	(0.166)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	638	638	638
Adj. R2	0.167	0.279	0.197

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1)	(2)	(3)	(4)	(5)	(6)
	Termination	Progress	New Patents	Termination	Progress	New Patents
R&D non-Inc				0.152**	-0.115*	-0.239***
				(0.075)	(0.070)	(0.062)
R&D Inc + rest	0.020	0.017	-0.268***	0.077	-0.081	-0.126
	(0.079)	(0.093)	(0.093)	(0.091)	(0.092)	(0.096)
R&D Inc + Twin + MoA	0.056	-0.234*	-0.144	0.053	-0.407***	-0.070
	(0.155)	(0.123)	(0.097)	(0.205)	(0.140)	(0.245)
Mkt Inc	-0.061	0.165**	-0.137**	-0.189***	0.259***	0.082
	(0.046)	(0.081)	(0.066)	(0.061)	(0.089)	(0.082)
R&D Inc + rest × Mkt Inc	0.066	-0.101	0.306***	0.148	-0.295**	-0.222
	(0.078)	(0.116)	(0.103)	(0.150)	(0.149)	(0.139)
R&D Inc + Twin + MoA × Mkt Inc	-0.178	-0.054	0.685***	-0.124	0.618***	0.344
	(0.177)	(0.193)	(0.197)	(0.222)	(0.206)	(0.287)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	334	334	334	304	304	304
Adj. R2	0.128	0.344	0.200	0.207	0.254	0.257

Notes: This table shows the results of the OLS regressions measuring the effect of M&A depending on the combination of technological closeness, MoA overlap, R&D and product market incumbency on all projects (panel A) and acquirer's and target projects (panel B). The treated projects are split in five exclusive groups and compared to the pooled counterfactual group, consisting of the technologically most similar, never treated projects of other firms in the same phase, assigned to each treated project. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 Treated : R\&DIncRest_i + \\ + \beta_3 Treated : R\&DIncTwinMoA_i + \beta_4 MktInc_i + \\ + \beta_5 Treated : R\&DIncRest_i \times MktInc_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed (no connection exists between the target and acquirer's antidiabetics projects). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.13: Robustness: Tech closeness, MoA overlap, R&D and market incumbency with exact matching

Panel A: All

	All projects		
	(1)	(2)	(3)
	Progress	Termination	New Patents
R&D non-Inc	0.286***	-0.163**	-0.272***
	(0.071)	(0.077)	(0.063)
R&D Inc + rest	0.025	-0.034	-0.157*
	(0.079)	(0.078)	(0.081)
R&D Inc + Twin + MoA	0.025	-0.196	-0.048
	(0.151)	(0.205)	(0.161)
Mkt Inc	-0.052	0.125*	-0.060
	(0.064)	(0.075)	(0.063)
R&D Inc + rest × Mkt Inc	0.038	-0.135	0.080
	(0.099)	(0.112)	(0.105)
R&D Inc + Twin + MoA × Mkt Inc	-0.365*	0.296	0.365*
	(0.206)	(0.263)	(0.207)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Match pair FE	Yes	Yes	Yes
Obs	638	638	638
Adj. R2	0.221	0.351	0.291

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1)	(2)	(3)	(4)	(5)	(6)
	Progress	Termination	New Patents	Progress	Termination	New Patents
R&D non-Inc				0.239**	-0.167	-0.286***
				(0.096)	(0.105)	(0.069)
R&D Inc + rest	0.093	0.001	-0.234*	0.004	-0.150	-0.142
	(0.088)	(0.109)	(0.135)	(0.134)	(0.132)	(0.121)
R&D Inc + Twin + MoA	0.007	0.079	-0.245	0.143	-0.476	-0.185
	(0.129)	(0.235)	(0.164)	(0.317)	(0.330)	(0.355)
Mkt Inc	0.035	0.114	-0.069	-0.120	0.155	-0.073
	(0.079)	(0.109)	(0.093)	(0.133)	(0.156)	(0.118)
R&D Inc + rest × Mkt Inc	-0.084	-0.098	0.148	0.084	-0.224	-0.103
	(0.110)	(0.145)	(0.148)	(0.222)	(0.232)	(0.217)
R&D Inc + Twin + MoA × Mkt Inc	0.015	-0.458	0.861**	-0.576	0.818**	0.255
	(0.189)	(0.353)	(0.345)	(0.371)	(0.396)	(0.363)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Match pair FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	334	334	334	304	304	304
Adj. R2	0.231	0.435	0.261	0.172	0.192	0.360

Notes: This table shows the results of the OLS regressions measuring the effect of M&A depending on the combination of technological closeness and MoA overlap, R&D and product market incumbency on all projects (panel A) and acquirer's and target projects (panel B). The treated projects are split in five exclusive groups and compared to their counterfactual counterparts (ie matched never treated technologically closest project of other firms in the same phase) The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 Treated : R\&DIncRest_i + \\ + \beta_3 Treated : R\&DIncTwinMoA_i + \beta_4 MktInc_i + \\ + \beta_5 Treated : R\&DIncRest_i \times MktInc_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed (no connection exists between the target and acquirer's antidiabetics projects). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.14: Robustness: Tech closeness, MoA overlap, R&D and market incumbency with Phase III included

Panel A: All

	All projects		
	(1)	(2)	(3)
	Termination	Progress	New Patents
R&D non-Inc	0.297***	-0.243***	-0.341***
	(0.053)	(0.044)	(0.038)
R&D Inc + rest	0.095**	-0.061	-0.279***
	(0.048)	(0.051)	(0.050)
R&D Inc + Twin + MoA	0.045	-0.426***	-0.306***
	(0.139)	(0.090)	(0.079)
Mkt Inc	-0.073***	0.072**	-0.074***
	(0.015)	(0.028)	(0.025)
R&D Inc + rest × Mkt Inc	-0.009	-0.019	0.081
	(0.057)	(0.068)	(0.067)
R&D Inc + Twin + MoA × Mkt Inc	-0.092	0.464***	0.499***
	(0.141)	(0.146)	(0.140)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	2028	2028	2028
Adj. R2	0.103	0.133	0.238

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1)	(2)	(3)	(4)	(5)	(6)
	Termination	Progress	New Patents	Termination	Progress	New Patents
R&D non-Inc				0.296***	-0.238***	-0.342***
				(0.053)	(0.044)	(0.038)
R&D Inc + rest	0.019	0.038	-0.284***	0.198**	-0.182***	-0.275***
	(0.055)	(0.068)	(0.062)	(0.078)	(0.070)	(0.081)
R&D Inc + Twin + MoA	0.018	-0.360***	-0.282***	0.058	-0.492***	-0.344*
	(0.171)	(0.115)	(0.058)	(0.234)	(0.137)	(0.181)
Mkt Inc	-0.076***	0.074**	-0.073***	-0.076***	0.081***	-0.087***
	(0.015)	(0.029)	(0.025)	(0.015)	(0.029)	(0.025)
R&D Inc + rest × Mkt Inc	0.047	-0.080	0.113	0.022	-0.138	-0.097
	(0.063)	(0.082)	(0.080)	(0.132)	(0.121)	(0.101)
R&D Inc + Twin + MoA × Mkt Inc	-0.064	0.093	0.504***	-0.093	0.744***	0.493**
	(0.173)	(0.195)	(0.154)	(0.235)	(0.182)	(0.249)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1871	1871	1871	1843	1843	1843
Adj. R2	0.061	0.128	0.238	0.110	0.132	0.272

Notes: This table shows the results of the OLS regressions measuring the effect of M&A depending on the combination of technological closeness, MoA overlap, R&D and product market incumbency on all projects (panel A) and acquirer's and target projects (panel B). The treated projects are split in five exclusive groups and compared to the set of all never treated projects. All Phase III projects are included in the sample. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 Treated : R\&DIncRest_i + \\ + \beta_3 Treated : R\&DIncTwinMoA_i + \beta_4 MktInc_i + \\ + \beta_5 Treated : R\&DIncRest_i \times MktInc_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed (no connection exists between the target and acquirer's antidiabetics projects). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.15: Robustness: Tech closeness, MoA overlap, R&D and market incumbency for non-terminated projects only

Panel A: All

	All projects	
	(1)	(2)
	Progress	New Patents
R&D non-Inc	-0.201***	-0.253***
	(0.069)	(0.055)
R&D Inc + rest	-0.017	-0.249***
	(0.062)	(0.057)
R&D Inc + Twin + MoA	-0.394***	-0.233***
	(0.086)	(0.090)
Mkt Inc	0.036	-0.078***
	(0.030)	(0.025)
R&D Inc + rest × Mkt Inc	-0.069	0.112
	(0.080)	(0.078)
R&D Inc + Twin + MoA × Mkt Inc	0.352**	0.504***
	(0.147)	(0.137)
MoA FE	Yes	Yes
Cohort	Yes	Yes
Stage	Yes	Yes
Pat stock FE	Yes	Yes
Obs	1730	1730
Adj. R2	0.131	0.255

Panel B: Acquirers and targets

	Acquirer's projects		Target projects	
	(1)	(2)	(3)	(4)
	Progress	New Patents	Progress	New Patents
R&D non-Inc			-0.195***	-0.249***
			(0.069)	(0.055)
R&D Inc + rest	0.051	-0.249***	-0.121	-0.255***
	(0.077)	(0.070)	(0.100)	(0.097)
R&D Inc + Twin + MoA	-0.323***	-0.243***	-0.466***	-0.214
	(0.098)	(0.065)	(0.112)	(0.198)
Mkt Inc	0.037	-0.078***	0.042	-0.090***
	(0.030)	(0.025)	(0.030)	(0.025)
R&D Inc + rest × Mkt Inc	-0.097	0.148	-0.284*	-0.171
	(0.093)	(0.090)	(0.157)	(0.119)
R&D Inc + Twin + MoA × Mkt Inc	0.020	0.471***	0.631***	0.513**
	(0.190)	(0.160)	(0.171)	(0.244)
MoA FE	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes
Obs	1638	1638	1589	1589
Adj. R2	0.132	0.255	0.126	0.283

Notes: This table shows the results of the OLS regressions measuring the effect of M&A depending on the combination of technological closeness, MoA overlap, R&D and product market incumbency on all projects (panel A) and acquirer's and target projects (panel B). The treated projects are split in five exclusive groups and compared to the set of all never treated projects. Terminated projects (both treated and controls) are excluded from the analysis. This restrict the outcome variables to *Progression* and *New Patents* only. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 Treated : R\&DIncRest_i + \\ + \beta_3 Treated : R\&DIncTwinMoA_i + \beta_4 MktInc_i + \\ + \beta_5 Treated : R\&DIncRest_i \times MktInc_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed (no connection exists between the target and acquirer's antidiabetics projects). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.16: Robustness: Tech closeness, MoA overlap, R&D and market incumbency with citation closeness measure

Panel A: All

	All projects		
	(1) Termination	(2) Progress	(3) New Patents
R&D non-Inc	0.286***	-0.256***	-0.343***
	(0.054)	(0.043)	(0.039)
R&D Inc + rest	0.083*	-0.094*	-0.271***
	(0.046)	(0.049)	(0.047)
R&D Inc + MoA + Cit	0.898***	-0.735***	-0.432***
	(0.046)	(0.069)	(0.061)
Mkt Inc	-0.073***	0.077***	-0.081***
	(0.015)	(0.029)	(0.025)
R&D Inc + rest × Mkt Inc	-0.004	-0.024	0.143**
	(0.057)	(0.066)	(0.066)
R&D Inc + MoA + Cit × Mkt Inc	-0.891***	0.710***	0.702***
	(0.051)	(0.179)	(0.088)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	2005	2005	2005
Adj. R2	0.102	0.132	0.247

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1) Termination	(2) Progress	(3) New Patents	(4) Termination	(5) Progress	(6) New Patents
R&D non-Inc				0.285***	-0.251***	-0.343***
				(0.054)	(0.043)	(0.039)
R&D Inc + rest	0.003	0.009	-0.266***	0.182**	-0.211***	-0.279***
	(0.053)	(0.065)	(0.058)	(0.075)	(0.067)	(0.076)
R&D Inc + MoA + Cit	0.901***	-0.784***	-0.461***	-0.000	0.276***	0.212***
	(0.045)	(0.072)	(0.065)	(0.031)	(0.059)	(0.077)
Mkt Inc	-0.076***	0.079***	-0.079***	-0.075***	0.081***	-0.088***
	(0.015)	(0.029)	(0.025)	(0.015)	(0.029)	(0.025)
R&D Inc + rest × Mkt Inc	0.059	-0.098	0.154**	-0.015	-0.051	0.071
	(0.063)	(0.081)	(0.077)	(0.118)	(0.114)	(0.114)
R&D Inc + MoA + Cit × Mkt Inc	-0.889***	0.556**	0.738***			
	(0.055)	(0.218)	(0.114)			
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1853	1853	1853	1838	1838	1838
Adj. R2	0.064	0.126	0.249	0.107	0.132	0.271

Notes: This table shows the results of the OLS regressions measuring the effect of M&A depending on the combination of technological closeness (citations), MoA overlap, R&D and product market incumbency on all projects (panel A) and acquirer's and target projects (panel B). The treated projects are split in five exclusive groups and compared to the set of all never treated projects. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 Treated : R\&DIncRest_i + \\ + \beta_3 Treated : R\&DIncTwinMoA_i + \beta_4 MktInc_i + \\ + \beta_5 Treated : R\&DIncRest_i \times MktInc_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed (no connection exists between the target and acquirer's antidiabetics projects). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.17: Robustness: Tech closeness, MoA overlap, R&D and market incumbency with Jaffe closeness measure

Panel A: All

	All projects		
	(1) Termination	(2) Progress	(3) New Patents
R&D non-Inc (Jaffe)	0.273***	-0.253***	-0.342***
	(0.054)	(0.044)	(0.039)
R&D Inc + rest (Jaffe)	0.103**	-0.087*	-0.271***
	(0.048)	(0.050)	(0.048)
R&D Inc+MoA+Twin (Jaffe)	0.188	-0.408***	-0.353***
	(0.205)	(0.131)	(0.083)
Mkt Inc	-0.073***	0.077***	-0.080***
	(0.015)	(0.029)	(0.025)
R&D Inc + rest (Jaffe) × Mkt Inc	-0.013	-0.045	0.123*
	(0.059)	(0.068)	(0.067)
R&D Inc+MoA+Twin (Jaffe) × Mkt Inc	-0.240	0.465***	0.594***
	(0.207)	(0.173)	(0.139)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	2005	2005	2005
Adj. R2	0.098	0.133	0.249

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1) Termination	(2) Progress	(3) New Patents	(4) Termination	(5) Progress	(6) New Patents
R&D non-Inc (Jaffe)				0.272***	-0.249***	-0.342***
				(0.054)	(0.044)	(0.040)
R&D Inc + rest (Jaffe)	0.018	0.015	-0.273***	0.207***	-0.199***	-0.272***
	(0.055)	(0.069)	(0.061)	(0.076)	(0.064)	(0.075)
R&D Inc+MoA+Twin (Jaffe)	0.069	-0.283*	-0.242***	0.364	-0.645***	-0.553***
	(0.254)	(0.159)	(0.080)	(0.349)	(0.169)	(0.130)
Mkt Inc	-0.076***	0.079***	-0.080***	-0.076***	0.081***	-0.088***
	(0.015)	(0.029)	(0.025)	(0.015)	(0.029)	(0.025)
R&D Inc + rest (Jaffe) × Mkt Inc	0.048	-0.105	0.159**	0.021	-0.184*	-0.079
	(0.065)	(0.084)	(0.080)	(0.135)	(0.106)	(0.096)
R&D Inc+MoA+Twin (Jaffe) × Mkt Inc	-0.121	0.112	0.402**	-0.402	0.846***	0.833***
	(0.257)	(0.231)	(0.183)	(0.350)	(0.209)	(0.190)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1853	1853	1853	1838	1838	1838
Adj. R2	0.060	0.125	0.249	0.107	0.135	0.275

Notes: This table shows the results of the OLS regressions measuring the effect of M&A depending on the combination of technological closeness (Jaffe), MoA overlap, R&D and product market incumbency on all projects (panel A) and on acquirer's and target projects (panel B). The treated projects are split in five exclusive groups and compared to the set of all never treated projects. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 Treated : R\&DIncRest_i + \\ + \beta_3 Treated : R\&DIncTwinMoA_i + \beta_4 MktInc_i + \\ + \beta_5 Treated : R\&DIncRest_i \times MktInc_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed (no connection exists between the target and acquirer's antidiabetics projects). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.18: Robustness: Tech closeness, MoA overlap, R&D and market incumbency - alternative twin cutoff (p90)

Panel A: All

	All projects		
	(1) Termination	(2) Progress	(3) New Patents
R&D non-Inc	0.286***	-0.255***	-0.343***
	(0.054)	(0.043)	(0.039)
R&D Inc + rest	0.100**	-0.058	-0.269***
	(0.050)	(0.052)	(0.051)
R&D Inc + Twin + MoA	0.033	-0.434***	-0.312***
	(0.128)	(0.082)	(0.075)
Mkt Inc	-0.073***	0.076***	-0.080***
	(0.015)	(0.029)	(0.025)
R&D Inc + rest × Mkt Inc	-0.016	-0.061	0.116*
	(0.060)	(0.070)	(0.070)
R&D Inc + Twin + MoA × Mkt Inc	-0.014	0.350**	0.520***
	(0.144)	(0.141)	(0.121)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	2005	2005	2005
Adj. R2	0.100	0.134	0.249

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1) Termination	(2) Progress	(3) New Patents	(4) Termination	(5) Progress	(6) New Patents
R&D non-Inc				0.285***	-0.251***	-0.343***
				(0.054)	(0.043)	(0.039)
R&D Inc + rest	0.025	0.050	-0.268***	0.196**	-0.184***	-0.269***
	(0.059)	(0.070)	(0.065)	(0.078)	(0.070)	(0.051)
R&D Inc + Twin + MoA	0.002	-0.376***	-0.288***	0.059	-0.496***	-0.312***
	(0.147)	(0.099)	(0.054)	(0.236)	(0.138)	(0.075)
Mkt Inc	-0.076***	0.078***	-0.079***	-0.076***	0.081***	-0.080***
	(0.015)	(0.029)	(0.025)	(0.015)	(0.029)	(0.025)
R&D Inc + rest × Mkt Inc	0.042	-0.134	0.148*	0.006	-0.173	0.116*
	(0.068)	(0.086)	(0.083)	(0.140)	(0.116)	(0.070)
R&D Inc + Twin + MoA × Mkt Inc	-0.048	0.114	0.500***	0.010	0.522***	0.520***
	(0.150)	(0.188)	(0.154)	(0.258)	(0.202)	(0.121)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1853	1853	1853	1838	1838	2005
Adj. R2	0.060	0.127	0.250	0.107	0.133	0.249

Notes: This table shows the results of the OLS regressions measuring the effect of M&A depending on the combination of technological closeness, MoA overlap, R&D and product market incumbency on all projects (panel A) and on acquirer's and target projects (panel B). The *twin* definition differs from the baseline and corresponds to threshold of the 90th percentile of all-pairs similarity distribution. The treated projects are split in five exclusive groups and compared to the set of all never treated projects. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 Treated : R\&DIncRest_i + \\ + \beta_3 Treated : R\&DIncTwinMoA_i + \beta_4 MktInc_i + \\ + \beta_5 Treated : R\&DIncRest_i \times MktInc_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed (no connection exists between the target and acquirer's antidiabetics projects). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.19: Robustness: Tech closeness, MoA overlap, R&D and market incumbency - alternative twin cutoff (p99)

Panel A: All

	All projects		
	(1) Termination	(2) Progress	(3) New Patents
R&D non-Inc	0.286***	-0.256***	-0.343***
	(0.054)	(0.043)	(0.039)
R&D Inc + rest	0.088*	-0.071	-0.263***
	(0.048)	(0.051)	(0.049)
R&D Inc + Twin + MoA	0.152	-0.493***	-0.410***
	(0.173)	(0.103)	(0.067)
Mkt Inc	-0.073***	0.076***	-0.080***
	(0.015)	(0.029)	(0.025)
R&D Inc + rest × Mkt Inc	0.002	-0.053	0.114*
	(0.059)	(0.068)	(0.068)
R&D Inc + Twin + MoA × Mkt Inc	-0.206	0.469***	0.664***
	(0.175)	(0.161)	(0.130)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	2005	2005	2005
Adj. R2	0.100	0.134	0.249

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1) Termination	(2) Progress	(3) New Patents	(4) Termination	(5) Progress	(6) New Patents
R&D non-Inc				0.285***	-0.251***	-0.343***
				(0.054)	(0.043)	(0.039)
R&D Inc + rest	0.015	0.024	-0.267***	0.183**	-0.182***	-0.263***
	(0.056)	(0.068)	(0.061)	(0.078)	(0.068)	(0.049)
R&D Inc + Twin + MoA	0.104	-0.386***	-0.308***	0.188	-0.617***	-0.410***
	(0.235)	(0.150)	(0.076)	(0.271)	(0.116)	(0.067)
Mkt Inc	-0.076***	0.078***	-0.079***	-0.076***	0.081***	-0.080***
	(0.015)	(0.029)	(0.025)	(0.015)	(0.029)	(0.025)
R&D Inc + rest × Mkt Inc	0.052	-0.108	0.148*	0.034	-0.165	0.114*
	(0.065)	(0.083)	(0.080)	(0.132)	(0.111)	(0.068)
R&D Inc + Twin + MoA × Mkt Inc	-0.149	0.125	0.521***	-0.235	0.806***	0.664***
	(0.237)	(0.219)	(0.163)	(0.272)	(0.186)	(0.130)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1853	1853	1853	1838	1838	2005
Adj. R2	0.060	0.126	0.250	0.108	0.135	0.249

Notes: This table shows the results of the OLS regressions measuring the effect of M&A depending on the combination of technological closeness, MoA overlap, R&D and product market incumbency on all projects (panel A) and on acquirer's and target projects (panel B). The *twin* definition differs from the baseline and corresponds to threshold of the 99th percentile of all-pairs similarity distribution. The treated projects are split in five exclusive groups and compared to the set of all never treated projects. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 Treated : R\&DIncRest_i + \\ + \beta_3 Treated : R\&DIncTwinMoA_i + \beta_4 MktInc_i + \\ + \beta_5 Treated : R\&DIncRest_i \times MktInc_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed (no connection exists between the target and acquirer's antidiabetics projects). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.20: Robustness: Tech closeness, MoA overlap, R&D and market incumbency (clustering at firm level)

Panel A: All

	All projects		
	(1)	(2)	(3)
	Termination	Progress	New Patents
R&D non-Inc	0.286***	-0.255***	-0.343***
	(0.054)	(0.047)	(0.041)
R&D Inc + rest	0.098*	-0.064	-0.269***
	(0.054)	(0.048)	(0.046)
R&D Inc + Twin + MoA	0.046	-0.426***	-0.310***
	(0.211)	(0.143)	(0.112)
Mkt Inc	-0.074***	0.076**	-0.080**
	(0.017)	(0.030)	(0.036)
R&D Inc + rest × Mkt Inc	-0.008	-0.064	0.117
	(0.061)	(0.061)	(0.076)
R&D Inc + Twin + MoA × Mkt Inc	-0.095	0.425**	0.569***
	(0.213)	(0.174)	(0.214)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	2005	2005	2005
Adj. R2	0.100	0.134	0.250

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1)	(2)	(3)	(4)	(5)	(6)
	Termination	Progress	New Patents	Termination	Progress	New Patents
Treated: no dia				0.285***	-0.251***	-0.343***
				(0.056)	(0.047)	(0.042)
E_MoA_Twin_rest	0.023	0.038	-0.269***	0.196**	-0.183***	-0.274***
	(0.070)	(0.072)	(0.065)	(0.086)	(0.070)	(0.087)
Treated: Twin + MoA overlap	0.018	-0.359**	-0.285***	0.058	-0.496***	-0.342*
	(0.195)	(0.144)	(0.085)	(0.252)	(0.148)	(0.178)
Mkt Inc	-0.076***	0.078***	-0.079**	-0.076***	0.081***	-0.087**
	(0.016)	(0.029)	(0.037)	(0.016)	(0.030)	(0.037)
E_MoA_Twin_rest × Mkt Inc	0.045	-0.123	0.149	0.033	-0.199*	-0.078
	(0.076)	(0.085)	(0.097)	(0.143)	(0.108)	(0.102)
Treated: Twin + MoA overlap × Mkt Inc	-0.064	0.098	0.497***	-0.098	0.696***	0.622**
	(0.199)	(0.192)	(0.171)	(0.252)	(0.175)	(0.258)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1853	1853	1853	1838	1838	1838
Adj. R2	0.060	0.126	0.250	0.108	0.135	0.275

Notes: This table shows the results of the OLS regressions measuring the effect of M&A depending on the combination of technological closeness, MoA overlap, R&D and product market incumbency on all projects (panel A) and acquirer's and target projects (panel B). The treated projects are split in five exclusive groups and compared to the set of all never treated projects. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 Treated : R\&DIncRest_i + \\ + \beta_3 Treated : R\&DIncTwinMoA_i + \beta_4 MktInc_i + \\ + \beta_5 Treated : R\&DIncRest_i \times MktInc_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed (no connection exists between the target and acquirer's antidiabetics projects). Standard errors clustered at firm level in parentheses. $^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

Table 3.E.21: Robustness: Tech closeness, MoA overlap, R&D and market incumbency (clustering at transaction level)

Panel A: All

	All projects		
	(1)	(2)	(3)
	Termination	Progress	New Patents
R&D non-Inc	0.201***	-0.161**	-0.287***
	(0.066)	(0.064)	(0.052)
R&D Inc + rest	0.012	0.032	-0.168***
	(0.059)	(0.066)	(0.062)
R&D Inc + Twin + MoA	0.019	-0.327***	-0.096
	(0.131)	(0.113)	(0.113)
Mkt Inc	-0.122***	0.215***	-0.069
	(0.037)	(0.051)	(0.049)
R&D Inc + rest × Mkt Inc	0.086	-0.156*	0.154
	(0.067)	(0.089)	(0.101)
R&D Inc + Twin + MoA × Mkt Inc	-0.139	0.356**	0.600***
	(0.149)	(0.163)	(0.184)
MoA FE	Yes	Yes	Yes
Cohort	Yes	Yes	Yes
Stage	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes
Obs	638	638	638
Adj. R2	0.167	0.279	0.197

Panel B: Acquirers and targets

	Acquirer's projects			Target projects		
	(1)	(2)	(3)	(4)	(5)	(6)
	Termination	Progress	New Patents	Termination	Progress	New Patents
R&D non-Inc				0.152*	-0.115	-0.239***
				(0.080)	(0.080)	(0.063)
R&D Inc + rest	0.020	0.017	-0.268***	0.077	-0.081	-0.126
	(0.075)	(0.086)	(0.097)	(0.094)	(0.092)	(0.096)
R&D Inc + Twin + MoA	0.056	-0.234**	-0.144**	0.053	-0.407***	-0.070
	(0.169)	(0.105)	(0.064)	(0.214)	(0.142)	(0.238)
Mkt Inc	-0.061	0.165**	-0.137**	-0.189***	0.259***	0.082
	(0.054)	(0.074)	(0.066)	(0.064)	(0.091)	(0.082)
R&D Inc + rest × Mkt Inc	0.066	-0.101	0.306**	0.148	-0.295*	-0.222*
	(0.074)	(0.116)	(0.143)	(0.155)	(0.160)	(0.131)
R&D Inc + Twin + MoA × Mkt Inc	-0.178	-0.054	0.685***	-0.124	0.618***	0.344
	(0.195)	(0.141)	(0.133)	(0.239)	(0.214)	(0.299)
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes	Yes	Yes
Pat stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	334	334	334	304	304	304
Adj. R2	0.128	0.344	0.200	0.207	0.254	0.257

Notes: This table shows the results of the OLS regressions measuring the effect of M&A depending on the combination of technological closeness, MoA overlap, R&D and product market incumbency on all projects (panel A) and acquirer's and target projects (panel B). The treated projects are split in five exclusive groups and compared to the pooled counterfactual group, consisting of the technologically most similar, never treated projects of other firms in the same phase, assigned to each treated project. The underlying regression has the following form:

$$Pr(DV_i = 1) = \beta_0 + \beta_1 Treated : R\&DNonInc_i + \beta_2 Treated : R\&DIncRest_i + \\ + \beta_3 Treated : R\&DIncTwinMoA_i + \beta_4 MktInc_i + \\ + \beta_5 Treated : R\&DIncRest_i \times MktInc_i + \beta_6 Treated : R\&DIncTwinMoA_i \times MktInc_i + \epsilon_i.$$

Please note that R&D non-incumbents have empty antidiabetics portfolio before a transaction and hence, the acquirer's side cannot be analysed (no connection exists between the target and acquirer's antidiabetics projects). Standard errors clustered at transaction level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Chapter 4

Away from competition, away from "defeat zone": VCs and the direction of R&D [1]

Abstract This paper shows that venture capital funds (VCs) actively steer the direction of early-stage innovation in their backed start-ups to avoid the *"defeat zone"* - an area around big firms with launched drugs (incumbents) where it is not worth competing. VC-backed startups are less likely to pursue and more likely to terminate projects in existing product markets, future product markets, and technology spaces where big incumbents already operate. The effects are stronger when technology spaces overlap and the market power of incumbents is high. Instead of these projects, VC-backed startups are more likely to pursue and less likely to terminate breakthrough projects in new-to-be markets, which they eventually sell to non-incumbent firms. These results are fully driven by larger startups with enough room to steer R&D. In smaller startups with little room for steering, no evidence of the "defeat zone" is found as such startups pursue all their projects. The number of VC rounds and the volume of VC funding are the relevant channels for R&D steering. The findings indicate that VC investments are likely to result in breakthrough innovation in new markets but are unlikely to challenge big incumbent firms with market power in product and technology markets where they operate. From a policy perspective, VC-backed innovation is unlikely to solve the problem of the rising concentration of markets.

[1] This chapter is single-authored. I thank my committee members and seminar participants at DIW Berlin and KU Leuven for their valuable comments and helpful suggestions. I thank Jan Hlousek for the fruitful discussions that led to the emergence of this project and for his dedicated effort with data support. I am also grateful to the company Redstone which provided venture capital data. I acknowledge support from FWO through project 3H190094.

4.1 Introduction

Venture capital funds (VCs) play a central role in the lives of pharmaceutical startups, which are responsible for a substantial part of innovation (Chandy & Tellis, 2000). Being more than just financial intermediaries, VCs not only provide the needed liquidity,[2] but also active involvement in a startup's business development and management.[3] Research shows that VCs positively affect innovation output,[4] but remains silent on what types of innovation VC-backed startups pursue, particularly with regard to competition in product markets and the market power of competitors they face. Thus, we do not know whether and how venture capital affects the direction of innovation depending on competition. Given VCs' prominent role in the innovation landscape,[5] their actions can though have a profound and salient impact on the long-term competitiveness of markets and the welfare of consumers.

Few pieces of evidence provide examples of VCs' thinking in this respect. Albert Wenger, a managing partner at Union Square Ventures, said that VCs "*are now wary of entering into direct competition with giants like Google and Facebook*" (Schechter, 2018, p.2). Even more strikingly, Silicon Valley venture capitalist Peter Thiel famously proclaimed that "*Competition is for losers. If you want to create and capture lasting value, look to build a monopoly*" (The Wall Steet Journal, 2014, p.1). A policy report form the Stigler Center finds that *"a VC will usually be wary of outright investing in an innovative startup that will implicitly or explicitly compete head-on with a tech giant. Given the tech incumbent's ability to block or foreclose a threatening entrant, the chance of successful entry is tiny. VCs would rather invest in businesses that are creating new categories or solving common technical issues"* (Scott Morton *et al.*, 2019, p.8). Lastly, Hellmann & Puri (2000), the only empirical work on this topic so far, finds a positive association between

[2] In the US alone, $17 billion were invested by VC in the biotech sector in 2019 (BiopharmaDive, 2020).

[3] VC's involvement spans a broad range of activities. VCs are active managers (Gorman & Sahlman, 1989; Sahlman, 1990), often represented in the board of directors (Lerner, 1995) or even replacing the original founder as CEO (Hellmann, 1998; Hellmann & Puri, 2002). VCs also help shape new strategies and provide advice (Hellmann & Puri, 2000; Bernstein *et al.*, 2016), using their networks to garner resources for the company (Hellmann & Puri, 2002; Gompers & Lerner, 2004), including the recruitment of talented managers (Gorman & Sahlman, 1989).

[4] Kortum & Lerner (2001) is a pioneering study linking increases in venture capital to higher rates of patenting at the industry level, later confirmed by Ueda & Hirukawa (2008). Ruling out selection effects, Bernstein *et al.* (2016) confirms these findings at the company level. Other studies focus on the characteristics of innovation that VC promotes. Howell *et al.* (2020) document that VC-backed firms are more likely to file influential patents. Schnitzer & Watzinger (2017) shows that VC-backed startups produce patents of higher quality, and novelty, with resulting substantial technology spillovers. Beyond innovation, Kerr *et al.* (2014); Chemmanur *et al.* (2011); Puri & Zarutskie (2012) find that VC-backed startups are more successful, typically experiencing sales growth; they are also less likely to fail. Da Rin *et al.* (2013); Engel & Keilbach (2007) provide evidence that VCs help with product commercialization. Samila & Sorenson (2011) show that the benefits of VCs stimulate firm starts, employment, and aggregate income. Kwon & Sorenson (2021) argue that VC funding leads to a less diverse tradable sector and increases in income inequality in the region.

[5] VC funding has little alternatives as startups are particularly difficult to finance due to their highly uncertain prospects, lack of tangible assets to serve as collateral, and severe asymmetric information (He & Tian, 2018; Hall & Lerner, 2010; Zider, 1998). Some research even labels VCs as gatekeepers of pharmaceutical R&D (Fernald *et al.*, 2015).

VC and the speed of entering new product markets but no such association in existing product markets.[6] Hence, it seems more than a theoretical option that VCs do actively react to competition and direct R&D to avoid areas where incumbents with market power operate and where it is hard to compete.

This paper empirically investigates whether VCs steer the direction of R&D in their backed startups away from areas where large incumbents operate and where it is not worth competing - thus whether VCs avoid a "defeat zone" around large incumbent firms. I investigate this question not only in the product market space but also in the R&D and technology spaces. To do so, this paper uses the setting of the pharmaceutical industry - specifically the antidiabetics market - where VCs typically invest very early in the development process (Chandra, 2020). The strongly regulated and standardized development process alongside R&D taking place at the level of individual projects allows for looking under the hood of firms and seeing individual projects that the startups pursue. The dataset employed collects the universe of the corporate preclinical antidiabetics R&D between 1997 and 2017 conducted by startups - corresponding to the set of firms that were involved in antidiabetics R&D but did not have any launched drugs in the entire pharmaceutical industry. Each project is tracked from its inception throughout its preclinical development, ending either with progression to testing in humans or by staying in the preclinical phase. Every project also contains information on the identity of its primary developer and is linked to a comprehensive venture capital database of global VC deals, which maps projects to all VC transactions that affected the project during its lifetime. Projects' bio-chemical characteristics, namely the mechanism of action ("MoA"), are used to define relevant product markets and closeness of activities in the R&D space.[7] Patents linked to each R&D project - a unique feature of this data - then measure projects' technological properties and define the relationships to competitors' patent portfolios.

Since the decision of VCs to back a particular firm is not random, the empirical section starts by presenting results from a stepwise selection regression. This regression tests whether the key competition and technology project variables, after controlling for a range of control variables and fixed effects, predict VC investment. Finding no evidence that selection is taking place on these key observables, the paper proceeds to establish the main findings.

VC-backed startups avoid "defeat zone" around big incumbents as they are less likely to pursue and more likely to terminate projects in product markets, R&D spaces, and

[6] Two related noteworthy contributions are the papers of Li *et al.* (2020) and Gans & Stern (2003). Li *et al.* (2020) focus on the role of VC and common ownership. They find that common ownership leads to reductions in the duplication of R&D across competing projects and that investors shut down lagging R&D projects, restrict their funding, and encourage their firms to pivot to new R&D projects. Gans & Stern (2003) then broadly discuss what drives the decisions of startups to compete against incumbents or not in the product market space.

[7] Mechanism of action (MoA) defines the biochemical process through which a drug produces the desired effect in the body. Specific MoAs are associated with the type of side effects they produce. Diabetologists perceive drugs with the same MoA targeting a specific disease as substitutable products.

technology spaces where big incumbents already operate. The reactions in the technology space are stronger with a narrower definition of technology links or with a rising market power of the incumbent firms. Instead of these, VC-backed startups pursue projects with breakthrough potential, particularly in new-to-be and monopolizable markets. VCs exit from these breakthrough projects by selling them to other firms, but not to the large incumbent ones. The number of VC rounds and the volume of VC funding are the relevant channels for the VCs steering of R&D. The results are fully driven by sufficiently large startups where there is room to steer and redirect R&D. In contrast, in small startups with little room for steering, no evidence of the "defeat zone" is found as all VC-backed projects are more likely to be pursued and not more likely to be terminated.

The main threat to these findings is the presence of unobserved factors that can be correlated with the error term, giving rise to endogeneity. To address this concern, I employ an instrumental variable approach to check the main effects of VC backing on the project's progression rate. The IV approach uses the well-established idea that VC capital investing is localized - thus geographical distance and location serve as shifters of the likelihood to receive VC funding, whilst at the same time affecting the progression of a project only through the proximity between the VC and the project's developer. Accounting for endogeneity, the main results remain robust.

This paper makes several important contributions. First, it contributes to the VCs literature by providing detailed evidence on how VCs influence the direction of innovation, helping to fill the gap in our understanding of what VCs do inside of firms (Gompers *et al.*, 2020). Second, it contributes to the emerging literature on the links between venture capital and competition. By finding evidence for the "defeat zone," it complements ongoing policy and academic discussions on the phenomenon of the "kill-zone" (Kamepalli *et al.*, 2020; Koski *et al.*, 2020; Motta & Shelegia, 2021). Third, from the policy perspective, this paper highlights that although VCs help to push frontiers of treatments forward, the pattern of VCs' behaviour neither helps to make existing markets more competitive, nor to challenge the incumbents in the product, R&D, and technology spaces where they operate. Policymakers should keep this trade-off in mind when providing public money to support venture capital in spurring innovation (Lerner, 2009).

The remainder of the paper is structured as follows. Section 4.2 present institutional details and discusses the hypotheses. Section 4.3 provides an overview of the construction of data. Section 4.4 defines key variables and provides basic summary statistics. Section 4.5 presents results and section 4.6 concludes.

4.2 VC setting and expected effects

4.2.1 Venture capital funds

Venture capital funds are professional investors who raise funds from institutions, such as pension funds, insurance companies, foundations, or high net worth individuals (NVCA, n.a.).[8] These funds are then invested in companies in return for a minority equity stake. VC funds are well known for their willingness to invest in young, uncertain, but promising ventures with the potential to generate multiples of the invested stakes when the venture capital fund exits the venture. By the nature of the VC's "high risk/high gain" business model, their returns are highly skewed. While the majority of risky investments fail and are written-off, a few deals generate multiples of the investment and account for the bulk of overall VC returns.[9] Harris *et al.* (2014) finds that these are large enough of to outperform public markets by more than 3% annually.

For startups receiving VC funding, the role of VCs is not only limited to providing financial backing. Hellmann & Puri (2002) provide evidence that VC's role goes beyond traditional financial intermediaries and is generally associated with a variety of professionalization measures. VCs are active investors involved with management and monitoring of their portfolio companies (Gorman & Sahlman, 1989; Sahlman, 1990), with greater VC control associated with increased management intervention (Kaplan & Strömberg, 2004). VCs are often represented in the board of directors (Lerner, 1995) and sometimes even replace the founder as CEO (Hellmann, 1998; Hellmann & Puri, 2002). VCs also help shaping new strategies (Hellmann & Puri, 2000; Bernstein *et al.*, 2016) and use their networks to garner resources for the company (Hellmann & Puri, 2002; Gompers & Lerner, 2004), including the recruitment of talented managers (Gorman & Sahlman, 1989). Further, these findings are refined in a large-scale survey by Gompers *et al.* (2020), which finds that VCs indeed provide many services to their portfolio companies, including post-investment strategic guidance (87%), connections to investors (72%), connections to customers (69%), operational guidance (65%), as well as board members (58%) and employee (46%) hiring. The VC's involvement is also substantial in terms of the allocated time - an average VC spends 18 hours per week working together with their portfolio firms. This evidence demonstrates that VCs are active investors with strong control rights who not only provide needed liquidity but also steer the activities of ventures they have invested in. Ultimately, VCs have significant decision power in determining the direction of startup's operations.

[8] Please refer to Sahlman (1990) for a more detailed overview of the VC institutional details.

[9] For example, Prencipe (2017) find that among EU VC investments, about 70% of investments are either written-off or sold for an amount below cost, 8% are sold at cost, and only the remaining 22% of the liquidity events were profitable. In addition, only a quarter of the profitable exits have returned more than 5 times the investment.

4.2.2 Expected effects

A central element of this paper is to investigate whether VCs avoid the "defeat zone" in the product market, R&D space, and technology space, thus pursuing other projects instead and influencing the direction of innovation. The next section provides an overview of the expected effects.

Product market. Upon VC's investment into a company, a startup might be in a possession of a preclinical R&D project facing existing competition in its relevant product market. Holding all other factors constant, the presence of a launched product in a market reduces profits that a venture capitalist can earn if it pursues such a project. For example, assume that we have two markets with identical sizes and other characteristics, except there is no product in one and an existing product in the other. A VC would be more likely to pursue a preclinical project in the market with no existing competition that it can monopolize, compared to the market where it would face existing product market competition and would have to share profits in a duopoly. Therefore, it is expected that VC-backed projects of startups are more (or less) likely to progress in development depending on the absence (presence) of a launched product, compared to such non-VC-backed projects. This would hint at avoiding the "defeat zone" in the product market space.

R&D space. Large firms have deeper financial pockets and can benefit from economies of scale and scope when it comes to running R&D, clinical trials, or obtaining regulatory approvals (Arroyabe, 2021; Bena & Li, 2014; Szücs, 2014; Danzon *et al.*, 2007).[10] Upon investing into a company, a VC fund might see projects that are developed in the future relevant markets where the incumbents operate as an indicator of hardship to compete since it is unlikely that the small VC-backed startup would be able to compete with these firms. Therefore, this decreases the expected value of projects in the R&D areas where big incumbents are active. Thus, like the product market, the expectation is that VC-backed preclinical projects of startups are more (less) likely to progress if a big incumbent firm does not develop (develops) projects in the same future relevant market, compared to such non-VC-backed projects. This would hint at avoiding the "defeat zone" in the R&D space.

Technology space. Almost any pharmaceutical R&D project is protected by patents (see section 4.3.1 for more detail). Patents reference other patents on the front page of every document to identify the relevant prior art that has already been claimed by others and, therefore, that cannot be claimed by the current patent.[11] Hence, playing an important legal function, backward patent citations delimit the scope of property rights and may be used to preclude the issuance of a patent or limit the scope of the protection. Czarnitzki *et al.* (2020) show that firms indeed use patents for blocking competitors' inno-

[10] Malek *et al.* (2021) show that leaders - the four largest firms in the antidiabetics industry - are substantially less likely to engage in acquisitions and rely on in-house R&D, indicating that they have a relative comparative advantage in pursing in-house R&D.

[11] For patents filed at the US patent office, the patent applicant has a legal duty to disclose any knowledge of the prior art that she may have, but it is ultimately the patent examiner - an expert in a field - who evaluates and supplements the final set of citations (Jaffe *et al.*, 1993).

vation activities as such strategies are associated with higher firm value. Since particularly large incumbent firms can leverage their scale and market power to protect their positions in a given technology area where they enjoy legal monopoly rights, for example by copying or engaging in anti-competitive practices (Motta & Shelegia, 2021), a VC can decide that pursuing a project in a technology area where big incumbent already operates is less worthy. Therefore, the expectation is that VC-backed preclinical projects of startups are less likely to progress if a big incumbent firm develops projects in the same technology space, compared to such non-VC-backed projects. These effects are expected to be stronger with the increasing market power of the big incumbent firms since such big incumbents would have even stronger incentives to protect their existing markets against cannibalization, making projects even less attractive from the VC's perspective. These effects are also expected to be stronger with a narrower definition of the technology space and, hence, higher relevance of the technology content of the patent that a particular project cites. All these effects combined hint at avoiding the "defeat zone" in the technology space.

Breakthrough. Patents also allow characterization of each project in terms of breakthrough, or "high risk/high gain" potential. Breakthrough inventions introduce technological novelty and are important improvements to the current state of treatment of any disease. While being riskier (Fleming, 2001; Hall & Lerner, 2010; Verhoeven *et al.*, 2016), such inventions are more valuable and earn higher revenues conditional on being successful (Krieger *et al.*, 2018). This aligns with VC's strategy that they generally pursue. Thus, it is expected that upon investment into a company, a VC fund will be more likely to pursue these breakthrough projects compared to those non-VC-backed projects. Similarly to the above, it is also expected that these effects will be stronger in cases where the breakthrough projects face no product market competition, as in such cases the new-to-be markets can be monopolized, yielding higher expected value than if a project has to compete.

4.3 Data and variables

This paper employs a granular dataset, tracking technological characteristics, market characteristics, and VC transaction information for 783 preclinical R&D projects developed by startups to treat diabetes between 1997 and 2017. Startups are firms purely conducting R&D with no launched products in the entire pharma sector. The focus on preclinical development relates to the fact that VC funding is most relevant in this space. First, VC funding in the healthcare sector is generally directed at earlier stage firms than VC funding in other sectors (Chandra, 2020).[12] Second, startup firms are financially restricted, incurring costs to run R&D but not yet having any revenue streams from launched drugs.

[12]The analysis of the full diabetes sample reveals that 80% of all first-round deals indeed happened in the preclinical phase. Anecdotal evidence from the industry highlights the importance of investing in preclinical assets. For example, Todd Foley from a VC fund MPM Capital Inc said that "*As early-stage investors, we're willing to take significant biology risk and early development risk [as] you start to see the value is quite high after proof of concept.*" (BiopharmaDive, 2017, p.2).

These firms are natural candidates to seek VC funding.[13] Last, VCs have greater control over the startup during the project's early stages, making any potential effects of VCs actions more visible.

4.3.1 Core database

The backbone of the dataset is the private Pharmaprojects database from Citeline, which comprises a list of global corporate R&D activity in the pharmaceutical industry at the project level between 1997 and 2017. Pharmaprojects also contain lists of several basic variables, including a brief project description, information on the project's therapeutic area, MoA, and the developer's name.

To recover information on the project's initiation and the outcome of each project (progression to Phase I clinical trials), projects were matched to clinical trial databases.[14] To characterize the technological properties of projects and their relationship to other projects in the technology space, each R&D project was matched to its underlying US patents. Since regulation only requires patent disclosure at the time of the market approval, the identity of the patents relating to preclinical projects is unknown. In this dissertation, an algorithm was developed that uses the (bio)chemical and pharmacological properties of projects to assign patents. In brief, for each project in development, we search for patents that were filed at the USPTO by its developers between the inception and termination dates of a project. Depending on the type of project and information available, we are complementarily using various matching techniques. For small-molecule chemical drugs, we employ several crossroads between chemical, patent, and medical databases to establish project-patent links. For large molecule drugs relying on proteins, we follow the approach of Sampat & Williams (2019) by linking gene identifiers from the Pharmaprojects database to a list of protein and nucleotide sequences; thereafter we match these sequences against the census of sequences disclosed in the US patents to establish patent links. To complement these approaches and increase the matching rates, we also use natural language processing methods and data from Arts *et al.* (2021) to connect projects to patents based on keywords relating to their MoA. Please see a detailed description in Appendix B of Chapter 2.

4.3.2 Venture capital deals

All preclinical projects of the startups were matched to a database of venture capital deals, allowing identification of projects that were subject to at least one round of VC funding and those that were not. The venture capital data was provided by a Berlin-based venture capital firm Redstone. Redstone is a data-driven VC fund owning a comprehensive database of global VC transactions, spanning the period from 1997 through at least 2017;

[13] This assertion is also supported by the analysis of the unrestricted dataset with all phases. Pipeline firms were primary candidates to seek VC support, accounting for 80% of VC transactions.

[14] Chapter 2.2 provides more details about the matching procedure.

the end date of the sample. The data sources used for the construction of Redstone's database include both private and public databases - for example, crunchbase.com and tracxn.com - but rely also on web scrapers and crawlers, complementing these databases with information freely available on the Internet. The main information contained in Redstone's database relates to the identity of VC funds involved in the transaction, the date of the transaction, the identity of the backed company, information on the funding amount and stage of the VC round, as well as geographical information on the location of VC funds and the backed firms.[15] Unfortunately, additional information, like the characteristics of the VC funds involved in each transaction, is not available.

To match the VC deals to the relevant funded firms in the sample, I first run a fuzzy string matching algorithm, finding matches between the names of the companies. An extensive manual check was performed to ensure that only correct matches are kept. Apart from the indicator for VC backing, this approach also allows for collecting the number of rounds occurring during the projects' development (both preclinical and beyond) as well as the volume of funding raised in these transactions.

4.4 Key variables and descriptive statistics

The resulting dataset entails 783 preclinical projects developed by 582 startups, of which 151 (20%) projects developed by 118 companies were VC backed.[16] All information about projects' characteristics along product, R&D, technology, and breakthrough characteristics as well as about project outcomes is condensed into a single observation per project.

The treatment variable in this paper is a binary indicator equal to one if a project received its first VC transaction during the project's preclinical stage.[17] The key outcome variable in this paper is a binary indicator marking project's progression into the next development phase - Phase I of the clinical testing. For successful transitions indicating active development, the variable takes a value of one. In the opposite case, it takes the value of zero.

4.4.1 Key explanatory variables and the "defeat zone"

To check the hypotheses set out in section 4.2, several explanatory variables must be created, allowing for exploration of the heterogeneity in the VC effect. The explanatory

[15] As explained later, geographical information is needed for the construction of instrumental variables. Information on the geographical location of the non-VC-backed startups was collected manually.

[16] The magnitude of the VC rate seems to be in line with other evidence from the literature. For example, Cunningham (2017) finds that, on average, 15% of patenting medical device startups eventually receive VC funding.

[17] Existing research routinely uses only the first VC transaction in various settings. For example, Hsu (2004) empirically evaluates the certification and value-added roles of reputable venture capitalists by analyzing financing offers made by competing VCs at the first professional round of startup funding. Cunningham (2017) investigates the role of novelty in VC funding of medical device startups, whereby VC funding is defined by the first VC rounds a firm obtained.

variables for the VC-backed projects mirror the state of the world before the first VC transaction. The explanatory variables for the non-VC-backed projects reflect the state of the world at the beginning of the project's preclinical phase since a unique point in time cannot be identified for these due to a lack of a VC transaction.[18]

Competition in R&D and product markets. Like the framework used by antitrust authorities, this paper assesses competition interactions within "relevant markets."[19] Following the established practice in merger proceedings,[20] relevant markets are defined using the drug's therapeutic area (antidiabetics) in combination with its underlying mechanism of action (Seldeslachts *et al.*, 2021).[21] To assess whether a particular preclinical project faces product market competition (variable *MoA launched*), I create an indicator variable equal to one if there is at least one launched drug in the project's MoA and zero otherwise.[22] To assess competitive interactions in the R&D space, I construct two indicator variables, taking into account both the presence of competition in the future relevant product market and the level of market power that a developer of such a competing project holds. The first variable, *MoA: R&D Big inc*, is equal to one if a big incumbent firm was developing a project based on the same MoA. Thus, this variable captures competition in the future relevant markets by big incumbent firms with market power. The second variable, *MoA: R&D Small/Big non-inc*, is equal to one if a small or big non-incumbent firm was developing a project based on the same MoA. Thus, this variable captures competition in the future relevant markets by firms with no market power in antidiabetics (but with existing market power in other therapeutic areas).

Technology links. To quantify technology links, this paper employs backward patent citations. Several integer variables are constructed for each project, capturing how the project relates to the antidiabetics patent portfolios of competitors. The first variable, *Cites big inc*, counts the backward references between project's patent(s) and the antidiabetics patent portfolio of the big incumbents, thus capturing links to technology areas where firms with market power operate. The second variable, *Cites small or big non-inc*, counts the backward references between project's patent(s) and the antidiabetics patent portfolio of big non-incumbent firms or small firms, thus capturing links to technology areas where firms with no market power in antidiabetics operate.

[18] All results remain robust when this definition is changed to an alternative one, particularly to a random time-point in each non-VC-backed project's preclinical development.

[19] The European Commission defines a relevant product market as a space comprising all products that are regarded as interchangeable or substitutable by the consumer, because of the products' characteristics, their prices, and their intended use (Commission *et al.*, 1997). Competition law also defines a relevant geographic market as an equally important element for the assessment of competitive interactions. However, in the context of this paper, a geographical definition is not relevant, since the paper only considers the US product market.

[20] See for example the decisions of the European Commission in cases M.7275 or M.9294

[21] An interview with a diabetologist confirms this definition. Diabetologists would typically switch to another drug in the same class as the first option. A switch to a drug in a different class would only occur if the treatment is not sufficiently effective or if it produces intolerable side effects.

[22] As of 2017, there were 68 brand name launched drugs on the US market, falling in 8 MoAs. The most numerous MoAs include Insulins, K-ATP channel antagonists, and DDP-4 inhibitors. A complete overview is provided in table 4.C.16 in Annex 4.A.

To examine how different levels of competitors' market power matter in the technology space, I construct two additional variables, using the market structure of the antidiabetics market and varying degrees of market shares between companies. Since only four firms dominate the antidiabetics market - namely Novo Nordisk, Sanofi, Merck & Co., and Eli Lilly; collectively labelled as "leaders" - the first variable *Cites leaders* counts the backward references between the project's patent(s) and the antidiabetics patent portfolio of these four firms. Therefore, it captures links to technology areas where the firms with the largest market power operate. The second variable *Cites other big-incumbents* then counts the backward references between project's patent(s) and the antidiabetics patent portfolio of other incumbents, excluding the leaders. Thus, this variable captures links to technology areas where firms with moderate levels of market power operate.

To assess how the strength of VC reactions varies with the degree of technological similarity between a project and the patents in the competitor's portfolio that the project cites, an analogue of *Cites big inc* and *Cites small or big non-inc* variables is constructed with a narrower definition of technology links. Backward references are only counted in cases when they refer to patents in the competitors' antidiabetics portfolio in the same IPC class. Thus, these variables capture links to the same technology areas where various competitors operate.

Breakthrough nature. The breakthrough property of projects is measured based on information contained in the project's patents. More specifically, we utilize the Novelty in Technological Origins (NTO) indicator developed by Verhoeven *et al.* (2016) to classify each project as being a breakthrough or not. A patent scores on NTO if it makes a combination between its own IPC code and an IPC code from its referenced patents that have not yet occurred in the years before the application year of the patent (Verhoeven *et al.*, 2016). Since a project can have more than one patent assigned, a project scores on the NTO if at least one of the project's patent scores on the NTO indicator.[23]

4.4.2 Descriptive analysis

Table 4.1 presents how the dependent, explanatory, and control variables differ depending on the treatment.[24] Consistent with the evidence from the pharma industry, around 40% of all preclinical projects of the startups moved forward when they were not VC backed and 34% moved forward in case they were VC backed. This difference is insignificant.[25]

For most of the key explanatory variables, the VC-backed and non-VC-backed projects are not significantly different. About a quarter of the projects faced competition in the product market (*MoA launched*) and close to 90% of the projects faced competition in their MoA either from the incumbents or non-incumbents/small firms. Note that these

[23] Since the probability to score on novelty mechanically increases with a rising number of assigned patents, I always control for the number of project's assigned patents in the regression analysis.

[24] Table 4.B.15 in Annex 4.A provides definitions for all variables.

[25] Takebe *et al.* (2018) report a preclinical continuation rate of 44% for endocrine diseases.

Table 4.1: Summary statistics

	Non VC		VC		
	(1)	(2)	(3)	(4)	(5)
	Mean	Count	Mean	Count	p-value
Progression	0.40	632	0.34	151	0.19
Key explantory variables:					
NTO	0.34	632	0.42	151	0.04
MoA launched	0.25	632	0.24	151	0.80
MoA: R&D Big inc	0.87	632	0.86	151	0.80
MoA: Small/Big non-inc R&D	0.91	632	0.91	151	0.97
Cites big inc	0.56	632	1.38	151	0.03
Cites big inc - leader	0.19	632	0.53	151	0.02
Cites big inc - other	0.37	632	0.85	151	0.08
Cites small or big non inc	0.64	632	1.11	151	0.29
Control variables:					
Age of firm	11.15	632	11.48	151	0.78
Single pharma	0.16	632	0.11	151	0.10
Single dia	0.63	632	0.62	151	0.84
New chem	0.50	632	0.54	151	0.44
Dia 1 and 2	0.10	632	0.07	151	0.34
Front runner	0.27	632	0.32	151	0.26
RD projects in pharma	5.98	632	5.62	151	0.43
RD projects in dia	2.36	632	1.95	151	0.04
Patent nb.	2.60	632	2.43	151	0.73

two variables are not exclusive; in fact, frequently both big incumbents and big non-incumbents/small firms are working on projects in a particular MoA. For this reason, both must be taken into account simultaneously in the analysis. The only exception in which VC projects are significantly different from non-VC projects is the breakthrough indicator (*NTO*). Forty-two per cent of projects receiving VC funding were NTO, whereas, for the non-VC sample, this figure corresponds to only 34%. For the links between projects and technology portfolios of competitors, the only significant difference emerges in cases where patents of big incumbent firms (and also leaders and other big incumbent firms) are cited. On average, a non-VC-backed project cited 0.5 patents of big incumbents, whereas a VC-backed project cited on average 1.11 patents of the big incumbents.[26] However, significant heterogeneity exists among the projects citing at least one patent, warranting analysis with continuous variables.[27]

4.5 Results

4.5.1 Selection

The decision of venture capitalists to back a firm is not random. Petty & Gruber (2011) provide an extensive overview of the literature on which factors play role in the decisions of VCs to select companies. Their review indicates that VCs emphasize characteristics related to the venture's management team, the market, the product or service, and the venture's financial potential when making investment decisions. In an extensive survey, Gompers *et al.* (2020) find that the management/founding team is what VCs put the greatest emphasis on. Business-related factors, like the business model, product, market, and industry, are also frequently mentioned as important but were rated as the most important by only a third of the VC firms. Using a randomized experiment, Bernstein *et al.* (2017) also confirm that investors strongly respond to information about founding teams and human assets in the early stage ventures.

Since this paper uses a dataset providing very detailed information on key observable characteristics of projects in technological space, R&D space, and product market space, the analysis starts by directly exploring if these observable project characteristics play a role in the VC backing decision. In contrast, since no information is available on founding teams and their characteristics - which seem to matter the most for VC selection in early development stages - the robustness of the main results is checked in an instrumental regression framework in section 4.5.6.

The descriptive evidence already indicated that only a few of the observed variables

[26] Please also note that a project might reference patents from multiple types of competitors at the same time. For this reason, the variables capturing links in technology space to different types of competitors also must be taken into account simultaneously.

[27] Table 4.E.18 in the appendix gives a distribution of the number of backward citations. A robustness analysis confirms that outliers are not driving any of the results in this paper.

seem to significantly influence the decision of the VCs to back projects. This section employs a regression framework and performs stepwise selection equations by controlling for the above variables and the number of additional fixed factors. Equation 4.1 presents the model employed:

$$Prob(VC_i = 1) = \alpha + \beta X_i + \gamma W_i + \xi FE_i + \epsilon_i \tag{4.1}$$

The dependent variable is a binary indicator equal to one if a project received its first VC transaction during the project's preclinical stage. The vector X includes explanatory variables and the vector W is a vector of control variables. The vector FE encompasses a series of fixed effects. First, I include a cohort fixed effects that groups together projects initiated around the same time. This controls for technological trends. The second set of fixed effects relates projects' mechanism of action and controls for fixed, distinctive features specific to each group of drugs with the same MoA.[28] The third set of fixed effects controls for the technological classes that the project's patents encompass. Each project scores on as many IPC dummies as the underlying patents refer to. The last set of fixed effects controls for variation related to the geographical area of the developer's headquarters, aggregated in four distinct regions - Southeast Asia (Japan, China, India, Singapore, and Taiwan), Europe, Northern America (Canada and the US), and the rest of the world.

Table 4.2 presents robust evidence that once controlling for additional factors and fixed effects, none of the key explanatory variables of interest drives the decision of VCs to invest in a project. The coefficients are both economically and statistically insignificant in all specifications and robust to alternative estimation methods.[29]

4.5.2 VC effects

4.5.2.1 Descriptive evidence

In line with the "defeat zone" hypothesis, a lower progression rate of VC-backed projects is expected where they face presence and competition from big incumbent firms. In contrast, a higher progression rate of VC-backed breakthrough projects is expected as those are developed instead. Figure 4.1 provides descriptive evidence by exploring the heterogeneity in the progressions rates depending on VC backing status and type of project. Whilst 40% of all non-VC-backed projects progressed, this rate drops slightly to 34% for all VC-backed projects. Table 4.1 demonstrates that this baseline difference is insignificant. Looking at the various dimensions, the most profound difference is found for projects facing competition in the product market space. For those, the progression rate of VC-backed

[28] For example, research shows that MoA is related to types of side effects (Berger & Iyengar, 2011) and suitability for treatment in different patient populations (Association *et al.*, 2019; Chaudhury *et al.*, 2017). Further, the development of every project is centred around its MoA.

[29] Table 4.E.19 in Section 4.A reports the full results, including the coefficients on all control variables employed in the estimation. Table 4.E.20 in Section 4.A performs a robustness check using a probit model.

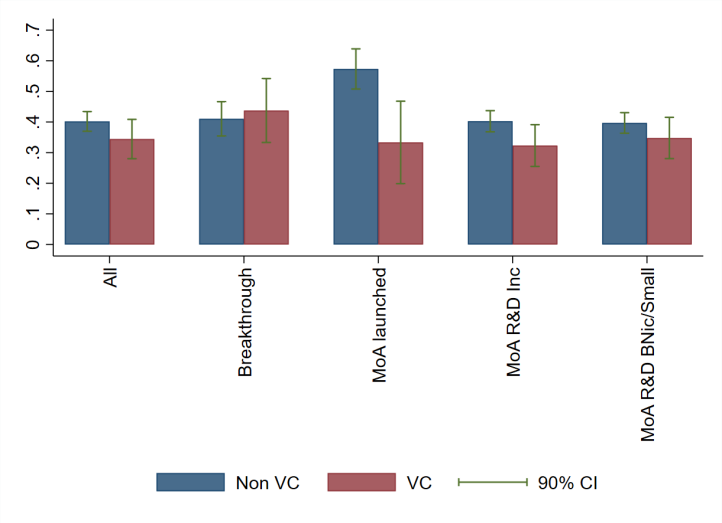
Table 4.2: Step-wise selection equation

	(1)	(2)	(3)	(4)	(5)	(6)
MoA launched	0.007	0.008	-0.011	0.012	0.012	0.012
	(0.037)	(0.040)	(0.040)	(0.064)	(0.063)	(0.063)
NTO	0.053	0.033	0.033	0.026	0.020	0.018
	(0.033)	(0.040)	(0.039)	(0.041)	(0.040)	(0.041)
MoA: R&D Big inc	-0.011	0.011	-0.007	-0.009	0.002	-0.001
	(0.048)	(0.051)	(0.052)	(0.059)	(0.059)	(0.060)
MoA: Small/Big non-inc R&D	0.006	-0.031	-0.030	-0.015	-0.025	-0.022
	(0.057)	(0.060)	(0.058)	(0.066)	(0.065)	(0.066)
Cites big inc - leader						0.009
						(0.013)
Cites big inc - other						0.002
						(0.006)
Cites small or big non inc						0.000
						(0.004)
Patent nb.	0.001	-0.003	-0.002	-0.002	-0.002	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
IPC FE	No	Yes	Yes	Yes	Yes	Yes
Cohort FE	No	No	Yes	Yes	Yes	Yes
MoA FE	No	No	No	Yes	Yes	Yes
Country FE	No	No	No	No	Yes	Yes
Obs	783	783	783	783	783	783
Adj. R2	-0.003	-0.001	0.022	0.025	0.035	0.033

Note: OLS regression. This table estimates how project characteristics affect the likelihood to receive VC backing. The dependent variable is a dummy variable equal to one if a project was VC-backed. The independent variables include various key predictors. Control variables include the age of the firm, number of projects in diabetes and pharmaceutical industry, indicator variables for projects based on new chemical compounds, targeting both types of diabetes, and being a front runner. The sample encompasses preclinical projects of startups. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

projects is substantially lower than for those not backed by VC. Similar patterns hold for projects facing incumbents in the same MoA in the R&D space. The reverse is then true for the breakthrough projects (NTO), as the average progression rate of the VC-backed is slightly higher compared to non-VC-backed projects. This descriptive evidence already hints at patterns consistent with the "defeat zone."

Figure 4.1: Heterogeneity in the progression rate



4.5.2.2 Econometric implementation

To empirically establish how VC backing affects the progression rates of the preclinical projects, I compare the mean differences in the progression rate between the VC-backed and non-VC-backed projects. To assess how heterogeneous the effects along various dimensions are, the explanatory variables are interacted with the treatment variable. The canonical model has the following form:

$$Prob(Progress_i = 1) = \beta_0 + \beta_1 VC_i + \beta_2 VC_i \cdot X_{ji} + \beta_4 X_{ji} + \beta_4 W_i + \beta_5 FE_i + \xi_i. \quad (4.2)$$

The VC treatment indicator corresponds to a binary variable equal to one if a project has received VC backing in the preclinical phase. The variable X_{ji} then stands for each of the key explanatory variables - the presence of existing competition ($j = 1$), the presence of big incumbent firms in the same future relevant market ($j = 2$), the presence of big non-incumbent or small firms in the same future relevant market ($j = 3$), breakthrough nature of the project ($j = 4$), technology links to big incumbents' technological portfolio ($j = 5$), and technology links to big non-incumbents' or small firms' technological portfolio

($j = 6$). The variables contained in the vectors W and FE coincide with those described in the section 4.5.1.

Similarly to Chapter 3, due to the relatively small sample, the necessity to accommodate a large set of fixed effects, and the main focus on the interaction and heterogeneity in the treatment, the regressions are estimated using ordinary least squares by default. Statistical inference is based on robust standard errors.[30]

Main results. VC backing in the preclinical phase is not associated with better or worse progression outcomes as such (see column (1), Table 4.3). However, the remainder of the table 4.3 clearly shows that VCs actively steer R&D of their backed startups depending on the type of project and provide strong support for the presence of a VC-induced "defeat zone."

Controlling for other factors, column (2) shows that VC-backed projects developed in relevant markets where big incumbents operate R&D are substantially less likely to be brought to clinical testing compared to such non-VC-backed projects. In contrast, no such effect is observed for VC-backed projects developed in relevant markets with competitors lacking market power. In column (3), the coefficient for the interaction term $VC \times$ *MoA launched* is negative and statistically significant at the 5% level. Thus, the presence of a launched product in a project's MoA substantially decreases the progression rates of VC-backed projects compared to such non-VC-backed projects. These results are highly indicative of a VC-induced "defeat zone" in the R&D and product market spaces.[31] Column (4) demonstrates that one reference to patents of big incumbent increases the progression rate by 1.4pp; however, the result reverses for projects that are VC-backed. In that case, one patent reference by a VC-backed project to patents of big incumbents decreases the progression rate by 1.4pp (-2.8+1.4). In contrast, referencing a patent of big/small firms lacking market power is not associated with any negative or positive impacts on the progression. These results hint at a VC-induced "defeat zone" in the technology space where big incumbent firms already operate. Unlike the previous cases, the results in column (5) show that VC-backed breakthrough projects are associated with a significantly higher progress rate compared to non-VC-backed projects. Therefore, these seem to be the only type of projects that VC investors actively pursue.

Taken together, the lack of progress for the VC-backed projects in the R&D, product market, and technology areas where big incumbents operate, and the excess of progress for breakthrough VC-backed projects, hints at active steering of R&D by the VC investors. Areas, where it is not worth competing against big incumbent firms, are avoided whilst

[30] The main results section presents a discussion of robust standard errors assumption as well as various sensitivity checks.

[31] The levels of the explanatory variables in columns (2) and (3) relating to competition in the relevant product markets are not associated with different progression rates. Thus, on average, startups are not significantly more (less) likely to develop their preclinical projects depending on the absence (presence) of competition in the relevant markets. The next sections illustrate that this average zero effect disappears if we look at the activities of firms based on their size, suggesting that smaller and larger startups follow different strategies.

Table 4.3: Effects of VC - various characteristics

	(1)	R&D (2)	Launch (3)	Citations (4)	Breakthrough (5)	All together (6)
VC	-0.027	0.051	0.021	-0.008	-0.088	0.011
	(0.044)	(0.159)	(0.050)	(0.045)	(0.054)	(0.157)
VC×MoA: R&D Big inc		-0.382***				-0.302**
		(0.145)				(0.153)
VC×MoA: Small/Big non-inc R&D		0.276				0.221
		(0.190)				(0.207)
VC×MoA launched			-0.205**			-0.164
			(0.101)			(0.104)
VC×Cites big inc				-0.029***		-0.029***
				(0.008)		(0.009)
VC×Cites small or big non inc				0.009		0.010
				(0.007)		(0.008)
VC×NTO					0.150*	0.197**
					(0.089)	(0.090)
MoA: R&D Big inc	-0.006	0.057	-0.006	0.002	0.003	0.061
	(0.068)	(0.070)	(0.068)	(0.069)	(0.068)	(0.070)
MoA: Small/Big non-inc R&D	0.001	-0.034	0.002	0.000	-0.003	-0.032
	(0.074)	(0.081)	(0.075)	(0.075)	(0.074)	(0.080)
MoA launched	0.065	0.064	0.109	0.075	0.063	0.107
	(0.076)	(0.077)	(0.079)	(0.076)	(0.075)	(0.079)
Cites big inc				0.014*		0.014*
				(0.008)		(0.008)
Cites small or big non inc				-0.009		-0.009*
				(0.006)		(0.006)
NTO	-0.019	-0.018	-0.014	-0.018	-0.051	-0.056
	(0.044)	(0.044)	(0.043)	(0.044)	(0.048)	(0.048)
Contols	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
IPC	Yes	Yes	Yes	Yes	Yes	Yes
Patent nb.	Yes	Yes	Yes	Yes	Yes	Yes
Obs	783	783	783	783	783	783
Adj. R2	0.163	0.167	0.167	0.169	0.165	0.178

Note: OLS regression. These regressions estimate the effects of VC backing (column 1) and VC backing in combination with project characteristics (columns 2-6) using the sample of preclinical projects of startups. The dependent variable is a binary indicator for progression from the preclinical phase to clinical testing. To quantify citation links, column (4) considers whether any projects' assigned patents refer to any patent in the respective competitor's antidiabetics patent portfolio. Control variables include the age of a firm, the number of projects in diabetes and pharmaceutical industry, indicator variables for projects based on new chemical compounds, targeting both types of diabetes, and being a front runner. Coefficients for those are not reported. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

breakthrough projects are pursued instead. The pattern of VC's activity suggests a VC-induced avoidance of the "defeat zone."

Defeat zone and firm size. Given that VCs observe project characteristics before they decide to back them and these characteristics do not drive the investment decisions (Table 4.2), one question arises: why would a VC halt the progression of an asset it has previously invested in? If VCs are rational, the above "defeat zone" hypothesis and the underlying redirection of R&D should be driven by sufficiently large startups where there is enough room for redirection in the first place. In contrast, no evidence of the "defeat zone" should be found in smaller firms where R&D cannot be redirected and the only option for the investors is to pursue the existing R&D that these firms have. To test this, table 4.4 splits the sample into two sub-samples, depending on the size of the startups. A cut-off of having 4 projects in the pharma industry was selected to define a startup as large. Given the limited number of treated observations, this definition helps to split the sample into two similarly sized groups and, at the same time, proxy sufficiently large startups.

Indeed, for large startups (column 2, Table 4.4), the above patterns are amplified and, in fact, drive the overall result (column 1). Absent VC backing, startups seem to push forward projects in those areas where big incumbent firms operate R&D compared to other projects.[32] With VC backing, projects are less likely to be pursued if a big incumbent firm runs R&D in the project's relevant market, if they face a launched drug in their relevant market, or if their technology relates to areas where big incumbents already operate, compared to respective non-VC backed projects. Only VC-backed breakthrough projects are more likely to be pursued compared to non-VC-backed breakthrough projects. In contrast to this, smaller startups working on a few projects (column 3) are generally more likely to pursue all their VC-backed projects, apart from those where big/small firms without market power operated R&D.

Hence, the results in table 4.4 provide evidence consistent with the above intuition on the nature of the "defeat zone." VCs steer R&D away from the "defeat zone" if the backed startups are sufficiently large and room to redirect the R&D exists. They do not avoid the "defeat zone" and pursue all startup's projects if the size of the firm does not allow for redirection or R&D. The differences to projects that non-VC backed startups pursue further suggest that this pattern of activity is connected to the VC strategy.

These results are robust in a series of sensitivity checks. First, estimation using a probit model yields similar findings (Table 4.E.22 in Appendix 4.E).[33] Second, the statistical inference does not depend on the assumptions about the nature of the correlation in the error terms. As shown in Table 4.E.21 in Appendix 4.E, several different assumptions

[32] Moraga-González & Motchenkova (2021) theoretically document the incentives of startups to strategically direct their portfolios to increase acquisition rents. This may result in an alignment or a misalignment of the direction of innovation relative to what is socially optimal.

[33] However, as explained in section 3.4, these results are not further interpreted as interaction terms in non-linear models are difficult to interpret.

Table 4.4: Effects of VC - all characteristics and heterogeneity by size

	All together (1)	$\geq 4proj$ (2)	< 4 proj (3)
VC	0.011	-0.006	1.114***
	(0.157)	(0.170)	(0.412)
VC×MoA: R&D Big inc	-0.302**	-0.408*	-0.321
	(0.153)	(0.218)	(0.325)
VC×MoA: Small/Big non-inc R&D	0.221	0.346	-0.846**
	(0.207)	(0.242)	(0.329)
VC×MoA launched	-0.164	-0.243*	-0.394*
	(0.104)	(0.132)	(0.207)
VC×NTO	0.197**	0.225**	0.268
	(0.090)	(0.113)	(0.170)
VC×Cites big inc	-0.029***	-0.029**	-0.018
	(0.009)	(0.012)	(0.044)
VC×Cites small or big non inc	0.010	0.009	0.014
	(0.008)	(0.025)	(0.025)
MoA: R&D Big inc	0.061	0.227**	-0.142
	(0.070)	(0.096)	(0.117)
MoA: Small/Big non-inc R&D	-0.032	-0.090	-0.080
	(0.080)	(0.118)	(0.148)
MoA launched	0.107	0.023	0.262*
	(0.079)	(0.107)	(0.143)
NTO	-0.056	-0.015	-0.152
	(0.048)	(0.062)	(0.093)
Cites big inc	0.014*	0.018	-0.005
	(0.008)	(0.011)	(0.024)
Cites small or big non inc	-0.009*	-0.011	-0.019
	(0.006)	(0.007)	(0.024)
Contols	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
IPC	Yes	Yes	Yes
Patent nb.	Yes	Yes	Yes
Obs	783	479	304
Adj. R2	0.178	0.128	0.235

Note: OLS regression. These regressions estimate the effects of VC backing in combination with project characteristics when controlling jointly for all characteristics. Column (1) presents results for the sample of preclinical projects of startups. Columns (2) and (3) analyze sub-samples of large and small startups as determined by the number of pharmaceutical projects in their portfolio. To quantify citation links, the regressions consider whether any projects' assigned patents referred to any patent in the respective competitor's antidiabetics patent portfolio. The dependent variable is a binary indicator for progression from the preclinical phase to clinical testing. Control variables include the age of a firm, number of projects in diabetes and pharmaceutical industry, indicator variables for projects based on new chemical compounds, targeting both types of diabetes, and being a front runner. Coefficients for those are not reported. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

about the possible correlation structured were examined which potentially relate to the VC backing decisions. The inference based on the default robust standard errors (assuming no correlation) is almost identical to the inference with standard errors clustered at the firm level (allowing for correlation in the error terms of projects of the same firm), standard errors clustered at the country level (allowing for correlation in error terms of projects developed by firms in the same country) as well as standard errors clustered at the country-cohort level (allowing for correlation in error terms of projects initiated in the same year and developed by firms in the same country).

4.5.3 VC effects: Extensions

Breakthrough and competition in product markets. The table 4.5 shows that the progression rate of VC-backed breakthrough projects varies with the presence of product market competition in the project's relevant market. In particular, the underlying breakthrough effect for VC-backed projects is entirely driven by those that face no product market competition. In contrast, breakthrough projects facing existing product competition are not different in terms of progress from non-breakthrough projects and non-VC-backed projects. Columns (2) and (3) show that, in line with the previous results, such projects are more likely to be pursued by both large and small startups. This further reinforces the main results and demonstrates that VC firms actively pursue the development of projects with breakthrough potential in markets with no competition, it is in markets that can be monopolized and where the expected payoff is the highest.

Narrower technology links. Columns (1) and (2) in Table 4.6 present evidence that the VCs avoid the technology "defeat zone" even more when the patent links only refer to technology areas common between the project and the competitor's patent portfolio. Column (1) counts all citations to the antidiabetic patent portfolio of competitors and closely resembles the findings in Table 4.4, indicating that VC-backed projects in the technology areas where big incumbents operate are less likely to progress. In contrast, column (2) counts only references to IPC groups shared between the projects and the competitors' patent portfolios. Thus, these citations refer to a much narrower, directly overlapping technology space. In these cases, one patent citation to a big incumbent's portfolio generates stronger negative effects and the difference compared to the non-VC-backed projects is also amplified. In addition, we also see that patent citation to firms without market power results in a lower progression rate, except for VC-backed projects where it results in a higher progression rate. Therefore, this finding provides further support that VCs avoid competition with big incumbents in the technology space, corroborating the "defeat zone" in the technology space.[34]

Technology links and the extent of market power. Columns (3) and (4) in

[34] Appendix 4.A provides several robustness checks. Table 4.E.23 estimates the effects for small and large startups. Table 4.E.24 confirms that the results are not driven by outliers.

Table 4.5: Effects of VC - new-to-be markets & breakthrough

	All (1)	$\geq 4proj$ (2)	< 4 proj (3)
VC	-0.074	-0.060	-0.071
	(0.059)	(0.077)	(0.105)
VC×NTO in not launched MoA	0.261***	0.259*	0.306*
	(0.100)	(0.137)	(0.176)
VC×NTO in launched MoA	-0.145	-0.140	-0.223
	(0.145)	(0.181)	(0.245)
VC×non NTO in launched MoA	-0.053	-0.168	-0.194
	(0.125)	(0.135)	(0.242)
NTO in not launched MoA	-0.055	0.019	-0.150
	(0.053)	(0.067)	(0.101)
NTO in launched MoA	0.071	-0.030	0.153
	(0.103)	(0.142)	(0.186)
non NTO in launched MoA	0.106	0.045	0.258
	(0.085)	(0.110)	(0.158)
Contols	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
IPC	Yes	Yes	Yes
Patent nb.	Yes	Yes	Yes
Obs	783	479	304
Adj. R2	0.172	0.122	0.226

Note: OLS regression. These regressions estimate the effects of VC backing depending on the combination of two project characteristics: product market competition (*MoA: launched* indicator) and breakthrough nature (*NTO* indicator). Column (1) presents results for the sample of preclinical projects of startups. Columns (2) and (3) analyze sub-samples of large and small firms as determined by the number of pharmaceutical projects in their portfolio. The dependent variable is a binary indicator for progression from the preclinical phase to clinical testing. Control variables include the age of a firm, the number of projects in diabetes and pharmaceutical industry, indicator variables for projects based on new chemical compounds, targeting both types of diabetes, and being a front runner. In addition, the regressions also control for the presence of R&D incumbents and small/big non-incumbent firms in an MoA. Coefficients for those are not reported. The baseline group are non NTO projects in not launched MoAs. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 4.6 find that the lower likelihood of VCs to pursue projects in the technology spaces with the presence of big incumbents varies with the extent of incumbents' market power. More specifically, the results show that for both the broader (column 3) and narrower definitions of technology links (column 4), the negative effects on the progression rates of the VC-backed projects in the technology space around big incumbents are driven by the leader firms - i.e. those firms with the largest market power. The progression rate for VC-backed projects citing leaders is even lower compared to incumbents as such and the results are also statistically much more significant. Thus, these results show that the VC's motivation to avoid the "defeat zone" in the technology space where it is not worth competing is connected to the market power of the big incumbent firms.[35]

Terminations. The main outcome variable in this paper - progression - measures whether (VC-backed) startups actively develop preclinical projects by starting clinical testing in humans. This section investigates the opposite angle and explores whether (VC-backed) startups terminate the development of projects they do not wish to pursue. To do so, this section employs an alternative outcome variable - *termination*. Like Chapter 3, termination is defined as a situation in which a preclinical project has not proceeded to the clinical testing and, at the same time, remained in the preclinical phase for an exceptionally long period, which means beyond the mean plus one standard deviation of the projects' preclinical development duration.

The results presented in Table 4.7 reinforce the main findings. Column (2) shows that, in sufficiently large startups, VC-backed projects are generally more likely to be terminated. An exception are projects outside of the "defeat zone" around big incumbents, ie. breakthrough projects in new-to-exist-markets and projects in the R&D areas where they only face competitors without market power. In contrast, for small firms where there is little room to steer R&D, terminations of VC-backed projects are generally not more likely. We only see evidence of lower termination rates for breakthrough projects in new-to-be markets, which are projects that these companies are also more likely to pursue. Combined, this analysis provides further evidence of the "defeat zone" in the product, R&D, and technology spaces.

Volume of funding. This section shows in Table 4.8 that the previous findings also extend beyond VC's yes/no backing decision. More specifically, using variation in the amounts that the VCs provided to ventures, this section shows that the strength of the reaction increases with the rising funding volume. This pattern implies that the larger the VC's investment stake (and the extent of control rights associated with it), the stronger the "defeat zone" effect. Thus, it is not only that VCs avoid "defeat zone", they avoid it more the larger their stakes are.

[35] Table 4.E.24 confirms that the results are not driven by outliers.

Table 4.6: Effects of VC - narrower technology links and incumbent's market power

	(1) All IPC	(2) Overlapping IPC	(3) All IPC	(4) Overlapping IPC
VC	-0.008	-0.012	-0.015	-0.014
	(0.045)	(0.045)	(0.045)	(0.045)
VC×Cites big inc	-0.029***	-0.075***		
	(0.008)	(0.029)		
VC×Cites big inc - leader			-0.076***	-0.134***
			(0.029)	(0.032)
VC×Cites big inc - other			0.011	0.008
			(0.019)	(0.024)
VC×Cites small or big non inc	0.009	0.028*	0.012**	0.021**
	(0.007)	(0.016)	(0.006)	(0.010)
Cites big inc	0.014*	0.029***		
	(0.008)	(0.011)		
Cites big inc - leader			-0.003	0.012
			(0.017)	(0.021)
Cites big inc - other			0.018**	0.029**
			(0.008)	(0.012)
Cites small or big non inc	-0.009	-0.022***	-0.011**	-0.021***
	(0.006)	(0.008)	(0.005)	(0.008)
Contols	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
IPC	Yes	Yes	Yes	Yes
Patent nb.	Yes	Yes	Yes	Yes
Obs	783	783	783	783
Adj. R2	0.169	0.167	0.172	0.171

Note: OLS regression. These regressions estimate the effects of VC backing based on whether projects cite competitors' antidiabetic patents. The regressions distinguish between broader and narrower technology links and between different levels of the incumbent's market power. Column (1) and column (3) consider whether any projects' assigned patents referred to any patent in the competitors' antidiabetics patent portfolio. Column (2) and column (4) narrow the definition and consider whether projects' assigned patents referred to patents in competitors' antidiabetics patent portfolio within overlapping IPC groups. To differentiate between the levels of market power that the incumbents hold, big incumbent firms are split into two categories, *big incumbent leaders* and *other big incumbents*. The estimation sample encompasses all preclinical projects of startups. The dependent variable is a binary indicator for progression from the preclinical phase to clinical testing. Control variables include the age of a firm, the number of projects in diabetes and pharmaceutical industry, indicator variables for projects based on new chemical compounds, targeting both types of diabetes, and being a front runner. In addition, the regressions also control for the presence of R&D incumbents and small/big non-incumbent firms in an MoA, a presence of a launched project in the MoA, and the breakthrough nature of a project. Coefficients for those are not reported. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 4.7: Effects of VC - Terminations

	(1) All together	(2) $\geq 4proj$	(3) < 4 proj
VC	0.235**	0.283**	0.201
	(0.116)	(0.135)	(0.197)
VC×MoA: R&D Big inc	0.048	0.188	-0.135
	(0.120)	(0.127)	(0.122)
VC×MoA: Small/Big non-inc R&D	-0.110	-0.338**	0.080
	(0.157)	(0.153)	(0.171)
VC×NTO in not launched MoA	-0.220***	-0.215**	-0.204**
	(0.067)	(0.098)	(0.097)
VC×Cites big inc or leader	0.013	0.040	0.042
	(0.036)	(0.040)	(0.042)
VC×Cites small or big non inc	-0.015	0.064	-0.050
	(0.042)	(0.048)	(0.046)
MoA: R&D Big inc	0.048	0.010	0.093*
	(0.053)	(0.075)	(0.054)
MoA: Small/Big non-inc R&D	-0.086	-0.024	-0.050
	(0.059)	(0.088)	(0.077)
NTO in not launched MoA	0.011	-0.037	0.029
	(0.035)	(0.057)	(0.049)
Cites big inc or leader	0.005	0.020	0.002
	(0.013)	(0.020)	(0.018)
Cites small or big non inc	0.007	-0.009	-0.000
	(0.016)	(0.021)	(0.028)
Contols	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
IPC	Yes	Yes	Yes
Patent nb.	Yes	Yes	Yes
Obs	783	479	304
Adj. R2	0.151	0.183	0.198

Note: OLS regression. These regressions estimate the effects of VC backing in combination with project characteristics. Column (1) presents results for the sample of preclinical projects of startups. Columns (2) and (3) analyze sub-samples of large and small firms as determined by the number of pharmaceutical projects in their portfolio. To quantify citation links, the regressions consider whether any projects' assigned patents referred to any patent in the respective competitor's antidiabetics patent portfolio. The dependent variable is a binary indicator for termination. A project is terminated if it has not progressed from preclinical to clinical testing and at the same time stayed in the preclinical exceptionally long - it is beyond the mean plus one standard deviation of the phase duration. Control variables include the age of a firm, the number of projects in diabetes and pharmaceutical industry, indicator variables for projects based on new chemical compounds, targeting both types of diabetes, and being a front runner. Coefficients for those are not reported. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 4.8: Effects of VC funding volume - all characteristics and heterogeneity by size

	All together (1)	$\geq 4proj$ (2)	< 4 proj (3)
ln(VC amount)	-0.001	-0.007	0.071**
	(0.011)	(0.012)	(0.029)
ln(VC amount)×MoA: R&D Big inc	-0.021**	-0.029*	-0.015
	(0.010)	(0.016)	(0.021)
ln(VC amount)×MoA: Small/Big non-inc R&D	0.015	0.031*	-0.063***
	(0.015)	(0.017)	(0.024)
ln(VC amount)×MoA launched	-0.010	-0.016*	-0.025
	(0.008)	(0.009)	(0.015)
ln(VC amount)×NTO	0.016***	0.018**	0.025**
	(0.006)	(0.008)	(0.011)
ln(VC amount)×Cites big inc	-0.002***	-0.002**	-0.002
	(0.001)	(0.001)	(0.003)
ln(VC amount)×Cites small or big non inc	0.001	0.001	0.001
	(0.001)	(0.002)	(0.002)
MoA: R&D Big inc	0.059	0.223**	-0.156
	(0.070)	(0.097)	(0.118)
MoA: Small/Big non-inc R&D	-0.039	-0.117	-0.074
	(0.081)	(0.120)	(0.148)
MoA launched	0.107	0.019	0.264*
	(0.079)	(0.107)	(0.143)
NTO	-0.063	-0.019	-0.167*
	(0.048)	(0.063)	(0.090)
Cites big inc	0.014*	0.017	-0.005
	(0.008)	(0.011)	(0.024)
Cites small or big non inc	-0.009*	-0.011	-0.018
	(0.006)	(0.007)	(0.023)
Contols	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
IPC	Yes	Yes	Yes
Patent nb.	Yes	Yes	Yes
Obs	783	479	304
Adj. R2	0.179	0.129	0.238

Note: OLS regression. These regressions estimate the effects of the volume of VC backing in combination with project characteristics. The dependent variable is a binary indicator for progression from the preclinical phase to clinical testing. The treatment variable is the funding volume. Funding volume equals zero if the project did not receive any funding. For funded projects, it equals the natural logarithm of the funding amount received. Column (1) presents results for the sample of preclinical projects of startups. Columns (2) and (3) analyze sub-samples of large and small startups as determined by the number of pharmaceutical projects in their portfolio. To quantify citation links, the regressions consider whether any projects' assigned patents referred to any patent in the respective competitor's antidiabetics patent portfolio. Control variables include the age of a firm, the number of projects in diabetes and pharmaceutical industry, indicator variables for projects based on new chemical compounds, targeting both types of diabetes, and being a front runner. Coefficients for those are not reported. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

4.5.4 VC rounds and funding volume

Unlike the previous section, in this analysis, I restrict the attention to VC-backed projects only. This allows studying whether the extent of VC activity - measured by the number of VC rounds and the volume of VC funding - serves as a channel for the steering of R&D.

In Table 4.9, the dependent variables differ from the binary progression indicators employed so far and correspond instead to the number of VC rounds (column 1), the cumulative amount of money that the project's developer received (column 2), or the cumulative amount of money that projects itself received (column 3).[36] The explanatory variables correspond to the standard set of controls, except for the exclusion of the IPC fixed effects, which I cannot control for due to the sample size issues, and the inclusion of a control variable for the highest achieved phase, controlling for the different development timelines.

Table 4.9: VC rounds and funding

		Lifetime	
	(1)	(2)	(3)
	# VC rounds	$ raised	$\frac{\$raised}{\#projects}$
MoA: R&D Big inc	-0.420**	-0.976***	-1.016***
	(0.183)	(0.341)	(0.353)
MoA: Small/Big non-inc R&D	0.788***	1.486***	1.637***
	(0.202)	(0.462)	(0.464)
NTO in not launched MoA	-0.184	-0.140	-0.173
	(0.147)	(0.292)	(0.297)
Cites big inc	-0.009*	-0.006	-0.009
	(0.005)	(0.010)	(0.010)
Cites small or big non inc	0.003	-0.013	-0.012
	(0.005)	(0.018)	(0.017)
Contols	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
Patent nb.	Yes	Yes	Yes
Highest phase	Yes	Yes	Yes
Obs	151	143	143
Pseudo R2	0.121		
Adjusted R2		0.385	0.351

Note: These regressions use only the sample of VC-backed preclinical projects of startups and estimate how various project characteristics affect the following outcomes. In column (1), the dependent variable is the number of funding rounds. Since this is a count variable, Poisson regression is used. In column (2), the dependent variable is the logarithm of the cumulative amount of dollars raised per project's owner. In column (3), the dependent variable is the logarithm of the cumulative amount of dollars raised per project. The outcomes are counted over the entire project's lifetime (with a variable controlling for the highest achieved phase). Control variables include the age of a firm, the number of projects in diabetes and pharmaceutical industry, indicator variables for projects based on new chemical compounds, targeting both types of diabetes, and being a front runner. Coefficients for those are not reported. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

[36] Direct allocation of money within a startup cannot be observed. To proxy this, I allocate the amount to each project by dividing the total amount of money the startup has raised by the number of pharmaceutical projects (antidiabetic ones and others) that it had in the portfolio at the time of the transaction.

The results show that, on average, startups developing projects citing patents of the big incumbents or in those R&D areas with the presence of big incumbents are likely to receive fewer funding rounds (Column 1). In the latter case, this relationship also extends to lower average cumulative funding raised by the startups (Column 2) and lower average funding per project (Column 3). In contrast, projects outside of the "defeat zone" are not less likely to receive fewer funding rounds and, on average, also raise the same or higher average amounts of money compared to the VC-backed projects in the "defeat zone." In addition, given that the average founding amount received by a company in this sample amounts to $2.2 million, and the development cost for a project's first phase of clinical testing amount to $1.4 million on average (Office of The Assistant Secretary for Planning and Evaluation of the US Department of Health, 2014), the cost for phase I represent a substantial part of the received amounts. From the perspective of a profit-seeking VC fund, it is optimal to concentrate efforts on projects with the highest expected return. These findings jointly indicate that VC funding is a relevant channel through which the VCs steer R&D away from the "defeat zone."

4.5.5 VCs and exit decisions for the breakthrough projects

Since the breakthrough projects outside the "defeat zone" progress further and are less likely to be terminated following a VC investment, this section examines how the venture capital funds exit from these investments and realize gains. A startup can be either sold to another company in an acquisition or the startup could become publicly traded on a stock exchange through an initial public offering (IPO), allowing insiders to sell their shares.[37]

Table 4.10: Exits: Number of observations

	IPO	Acquired	Acquired - no big incumbent
All	68	73	61
VC	15	13	10
NTO	25	45	38
VC & NTO	6	8	5

Note: The table shows the number of preclinical projects developed by startups subject to experiencing an exit via IPO (column 1), M&A (column 2), or M&A where the acquirer was other than a big incumbent firm (column 3). The first row gives the total number of projects scoring on either outcome. The second row gives the number of VC-backed projects scoring on either outcome. The third row gives the number of NTO projects scoring on either outcome. The fourth row gives the number of VC-backed NTO projects scoring on either outcome.

Table 4.10 first tabulates the number of exits in the sample, whether through IPO (column 1), M&A (column 2), or M&A conditional on the acquirer being another firm than a big incumbent firm (Column 3). Overall, 73 preclinical projects of startups were acquired

[37] The exit strategy and its interactions with VC funding are thoroughly studied in the literature. Examples include Ragozzino & Blevins (2016); Amor & Kooli (2020); Giot & Schwienbacher (2007); Behnke & Hültenschmidt (2007).

at some point in their lifetime, whereas 68 were subject to an IPO. Further, 13 acquired projects and 15 IPO projects were VC-backed. A total of 45 acquired projects were VC-backed and 25 IPO projects were VC-backed. Combining the VC and breakthrough status, this number reduces to 8 projects (M&As) and 6 projects (IPO). The numbers of projects acquired by other than big-incumbent firms largely follow the total number of acquired projects, indicating a very limited role of big incumbent acquirers in these transactions in the first place.

The econometric results in Table 4.11 show that breakthrough projects are generally more likely to be acquired; this is true independently of whether the project was VC-backed or not. In contrast, IPOs are not an attractive exit option if a breakthrough project is involved. Interestingly, the M&A results are fully driven by acquisitions by other than big incumbents firms.[38] Therefore, these results indicate that breakthrough projects are generally attractive for acquisitions, particularly by firms that are not big incumbents, securing a M&A exit option for the VC's investments.[39]

Table 4.11: Exits: M&As and IPOs

	(1) IPOs	(2) M&A	(3) M&A - no big inc
VC	0.022	-0.011	0.004
	(0.035)	(0.033)	(0.032)
NTO	0.000	0.075**	0.081***
	(0.028)	(0.032)	(0.031)
VC×NTO	-0.006	-0.009	-0.059
	(0.053)	(0.056)	(0.052)
Contols	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
IPC	Yes	Yes	Yes
Patent nb.	Yes	Yes	Yes
Obs	783	783	783
Adj. R2	0.081	0.160	0.102

Note: OLS regression. This regression uses the sample of all preclinical projects of the startups and estimates how VC funds exit investments from their NTO projects. In column (1) the dependent variable equals one if the project was subject to an IPO during its lifetime. In column (2), the dependent variable equals one if the project was subject to a M&A deal. In column (3), the dependent variable equals one if the project was subject to a M&A deal and at the same time the acquirer was other than a big incumbent firm. Control variables include the age of a firm, the number of projects in diabetes and pharmaceutical industry, indicator variables for projects based on new chemical compounds, targeting both types of diabetes, and being a front runner. Coefficients for those are not reported. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

[38] An additional unreported descriptive analysis shows that these results also hold for the breakthrough projects developed in the areas with no launched projects.

[39] These findings are fully in line with Malek *et al.* (2021), who show that preclinical breakthrough projects developed by mature startups are particularly attractive targets in M&A deals and these projects are generally less likely to be acquired by big firms. In addition, several papers theoretically point out that VCs pursue R&D in areas that are particularly interesting for acquisitions (Lemley & McCreary, 2021; Norbäck & Persson, 2009; Moraga-González & Motchenkova, 2021).

4.5.6 Checking for endogeneity

Although the key explanatory variables are not driving the selection into the VC treatment, there might still be unobserved factors that simultaneously drive the probability of being VC-backed and the probability of progressing, thereby confounding the estimates. In line with the previous research (see section 4.5.1), an obvious candidate is the unobserved quality or experience of the founding team. All else equal, more capable teams might be able to attract investment and, at the same time, increase the odds of the project proceeding to the next development phase.

To check the robustness of the baseline results against endogeneity, this section employs an instrumental variable approach. The geographic distance between the VC's headquarters and the headquarters of the startup in charge of project's development serves as an instrumental variable.[40] This instrument relies on the local nature of VC investing, as demonstrated in several papers.[41] The main identification assumption is that geographic proximity between the VC fund and the firm is not correlated with the error term of the outcome regression and affects our dependent variable - progression - only through the presence of a VC fund. In particular, the papers of Bernstein *et al.* (2016) and Kang *et al.* (2019) provide evidence that the channel through which VC produces its effect relates to the geographical distance.[42]

At the project level, the distance between a VC and a startup is undefined in cases where the project never received VC backing. Therefore, to implement the IV approach, the dataset was transformed into a pair-wise structure. For every VC transaction, a set of VC-project pairs is constructed, where one pair is the truly funded one (true pairs scoring one on the VC variable) and other pairs between VC fund and never treated project are the control pairs (false pairs scoring zero on the VC variable).[43] The set of control pairs for each VC-backed project includes all never treated projects that (i) did not belong to the same startup; (ii) were actively developed at the time of the VC investment; and (iii) belonged to the same cohort. To measure the distance between the VC office and the startups' headquarters for every project pair, the haversine formula shown in equation 4.3 was employed, taking into account the curvature of Earth and computing the shortest possible aerial distances:

[40] This instrument is frequently used in the VC literature and is applied in several other studies in different contexts, for example by Hsu (2006) and Li *et al.* (2020).

[41] For example, Lerner (1995) shows that VCs tend to invest in nearby companies to reduce the costs of search and monitoring. Gompers *et al.* (2020) finds that most deal flow comes from the VC's contact networks, which are often local. Hong *et al.* (2005) shows a fund manager is more likely to make an investment decision if provided with localized "word of mouth" information.

[42] Bernstein *et al.* (2016) use the exogenous introduction of new airline routes to show the easier the access of the VC to its portfolio company, the better the innovation performance and the likelihood of a successful exit. Kang *et al.* (2019) then finds that firms backed by syndicates of geographically more concentrated VCs experienced a greater likelihood of successful exits.

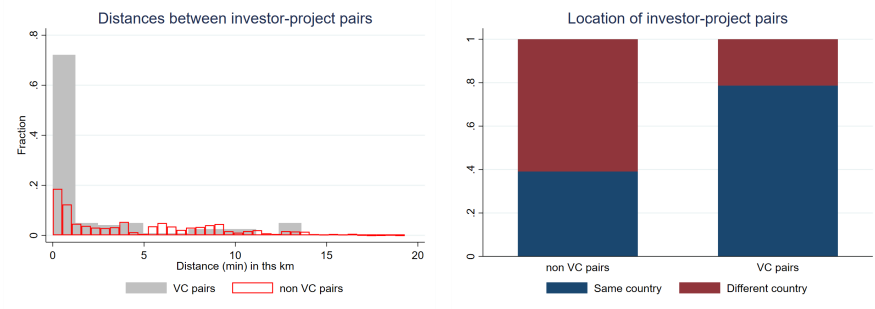
[43] This approach closely mirrors Bottazzi *et al.* (2008) and Sørensen (2007).

$$d = 2r \cdot \arcsin\left(\sqrt{\sin^2\left(\frac{\phi_j - \phi_i}{2}\right) + \cos(\phi_i)\cos(\phi_j)\sin^2\left(\frac{\lambda_j - \lambda_i}{2}\right)}\right). \tag{4.3}$$

ϕ represents the latitude, λ represents the longitude, and i, j stands for the VC headquarters and startup headquarters.

To strengthen the identification, I also employed an additional instrument - a binary variable that equals one if the VC-firm pair resided in the same country. This variable uses the global nature of this database and captures yet another dimension of the localized nature of VC investing. Figure 4.2 shows a distribution of the geographical distance and same-country residence between the VC-funded and non-funded pairs. A clear picture emerges: truly funded VC-funded pairs are much closer to each other and reside disproportionately more often in the same country. This figure provides strong support for the relevance of the instruments. Further confirmation is provided by Table 4.12, which tests whether the instruments significantly predict the VC treatment. Both instruments are statistically significant and the coefficients move in the direction expected for the local nature of VC backing: lower distance is associated with more VC backing, and residence in the same country is associated with more VC backing.

Figure 4.2: Instrument relevance - visualization



Analogous to the results in Tables 4.3, the interaction term between the VC indicator and the key explanatory variable is of main interest. However, the interaction terms also suffer from endogeneity. Therefore, I instrument them by interacting the distance with a particular explanatory variable. This produces as many instruments as there are endogenous regressors, plus the location instrument, strengthening the identification. Table 4.13 presents the results of the IV regressions. The Anderson-Rubin p-value indicates that the coefficients of the excluded instruments are jointly different from zero. Columns (1)-(4) indicate that the main results are robust when accounting for the endogeneity of the treatment and again point in the direction of avoiding a "defeat zone."

Table 4.12: Instrument relevance

	(1)	(2)	(3)
log Distance	-0.020***		-0.015***
	(0.003)		(0.003)
Same country		0.090***	0.052***
		(0.013)	(0.013)
Controls	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes
IPC FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
Matched pair FE	Yes	Yes	Yes
Obs	5011	5011	5011
Adj. R2	0.066	0.058	0.074

Note: OLS regressions. This table tests the relevance of instrumental variables. The sample used for the estimation corresponds to a pairwise dataset. For each VC transaction, a set of control VC-project pairs and the truly backed VC-project pair involved in the transaction is constructed. The dependent variable, VC backing, equals one for the truly backed VC-project pair and zero for all other control pairs. The explanatory variables include two instrumental variables: a logarithm of the distance between the VC investor and the headquarters of the firm developing the project, and an indicator variable equal to one if the VC investor and the project's developer originated from the same country. Control variables include the age of a firm, the number of projects in diabetes and pharmaceutical industry, indicator variables for projects based on new chemical compounds, targeting both types of diabetes, and being a front runner. Coefficients for those are not reported. The fixed effects include the standard set and in addition fixed effects for VC-project pairs relating to the same VC transaction. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 4.13: Instrumental variable regressions

	(1)	(2)	(3)	(4)
VC	-0.507**	0.304	-0.328*	-4.076***
	(0.228)	(0.207)	(0.193)	(0.949)
VC x NTO	0.817**			
	(0.367)			
VC x MoA launched		-3.275***		
		(0.971)		
VC x R&D MoA Inc				0.241
				(1.454)
VC x R&D MoA small/big non-inc				3.956**
				(1.778)
VC x Cites leader			-0.636*	
			(0.361)	
VC x Cites big inc			0.165	
			(0.133)	
VC x Cites small or big non inc			0.106	
			(0.070)	
NTO	-0.044*	0.000	-0.007	-0.009
	(0.024)	(0.020)	(0.017)	(0.020)
MoA launched	0.208***	0.308***	0.157***	0.209***
	(0.035)	(0.049)	(0.044)	(0.039)
Cites leader			0.039*	
			(0.021)	
Cites big inc			0.035***	
			(0.006)	
Cites small or big non inc			-0.010***	
			(0.003)	
Controls	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes	Yes
IPC FE	Yes	Yes	No	Yes
Country FE	Yes	Yes	Yes	Yes
Matched pair FE	Yes	Yes	Yes	Yes
Obs	5011	5011	5011	5011
Adj. R2	0.308	0.147	0.102	0.155
Endog. test	0.03	0.00	0.09	0.00
Cragg-Donald F	68.88	15.59	4.54	10.16
Anderson-Rubin p-value	0.00	0.00	0.01	0.00

Note: 2SLS regressions. This table reports the results of 2SLS regressions, where the endogenous treatment variable (VC) and endogenous interactions with the explanatory variables (NTO, R&D competition, MoA launched, and citations) are instrumented by distance-based instruments. The VC variable is instrumented by the logarithm of the minimal geographic distance between the HQ of the company and its VC investors. The interaction terms are instrumented by the product of the given explanatory variable and the logarithm of the minimal geographic distance. To strengthen identification, an additional instrument is added to each regression; specifically a dummy variable indicating whether the VC-firm pair was from the same country. The dependent variable is a dummy variable equal to one if a project progressed into the clinical stage. The sample used for the estimation corresponds to a pair-wise dataset. For each VC transaction, a set of control VC-project pairs and the truly backed VC-project pair involved in the transaction is constructed. Control variables include the age of a firm, the number of projects in diabetes and pharmaceutical industry, indicator variables for projects based on new chemical compounds, targeting both types of diabetes, and being a front runner. Coefficients for those are not reported. The fixed effects include the standard set and in addition fixed effects for VC-project pairs relating to the same VC transaction. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

4.6 Conclusion

This paper studies how venture capitalists affect the direction of innovation depending on product market competition and the market power of competitors they face. Using a granular dataset tracking preclinical R&D of startups at the project level, I show that VC funds actively steer the direction of the R&D of startups they have invested in. VC-backed startups are less likely to pursue and more likely to terminate the development of projects in product, R&D, and technology markets where they face the presence of big incumbent firms with market power. Instead, they pursue the development of breakthrough projects in new-to-be markets with no product market competition and sell these projects to firms that are not big incumbents. In this way, VC steers the R&D to avoid the "defeat zone" - an area where it is not worth competing with powerful incumbents.

The results of this paper provide the first detailed empirical evidence on the phenomenon of "defeat zone" and contribute to several streams of literature, particularly to the emerging literature connecting venture capital to competition and the ongoing policy and academic discussions over the related phenomenon of the "kill-zone" (Kamepalli *et al.*, 2020; Koski *et al.*, 2020; Motta & Shelegia, 2021). In terms of policy, the findings of this paper provide evidence that VC-backed innovation is unlikely to challenge big powerful firms with market power in the technology, product, and R&D spaces where they operate. VC-backed innovation seems to have limits in its ability to address the rising concentration of markets.

There is room for future work on this topic. If more comprehensive VC data were available, it would be possible to explore in more detail the mechanisms behind these actions. For example, it could be interesting to determine if specific VCs drive these patterns or whether the effects depend on the strength of governance as determined by the size of stakes that the VCs have acquired. Given the trade-off between what VC-backed companies pursue and what they do not, it would be useful to assess the actual welfare impacts of the VC's actions. A structural model would be likely needed where VC backing decisions, R&D development, and competition would be explicitly modelled and welfare under various scenarios quantified. Whilst this paper takes the decision of startups on which projects to pursue as given, a possible extension of this paper would be to investigate whether startups are likely to enter a particular development area. Finally, since this study is only limited to the diabetes market, it would be interesting to see whether these results extend also to the whole pharmaceutical industry and industries beyond pharmaceuticals.

4.A Definition of firm bins

Table 4.A.14: Bins definition

		Diabetes		
		Non-incumbent	**Incumbent**	**Leader**
Pharma	**Big**	*Big non-incumbent*	*Big incumbents*	*Leader*
	Small	*Small non-incumbents*	*Small incumbents*	-
	No marketed	*Startups*	-	-

Note: The table provides an overview of the definition of bins for the firms in the sample. Each firm belongs to a single bin, but bins can be changed over time. **Leaders** are firms with a market share in diabetes larger than 10 percent over the sample. **Big incumbents** are firms with launched dia projects and an average market share in the pharma industry above 1 percent over the sample. **Big non-incumbents** are firms without launched dia project with an average market share in the pharma industry above 0.75 percent over the sample. **Small incumbents** are all firms that are not big, are not leaders and have at least one launched project in any pharmaceutical market AND diabetes. **Small non-incumbents** are all firms that are not big, are not leaders and have at least one launched project in any pharmaceutical market BUT NOT IN diabetes. **Startups** have no launched products in any pharmaceutical market (ie only have R&D projects)

4.B Definition of variables

Table 4.B.15: Definition of variables

Variable	Type	Definition
Progression	Binary	Outcome variable; equals one if a particular project i progressed from preclinical to clinical phase and zero otherwise. Information on the initiation of Phase I clinical trials was collected from the database of clinical trials.
Termination	Binary	Outcome variable; equals one if a particular project i has not progressed and at the same time stayed in the phase beyond the mean + one standard deviation of the development time. In all other cases, it equals zero.
VC	Binary	Treatment variable; equals one if project i received its first VC backing while in preclinical phase and zero otherwise.
MoA launched	Binary	Equals one if a project i faced product market competition in its relevant market and equals zero otherwise. Product markets are defined on the basis of MoA, used by competition authorities to delineate relevant market in the antitrust proceedings.
MoA: R&D Big inc	Binary	Equals one if a big incumbent firms was working on R&D in the project's i relevant market and zero otherwise.
MoA: R&D Small/Big non-inc	Binary	Equals one if a big non-incumbent or small firm was working on R&D in the project's i relevant market and zero otherwise.
NTO	Binary	Equals one if a project i was categorized as a breakthrough project and zero otherwise. To measure breakthrough, the NTO indicator of Verhoeven *et al.* (2016) is used. A patent scores on NTO if at least one combination of IPC classes between the focal patent and its backward reference exists that has never occurred before. A project i scores on NTO if at least one project's patent scores on NTO.
Age of firm	Integer	Counts the number of semesters since the filing date of the firm's first patent.
RD projects in pharma	Integer	Number of R&D projects that the firm was working on at a particular point in time in the entire pharmaceutical industry.
RD projects in dia	Integer	Number of R&D projects that the firm was working on at a particular point in time in diabetes.
Single dia	Binary	Equals one it the firm was wokring on a development of a single project in diabetes and zero otherwise.
Single pharma	Binary	Equals one it the firm was wokring on a development of a single project in the pharmaceutical industry and zero otherwise.
New chem	Binary	Equals one if the project was based on a chemical entity not yet approved by the FDA to treat diseases (New chemical entity) and zero otherwise.
Dia 1 and 2	Binary	Equals one if the project's development targeted both types of diabetes and zero otherwise.
Front runner	Binary	Equals one for the most advanced project in a particular MoA at given time and zero otherwise.
Patent nb.	Integer	Number of patents assigned to a project at a particular point in time. This variable is a result of the patent matching procedure developed in this thesis. Every projects is matched to patent(s) that (i) belonged to the firm(s) developing it, (ii) were filed within project's development window, and (iii) related to project's chemical compounds, genetic sequences of therapeutic proteins, or mechanism of action of the project.
Cites big inc or leader	Integer	Counts the number of times a project cites patents belonging to the firms categorized as leaders or big incumbents.
Cites big inc - leader	Integer	Counts the number of times a project cites patents belonging to the firms categorized as leaders.
Cites big inc - other	Integer	Counts the number of times a project cites patents belonging to the firms categorized as big incumbents.
Cites small or big non-incumbent	Integer	Counts the number of times a project cites patents belonging to the firms categorized as big non-incumbents or small firms.

4.C Overview of relevant product markets by MoA

Table 4.C.16: Categorization of launched drugs by MoA

Market	**MoA**	**Drugs examples (and approval date by FDA)**
Rapid and medium acting insulins	Insulin	Humulin (1982q4), Novolin (1991q2), Humalog (1996q2), Novolog (2000q2), Apidra (2004q2), Afrezza (2014q2)
Long acting insulins	Insulin	Lantus (2000q2), Levemir (2005q2), Toujeo (2015q1), Tresiba (2015q3), Basaglar (2015q4), Xultophy (2016q4), Soliqua (2016q4)
Insulin sensitizers	AMPK inhibitor	Glucophage (1995q1), Riomet (2003q3), Fortamet (2004q2), Glumetza, (2005q2), Actoplus Met (2005q3)
Thiazolidinediones	PPAR-gamma agonist	Avandia (1999q2), Actos (1999q3), Avandamet (2002q4), Avandaryl (2005q4), Duetact (2006q3)
Sulfonylureas + meglitinides	K-ATP channel antagonist	Glucotrol (1984q2), Diabeta (1984q2), Glynase (1992q1), Glyburide Micronized (1992q2), Amaryl (1995q4), Prandin (1997q4), Glucovance (2000q3), Starlix (2000q4), Metaglip (2002q4), Prandimet (2008q2)
Incretin mimetics	GPL-1 agonist	Byetta (2005q2), Victoza (2010q1), Bydureon (2012q1), Tanzeum (2014q2), Trulicity (2014q3), Adlyxin (2016q3)
Gliptins	DDP-4 inhibitors	Januvia (2006q4), Janumet (2007q1), Onglyza (2009q3), Kombiglyze XR (2010q4), Tradjenta (2011q2), Jentadueto (2012q1), Nesina (2013q1), Oseni (2013q1), Kazano (2013q1)
Glifozins	SGTL-2 inhibitor	Invokana (2013q1), Farxiga (2014q1), Invokamet (2014q3), Jardiance (2014q3), Xigduo (2014q4), Glyxambi (2015q1), Synjardy (2015q3)
Alpha-glucosidas	Alpha-glucosidase inhibitors	Precose (1995q3), Glyset (1996q4)

4.D List of the largest investors

Table 4.D.17: List of investors

Investors	Invests alone	Nb. backed projects
TVM Capital	0	8
MPM Capital	0	8
Alta Partners	1	8
Novartis Venture Fund	0	7
Johnson & Johnson Development Corporation	0	6
Venrock	0	5
SR One	0	5
Versant Ventures	0	5
NovaQuest	1	4
Bay City Capital	0	4
Wellcome Trust	0	4
Arch Venture Partners	1	4
Sofinnova Ventures	0	4
Abingworth	1	4
New Enterprise Associates	0	4
Novo Holdings	0	3
HBM Healthcare Investments	0	3
KfW Bankengruppe	0	3
Index Ventures	1	3
Oxford Bioscience Partners	0	3
Sofinnova Partners	0	3
Tall Oaks Capital Partners	1	3
Tavistock Life Sciences	0	3
Bay City Capital	1	3
Innovations Kapital	0	3
Global Life Science Ventures	0	3
Index Ventures	0	3
Mitsubishi UFJ Capital	0	3
Arch Venture Partners	0	3
Deutsche Venture Capital	0	3

4.E Additional results

Table 4.E.18: Summary statistics - at least one patent reference

	Non VC		VC	
	(1)	(2)	(3)	(4)
	Mean	Count	Mean	Count
Cites big inc	7.12	50	10.95	19
Cites big inc - leader	4.29	28	7.27	11
Cites big inc - other	7.87	30	8.00	16
Cites small or big non inc	7.90	51	9.88	17

Note: The sum of the number of projects with non-zero citations for leaders and big incumbents exceeds the number of projects with non-zero citations for the big incumbents combined. This is the case since when separating the incumbent firms into the two subgroups, some project might refer to both groups at the same time.

Table 4.E.19: Step-wise selection equation (full specification)

	(1)	(2)	(3)	(4)	(5)	(6)
MoA launched	0.007	0.008	-0.011	0.012	0.012	0.012
	(0.037)	(0.040)	(0.040)	(0.064)	(0.063)	(0.063)
NTO	0.053	0.033	0.033	0.026	0.020	0.018
	(0.033)	(0.040)	(0.039)	(0.041)	(0.040)	(0.041)
MoA: R&D Big inc	-0.011	0.011	-0.007	-0.009	0.002	-0.001
	(0.048)	(0.051)	(0.052)	(0.059)	(0.059)	(0.060)
MoA: Small/Big non-inc R&D	0.006	-0.031	-0.030	-0.015	-0.025	-0.022
	(0.057)	(0.060)	(0.058)	(0.066)	(0.065)	(0.066)
Cites big inc - leader						0.009
						(0.013)
Cites big inc - other						0.002
						(0.006)
Cites small or big non inc						0.000
						(0.004)
Age of firm	-0.000	-0.001	-0.001	-0.002	-0.002	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Single pharma	-0.068*	-0.073*	-0.068	-0.068	-0.070*	-0.070*
	(0.039)	(0.042)	(0.042)	(0.042)	(0.041)	(0.041)
Single dia	0.005	0.017	0.032	0.039	0.045	0.048
	(0.032)	(0.033)	(0.033)	(0.034)	(0.034)	(0.035)
New chem	0.010	0.022	0.018	0.028	0.027	0.028
	(0.030)	(0.036)	(0.036)	(0.037)	(0.037)	(0.037)
Dia 1 and 2	-0.036	-0.029	-0.021	0.033	0.025	0.022
	(0.047)	(0.050)	(0.049)	(0.056)	(0.056)	(0.057)
Front runner	0.033	0.021	0.025	-0.047	-0.047	-0.047
	(0.035)	(0.038)	(0.037)	(0.058)	(0.057)	(0.057)
Patent nb.	0.001	-0.003	-0.002	-0.002	-0.002	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
IPC FE	No	Yes	Yes	Yes	Yes	Yes
Cohort FE	No	No	Yes	Yes	Yes	Yes
MoA FE	No	No	No	Yes	Yes	Yes
Country FE	No	No	No	No	Yes	Yes
Obs	783	783	783	783	783	783
Adj. R2	-0.003	-0.001	0.022	0.025	0.035	0.033

Note: OLS regression. This table estimates how project characteristics affect the likelihood to receive VC backing. The dependent variable is a dummy variable equal to one if a project was VC backed. The independent variables include various key predictors. The sample encompasses preclinical projects of startups. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 4.E.20: Step-wise selection equation (probit specification)

	(1)	(2)	(3)	(4)	(5)	(6)
VC						
MoA launched	0.023	0.050	-0.024	0.121	0.147	0.135
	(0.134)	(0.148)	(0.155)	(0.285)	(0.285)	(0.285)
NTO	0.186	0.140	0.132	0.086	0.056	0.048
	(0.115)	(0.138)	(0.140)	(0.148)	(0.148)	(0.149)
MoA: R&D Big inc	-0.036	0.055	-0.011	-0.106	-0.040	-0.050
	(0.177)	(0.197)	(0.212)	(0.257)	(0.272)	(0.272)
MoA: Small/Big non-inc R&D	0.025	-0.065	-0.044	0.143	0.099	0.105
	(0.214)	(0.231)	(0.239)	(0.295)	(0.314)	(0.315)
Cites big inc - leader						0.032
						(0.048)
Cites big inc - other						0.004
						(0.024)
Cites small or big non inc						0.001
						(0.013)
Age of firm	-0.001	-0.006	-0.007	-0.008*	-0.008*	-0.008*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Single pharma	-0.258	-0.302*	-0.323*	-0.340*	-0.330*	-0.328*
	(0.164)	(0.178)	(0.184)	(0.187)	(0.186)	(0.186)
Single dia	0.011	0.081	0.162	0.181	0.193	0.200
	(0.112)	(0.120)	(0.125)	(0.129)	(0.132)	(0.133)
New chem	0.039	0.089	0.097	0.116	0.104	0.108
	(0.109)	(0.132)	(0.137)	(0.144)	(0.145)	(0.145)
Dia 1 and 2	-0.148	-0.085	-0.038	0.207	0.160	0.154
	(0.194)	(0.196)	(0.199)	(0.219)	(0.221)	(0.221)
Front runner	0.115	0.089	0.121	-0.189	-0.173	-0.176
	(0.121)	(0.133)	(0.138)	(0.207)	(0.207)	(0.208)
Patent nb.	0.003	-0.009	-0.008	-0.008	-0.009	-0.010
	(0.008)	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)
IPC FE	No	Yes	Yes	Yes	Yes	Yes
Cohort FE	No	No	Yes	Yes	Yes	Yes
MoA FE	No	No	No	Yes	Yes	Yes
Country FE	No	No	No	No	Yes	Yes
Obs	783	783	783	756	756	756
Pseudo R2	0.011	0.076	0.113	0.134	0.152	0.153

Note: Probit regression. This table estimates how project characteristics affect the likelihood to receive VC backing. The dependent variable is a dummy variable equal to one if a project was VC backed. The independent variables include various key predictors. The sample encompasses preclinical projects of startups. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 4.E.21: Robustness: Effects of VC - all characteristics and heterogeneity by size (Clustering of errors)

	Cluster SE: Firm			Cluster SE: Country			Cluster SE: Country-Time		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All together	$\geq 4proj$	< 4 proj	All together	$\geq 4proj$	< 4 proj	All together	$\geq 4proj$	< 4 proj
VC	0.011	-0.006	1.114***	0.011	-0.006	1.114***	0.011	-0.006	1.114***
	(0.151)	(0.171)	(0.416)	(0.101)	(0.184)	(0.352)	(0.160)	(0.167)	(0.399)
VC×MoA: R&D Big inc	-0.302**	-0.408**	-0.321	-0.302**	-0.408*	-0.321	-0.302*	-0.408*	-0.321
	(0.153)	(0.206)	(0.329)	(0.137)	(0.218)	(0.313)	(0.170)	(0.245)	(0.321)
VC×MoA: Small/Big non-inc R&D	0.221	0.346	-0.846**	0.221**	0.346**	-0.846***	0.221	0.346	-0.846***
	(0.199)	(0.211)	(0.337)	(0.092)	(0.168)	(0.188)	(0.202)	(0.238)	(0.300)
VC×MoA launched	-0.164	-0.243*	-0.394*	-0.164**	-0.243**	-0.394***	-0.164	-0.243*	-0.394*
	(0.107)	(0.133)	(0.208)	(0.074)	(0.105)	(0.077)	(0.122)	(0.146)	(0.212)
VC×NTO	0.197**	0.225**	0.268	0.197*	0.225	0.268**	0.197**	0.225*	0.268*
	(0.085)	(0.106)	(0.169)	(0.098)	(0.134)	(0.111)	(0.081)	(0.114)	(0.145)
VC×Cites big inc	-0.029***	-0.029**	-0.018	-0.029***	-0.029**	-0.018	-0.029***	-0.029**	-0.018
	(0.009)	(0.013)	(0.044)	(0.010)	(0.011)	(0.039)	(0.009)	(0.013)	(0.041)
VC×Cites small or big non inc	0.010	0.009	0.014	0.010**	0.009	0.014	0.010	0.009	0.014
	(0.008)	(0.014)	(0.027)	(0.004)	(0.014)	(0.041)	(0.008)	(0.028)	(0.024)
MoA: R&D Big inc	0.061	0.227**	-0.142	0.061	0.227***	-0.142	0.061	0.227**	-0.142
	(0.076)	(0.104)	(0.119)	(0.054)	(0.053)	(0.094)	(0.066)	(0.095)	(0.119)
MoA: Small/Big non-inc R&D	-0.032	-0.090	-0.080	-0.032	-0.090	-0.080	-0.032	-0.090	-0.080
	(0.081)	(0.118)	(0.151)	(0.040)	(0.073)	(0.121)	(0.079)	(0.122)	(0.160)
MoA launched	0.107	0.023	0.262*	0.107	0.023	0.262*	0.107	0.023	0.262*
	(0.081)	(0.105)	(0.139)	(0.084)	(0.079)	(0.128)	(0.079)	(0.102)	(0.144)
NTO	-0.056	-0.015	-0.152*	-0.056	-0.015	-0.152	-0.056	-0.015	-0.152
	(0.048)	(0.061)	(0.090)	(0.045)	(0.064)	(0.097)	(0.049)	(0.060)	(0.100)
Cites big inc	0.014*	0.018	-0.005	0.014*	0.018	-0.005	0.014*	0.018	-0.005
	(0.008)	(0.011)	(0.024)	(0.008)	(0.011)	(0.019)	(0.008)	(0.011)	(0.022)
Cites small or big non inc	-0.009*	-0.011	-0.019	-0.009***	-0.011***	-0.019	-0.009*	-0.011	-0.019
	(0.006)	(0.007)	(0.026)	(0.003)	(0.003)	(0.031)	(0.005)	(0.007)	(0.023)
Contols	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IPC	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patent nb.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	783	479	304	783	479	304	783	479	304
Adj. R2	0.178	0.128	0.235	0.178	0.128	0.235	0.178	0.128	0.235

Note: OLS regression. These regressions estimate the effects of VC backing in combination with project characteristics and allow for different assumption about the structure of the error terms. Columns (1) - (3) cluster standard errors at the firm level. Columns (4) - (6) cluster standard errors at the country level. Columns (7) - (9) cluster standard errors at the country-cohort level. For each of these groups, results are presented separately for the sample of preclinical projects of startups (columns (1), (4), and (7)), sub-sample of large startups (columns (2), (5), and (8)) and sub-sample of small startups (columns (3), (6), and (9)). The dependent variable is a binary indicator for progression from preclinical phase to clinical testing. Control variables include the age of firm, number of projects in diabetes and pharmaceutical industry, indicator variables for projects based on new chemical compounds, targeting both types of diabetes, and being a front runner. Coefficients for those are not reported. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 4.E.22: Robustness: Effects of VC - all characteristics and heterogeneity by size (Probit)

	All together (1)	$\geq 4proj$ (2)	< 4 proj (3)
Progression			
VC	0.073	-0.133	25.634***
	(0.481)	(0.575)	(4.469)
VC×MoA: R&D Big inc	-1.067**	-1.578**	-19.428***
	(0.501)	(0.776)	(4.158)
VC×MoA: Small/Big non-inc R&D	0.710	1.400	-6.823***
	(0.644)	(0.899)	(2.461)
VC×MoA launched	-0.609*	-1.452**	-1.327
	(0.332)	(0.588)	(0.971)
VC×NTO	0.748***	1.205***	1.566**
	(0.283)	(0.391)	(0.658)
VC×Cites big inc	-0.106***	-0.129**	-0.325
	(0.032)	(0.054)	(0.431)
VC×Cites small or big non inc	0.038	0.040	1.001**
	(0.024)	(0.073)	(0.465)
MoA: R&D Big inc	0.171	0.862**	-0.540
	(0.223)	(0.351)	(0.622)
MoA: Small/Big non-inc R&D	-0.099	-0.325	-0.718
	(0.266)	(0.409)	(0.853)
MoA launched	0.348	0.048	1.079*
	(0.238)	(0.383)	(0.584)
NTO	-0.175	-0.085	-0.622
	(0.155)	(0.215)	(0.393)
Cites big inc	0.045*	0.061**	0.100
	(0.025)	(0.030)	(0.134)
Cites small or big non inc	-0.028*	-0.038**	-0.993**
	(0.016)	(0.019)	(0.450)
Contols	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
IPC	Yes	Yes	Yes
Patent nb.	Yes	Yes	Yes
Obs	783	479	281
Pseudo R2	0.245	0.296	0.467

Note: Probit regression. These regressions estimate the effects of VC backing in combination with project characteristics when controlling jointly for all characteristics. Column (1) presents results for the sample of preclinical projects of startups. Columns (2) and (3) analyze sub-samples of large and small startups as determined by the number of pharmaceutical projects in their portfolio. To quantify citation links, the regressions considers whether any projects' assigned patents referred to any patent in the respective competitor's antidiabetics patent portfolio. The dependent variable is a binary indicator for progression from preclinical phase to clinical testing. Control variables include the age of firm, number of projects in diabetes and pharmaceutical industry, indicator variables for projects based on new chemical compounds, targeting both types of diabetes, and being a front runner. Coefficients for those are not reported. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 4.E.23: Robustness: Effects of VC - technology links and split by size

	All IPC			Overlapping IPC		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	$\geq$ 4 proj	$<$ 4 proj	All	$\geq$ 4 proj	$<$ 4 proj
VC	-0.008	-0.019	0.008	-0.012	-0.017	-0.010
	(0.045)	(0.059)	(0.082)	(0.045)	(0.059)	(0.083)
VC×Cites big inc	-0.029***	-0.029**	-0.025	-0.075***	-0.100***	0.057
	(0.008)	(0.011)	(0.043)	(0.029)	(0.027)	(0.061)
VC×Cites small or big non inc	0.009	0.009	0.012	0.028*	0.032	-0.012
	(0.007)	(0.025)	(0.026)	(0.016)	(0.038)	(0.036)
Cites big inc	0.014*	0.017	0.006	0.029***	0.037**	0.018
	(0.008)	(0.011)	(0.026)	(0.011)	(0.017)	(0.031)
Cites small or big non inc	-0.009	-0.011	-0.017	-0.022***	-0.027**	-0.031
	(0.006)	(0.007)	(0.024)	(0.008)	(0.011)	(0.034)
Contols	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
IPC	Yes	Yes	Yes	Yes	Yes	Yes
Patent nb.	Yes	Yes	Yes	Yes	Yes	Yes
Obs	783	479	304	783	479	304
Adj. R2	0.169	0.116	0.220	0.167	0.120	0.221

Note: OLS regression. These regressions estimate the effects of VC backing depending whether projects cite any competitor's antidiabetic patent. To quantify citation links, columns (1) - (3) consider whether any projects' assigned patents referred to any patent in the competitors antidiabetics patent portfolio. Columns (4) - (6) narrows the definition and considers whether projects' assigned patents referred to patents in competitors antidiabetics patent portfolio within overlapping IPC groups. Whilst the columns (1) and (4) present results for the sample of preclinical projects of startups, columns (2), (5), and (3), (6) analyze sub-samples of large and small startups as determined by the number of pharmaceutical projects in their portfolio, respectively. The dependent variable is a binary indicator for progression from preclinical phase to clinical testing. Control variables include the age of firm, number of projects in diabetes and pharmaceutical industry, indicator variables for projects based on new chemical compounds, targeting both types of diabetes, and being a front runner. In addition, the regressions also controls for the presence of R&D incumbents and small/big non-incumbent firms in an MoA, a presence of a launched project in the MoA and breakthrough nature of a project. Coefficients for those are not reported. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 4.E.24: Robustness: Effects of VC - narrower technology links and incumbent's market power

	(1) All IPC	(2) Overlapping IPC	(3) All IPC	(4) Overlapping IPC
VC	-0.008	-0.012	-0.015	-0.014
	(0.045)	(0.045)	(0.045)	(0.045)
VC×Cites big inc	-0.029***	-0.075***		
	(0.008)	(0.029)		
VC×Cites big inc - leader			-0.076***	-0.134***
			(0.029)	(0.032)
VC×Cites big inc - other			0.011	0.008
			(0.019)	(0.024)
VC×Cites small or big non inc	0.009	0.028*	0.012**	0.021**
	(0.007)	(0.016)	(0.006)	(0.010)
Cites big inc	0.014*	0.029***		
	(0.008)	(0.011)		
Cites big inc - leader			-0.003	0.012
			(0.017)	(0.021)
Cites big inc - other			0.018**	0.029**
			(0.008)	(0.012)
Cites small or big non inc	-0.009	-0.022***	-0.011**	-0.021***
	(0.006)	(0.008)	(0.005)	(0.008)
Contols	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
MoA FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
IPC	Yes	Yes	Yes	Yes
Patent nb.	Yes	Yes	Yes	Yes
Obs	783	783	783	783
Adj. R2	0.169	0.167	0.172	0.171

Note: OLS regression. These regressions estimate the effects of VC backing depending whether projects cite competitor's antidiabetic patents. The regressions distinguish between broader and narrower technology links and between different levels of the incumbent's market power. Column (1) and column (3) consider whether any projects' assigned patents referred to any patent in the competitors antidiabetics patent portfolio. Column (2) and column (4) narrow the definition and consider whether projects' assigned patents referred to patents in competitors antidiabetics patent portfolio within overlapping IPC groups. To limit the influence of outliers, the variables are capped at maximum of 5 citations – i.e. all values above 5 citations are set at the value of 5. To differentiate between the levels of the market power that the incumbents hold, big incumbent firms are split in two categories, *big incumbent leaders* and *other big incumbents*. The estimation sample encompasses all preclinical projects of pipeline firms. The dependent variable is a binary indicator for progression from preclinical phase to clinical testing. Control variables include the age of firm, number of projects in diabetes and pharmaceutical industry, indicator variables for projects based on new chemical compounds, targeting both types of diabetes, and being a front runner. In addition, the regressions also controls for the presence of R&D incumbents and small/big non-incumbent firms in an MoA, a presence of a launched project in the MoA and breakthrough nature of a project. Coefficients for those are not reported. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Chapter 5

General conclusions

This dissertation provides new empirical insights and deepens our understanding of how *competition* interacts with *innovation* in the pharmaceutical industry by focusing on two topics. The first two chapters study the interplay between changes in the competitive landscape (through mergers and acquisitions) and innovation. The third chapter then focuses on how the direction of innovation changes when venture capitalists invest in startups and take into consideration the competition they face from other players. Jointly, these chapters advance the literature and highlight important yet so far omitted considerations that need to be taken into account to better understand the complex and ambiguous relationship between innovation and competition. The findings also serve as a basis for concrete policy recommendations.

Main findings and implications

Chapter 2 starts by analyzing the landscape of M&A deals, happening whilst drugs are still under development. Mapping activities of all firms in antidiabetics R&D along the entire pipeline at the project level, this paper can look *inside* of the firms and study how characteristics of individual projects matter in M&A deals. This deeper dive at the project level and the focus on three key characteristics of projects along product and technology dimensions (who?, when?, what?) provide unique insights into the ownership changes that occur in the pharmaceutical markets. The majority of M&A activity takes place between small and research-focused firms and in early stages when projects are still very uncertain and far from product markets. Particularly the high-risk/high-gain projects are likely to change hands soon after their initiation and play important role in the transactions between the research-focused firms. In contrast to common narratives, incumbents with the largest market power are the least likely acquirers and rely primarily on in-house R&D. Lastly, large non-incumbents acquire fairly advanced late-stage projects from more established small companies, thus resorting to less risky acquisition strategies.

Chapter 3 utilizes these insights and investigates how acquisitions of small and research-focused firms impact innovation. Uniquely, this paper not only looks at the position of targets and acquirers in the product markets but uses the text of patents assigned to projects to measure the presence and closeness of the merging parties in the technology space. Assessing impacts on the projects of acquirers, projects of targets, and projects overall, this chapter shows that acquisitions negatively impact innovation on average. Innovation concerns are occurring primarily in early and uncertain development phases. However, there is substantial heterogeneity in the effects depending on the position of the merging parties in both product and technology markets. Most negative impacts seem connected to the lack of technological competence. There is also a small pool of transactions (approximately 4%) where synergies are realized by product market incumbents, yielding important positive impacts on innovation. This only happens when the product market incumbents are also technology incumbents, the target projects are technologically very close to the acquirer's own projects and developed in the same relevant product markets. The scope for technological synergies thus seems to dominate any market power-driven "killer" motives, at least when innovation is still very uncertain and far from product markets.

Combined, the findings fill an important gap in the literature and infuse novel elements into the empirical work in industrial organization. The chapters particularly show that the so far unexplored technological uncertainty (early stages) and technological characteristics of projects are an essential part of the changing-of-ownership stories. Market power, dominating all ongoing debates, is only a small part of what is ongoing in pharmaceutical M&As. The key to uncovering richer patterns and obtaining a much more comprehensive understanding of the implications of M&As for innovation in the pharma sector hinges on considering technological uncertainty, presence, and closeness between merging parties in both product and technology markets.

From the policy perspective, innovation concerns have been expressed very rarely in merger proceedings (Veugelers, 2012), and if so, typically only in cases where the acquiring firm is an incumbent and activities of targets and acquirers overlap in the product market space.[1] However, authorities have demonstrated substantial interest in this topic.[2] The presented empirical evidence is thus also important for policymakers and antitrust authorities, yielding several direct policy implications.

First, revenue thresholds should not limit the scope of antitrust scrutiny when inno-

[1] For example, between 2015 and 2017, European Commission received 1070 merger notifications. Innovation concerns were identified in 10 cases, usually in addition to static price concerns (Esteva Mosso, 2018, p.6).

[2] For example, in 2020, a study by Pang *et al.* (2020) was commissioned by the European Commission to assess the impacts of M&As on innovation in the pharma sector. In 2021, several competition authorities formed a working group to exchange experience and develop new approaches to deal with pharmaceutical M&As (DG Competition, 2021). The acquisition of Grail by Illumina prohibited by the European Commission is then an interesting example of the enforcement focus shifting toward target firms with no marketed products.

vation is at stake. The current merger notification criteria based on revenue thresholds miss a substantial part of important M&A activity. Firms with no launched drugs (and no revenues) or only small revenues are responsible for the majority of R&D and are not only frequent targets but also active acquirers. Since transactions between these firms often involve potentially very valuable disruptive projects, the scope of enforcement should be broadened to cover those. Second, the merger review focus should not be limited to later stages only. More than half of the deals happen in early development, and early development is where the most negative innovation effects are generally concentrated. Hence, the merger review process should be broadened in scope also in this dimension and should consistently look at projects involved in the transactions along the entire R&D pipeline. Third, the enforcement framework aimed at innovation concerns should routinely involve an assessment of technological as well as product market closeness between assets of the merging parties. This dissertation shows that technology incumbency and technological closeness of acquirer's and target's projects are at least as important as the positions of the merging parties in the product markets. It is the combination of closeness in technology *and* product markets that allows disentangling transactions with positive innovation effects from those with negative innovation effects in a specific therapeutic area.

Chapter 4 departs from the M&A topic. Rather than measuring how changes in the competitive landscape impact innovation, the primary objective of this chapter is to understand how venture capital investments shape the *direction of innovation*, depending on *competition* and the market power of the competitors they face. The chapter finds that venture capital funds actively steer the direction of early-stage innovation in their backed startups to avoid the *"defeat zone"* - an area where big incumbent firms are present and where it is not worth competing. VC-backed firms specifically do not pursue projects in product markets, R&D spaces, and technology spaces where big incumbents operate; instead, they pursue breakthrough projects in markets without product competition. Eventually, they sell these to non-incumbent firms. This behaviour is driven by sufficiently large startups, with room to steer R&D.

This paper advances the innovation literature by providing detailed evidence on how VCs influence the direction of innovation, helping to fill a gap in our understanding of what VCs do inside of firms (Gompers *et al.*, 2020). It also pioneers the empirical literature linking venture capital and innovation to competition. Also, this paper offers insights for policymakers. From the competition perspective, it shows that venture capital responds to market power and directs innovation away from it - an important consideration in the ongoing discussions on the market power of "superstar" firms. In terms of broader policy implications, recommendations on how to spur innovation, particularly in Europe, often emphasize the crucial role played by venture capital (see e.g. Popov & Roosenboom, 2009; Veugelers *et al.*, 2015; Acevedo *et al.*, 2016). The important insight is that not all early-stage innovation flourishes when backed by venture capitalists. VC funding pushes

breakthrough innovation in new markets, but not innovation in areas where big incumbent firms with market power operate. These considerations have been largely neglected in the debates but can have profound, long-term implications for welfare.

Avenues for future research

This dissertation sets the foundation for further research. Firstly, the unique algorithm developed in this thesis to match individual projects to patents offers a wide range of possibilities and applications. The rich information contained in patents can be utilized to create various project-level measures, capturing different technological aspects of projects, for example, their originality, similarity, novelty, centrality, or technological relationships between the parties involved in the deals. This can be used to shed further light on how competition interacts with innovation or employed to investigate entirely new research questions. Secondly, while the scope of current work is only limited to antidiabetics, the matching methodology as well as the framework defining positions of projects and firms in technology and product markets is universal and can be scaled to the entire industry with sufficient resources.

There are several avenues for further research in each chapter. A natural extension of chapter 2 is to investigate even deeper the transaction dynamics between the purely innovating pipeline firms, with a particular emphasis on the question of *what* is involved in these deals. The patent-based characterization of projects in the technology space offers a unique tool to do so. Thus, it would be interesting to see how close or far the acquired projects are to the technologies of their acquirers, how mainstream or niche the acquired projects are within the realm of antidiabetic R&D, or how complex and original the projects are. Since this topic remains largely unexplored, further avenues of research include also a deeper exploration of motives behind the deals, such as the removal of potential competition (consolidation), the motivation to tap and access technologically complementary resources, or the target's financial distress (Danzon *et al.*, 2007). In this respect, linking this topic to venture capital funding would be of particular relevance.

In chapter 3, the dynamics in the transactions involving terminations are specifically worth additional exploration. Given that most negative effects on innovation occur in transactions involving technology non-incumbents, it would be interesting to see whether these relate to the lack of acquirer's experience, the lack of resources, or simply stem from the fact that antidiabetics were not the area of acquirer's interest, resulting in limited willingness to incur costs and pursue antidiabetic R&D. Such analysis would require data encompassing also other therapeutic markets. Although not an easy task, another interesting angle worth further exploration relates to the types of terminations happening when acquiring firms that already worked on their own projects before buying others. Those can, for example, relate to the elimination of duplication, discontinuation of poor-quality

projects, or to the market-power-related killing of competitors, all having different implications for innovation. Posing a different but related question, further research could also consider our framework to investigate more long-term impacts on innovation, particularly whether and how M&As affect the starts of new R&D endeavours of the merging parties and/or rivals.

In chapter 4, the scope for further research starts with a more detailed exploration of the mechanism through which the VC generate the "defeat zone" effect. If more comprehensive VC data was available, it would be possible to see whether specific VCs drive these patterns. Given the trade-off between what VC-backed companies pursue and what they do not, it would be useful to assess the actual welfare impacts of the VC's actions. For this, a structural model would likely need to be developed where VC backing decisions, R&D development, and competition would be explicitly modelled and welfare under various scenarios quantified. In this respect, another interesting extension relates to an analysis of the longer-term dynamics - particularly what happens to projects in the clinical stages and how that depend on them being acquired or not. This is particularly relevant in the context of findings in other parts of this dissertation. Lastly, this paper takes the decision of startups on which projects to pursue as predetermined. A possible extension of this paper would be to thus to investigate whether VC-backed startups are likely to enter a particular development area.

References

Abadie, A., S. Athey, G. W. Imbens, & J. Wooldridge (2017): "When should you adjust standard errors for clustering?" *Technical report*, National Bureau of Economic Research.

Abrantes-Metz, R. M., C. Adams, & A. Metz (2004): "Pharmaceutical development phases: a duration analysis." *Technical report*, FTC, Bureau of Economics Working Paper.

Acevedo, M., M. Adey, C. Bruno, G. del Bufalo, A. Gazaniol, V. Lo, G. Metzger, B. Perez, D. van der Schans, & B. Thornary (2016): "France, Germany, Italy, Spain and the United Kingdom–Building Momentum in Venture Capital across Europe. Published by Bpifrance, CDP, ICO." *British Business Bank, KfW* .

Adams, C. & V. V. Brantner (2003): "New Drug Development: Estimating entry from human clinical trials." *Technical report*, FTC Bureau of Economics Working Paper.

Adams, C. P. & V. V. Brantner (2006): "Estimating the cost of new drug development: is it really $802 million?" *Health affairs* **25(2)**: pp. 420–428.

Affeldt, P. & R. Kesler (2021a): "Big Tech Acquisitions: Towards Empirical Evidence." *Journal of European Competition Law & Practice* **12(6)**: pp. 471–478.

Affeldt, P. & R. Kesler (2021b): "Competitors' Reactions to Big Tech Acquisitions: Evidence from Mobile Apps." *Technical report*, DIW Working paper.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, & P. Howitt (2005): "Competition and innovation: An inverted-U relationship." *The quarterly journal of economics* **120(2)**: pp. 701–728.

Aharonson, B. S. & M. A. Schilling (2016): "Mapping the technological landscape: Measuring technology distance, technological footprints, and technology evolution." *Research Policy* **45(1)**: pp. 81–96.

Ahuja, G. & R. Katila (2001): "Technological acquisitions and the innovation performance of acquiring firms: A longitudinal study." *Strategic management journal* **22(3)**: pp. 197–220.

Ai, C. & E. C. Norton (2003): "Interaction terms in logit and probit models." *Economics letters* **80(1)**: pp. 123–129.

Alcácer, J., M. Gittelman, & B. Sampat (2009): "Applicant and examiner citations

in us patents: An overview and analysis." *Research Policy* **38(2)**: pp. 415–427.

Amor, S. B. & M. Kooli (2020): "Do m&a exits have the same effect on venture capital reputation than ipo exits?" *Journal of Banking & Finance* **111**: p. 105704.

Andersson, M. & J. Xiao (2016): "Acquisitions of start-ups by incumbent businesses: A market selection process of "high-quality" entrants?" *Research Policy* **45(1)**: pp. 272–290.

Argente, D., S. Baslandze, S. Moreira, & D. Hanley (2019): "Patents to products: innovation, product creation, and firm growth." *Unpublished, Penn State* .

Argentesi, E., P. Buccirossi, E. Calvano, T. Duso, A. Marrazzo, & S. Nava (2021): "Merger policy in digital markets: An ex post assessment." *Journal of Competition Law & Economics* **17(1)**: pp. 95–140.

Arora, A., A. Gambardella, L. Magazzini, & F. Pammolli (2009): "A breath of fresh air? Firm type, scale, scope, and selection effects in drug development." *Management Science* **55(10)**: pp. 1638–1653.

Arroyabe, M. F. (2021): "The role of patent expiration in acquisition decision and target selection in the pharmaceutical industry." *R&D Management* **51(5)**: pp. 521–537.

Arts, S., B. Cassiman, & J. C. Gomez (2018): "Text matching to measure patent similarity." *Strategic Management Journal* **39(1)**: pp. 62–84.

Arts, S., J. Hou, & J. C. Gomez (2021): "Natural language processing to identify the creation and impact of new technologies in patent text: Code, data, and new measures." *Research Policy* **50(2)**: p. 104144.

Association, A. D. *et al.* (2019): "9. Pharmacologic approaches to glycemic treatment: Standards of Medical Care in Diabetes-2019." *Diabetes care* **42(Supplement 1)**: pp. S90–S102.

Behner, P. & P. Spence (2020): "2020 EY M&A Firepower report." *Technical report*, E&Y.

Behnke, N. & N. Hültenschmidt (2007): "New path to profits in biotech: Taking the acquisition exit." *Journal of Commercial Biotechnology* **13(2)**: pp. 78–85.

Bena, J. & K. Li (2014): "Corporate innovations and mergers and acquisitions." *The Journal of Finance* **69(5)**: pp. 1923–1960.

Berger, S. I. & R. Iyengar (2011): "Role of systems pharmacology in understanding drug adverse events." *Wiley interdisciplinary reviews: systems biology and medicine* **3(2)**: pp. 129–135.

Bernstein, S., X. Giroud, & R. R. Townsend (2016): "The impact of venture capital monitoring." *The Journal of Finance* **71(4)**: pp. 1591–1622.

Bernstein, S., A. Korteweg, & K. Laws (2017): "Attracting early-stage investors: Evidence from a randomized field experiment." *The Journal of Finance* **72(2)**: pp. 509–538.

BiopharmaDive (2017): "What does venture capital look for when investing in bio-

pharma?"

BiopharmaDive (2020): "Venture capital found its footing in biotech. then came the virus."

Blume-Kohout, M. E. & N. Sood (2013): "Market size and innovation: Effects of Medicare Part D on pharmaceutical research and development." *Journal of public economics* **97**: pp. 327–336.

Bottazzi, L., M. Da Rin, & T. Hellmann (2008): "Who are the active investors?: Evidence from venture capital." *Journal of Financial Economics* **89(3)**: pp. 488–512.

Branstetter, L., C. Chatterjee, & M. J. Higgins (2014): "Starving (or fattening) the golden goose?: Generic entry and the incentives for early-stage pharmaceutical innovation." *Technical report*, National Bureau of Economic Research.

Bryan, K. A. & E. Hovenkamp (2020): "Antitrust limits on startup acquisitions." *Review of Industrial Organization* pp. 1–22.

Caffarra, C., G. Crawford, & T. Valletti (2020): "How tech rolls: Potential competition and reverse killer acquisitions." *VOX.EU* .

Cassiman, B., M. G. Colombo, P. Garrone, & R. Veugelers (2005): "The impact of M&A on the R&D process: An empirical analysis of the role of technological-and market-relatedness." *research policy* **34(2)**: pp. 195–220.

Caudill, S. B. *et al.* (1988): "An advantage of the linear probability model over probit or logit." *Oxford Bulletin of Economics and Statistics* **50(4)**: pp. 425–427.

Chandra, A. (2020): "Innovation and Entrepreneurship in Health Care." In "The Role of Innovation and Entrepreneurship in Economic Growth," University of Chicago Press.

Chandy, R. K. & G. J. Tellis (2000): "The incumbent's curse? Incumbency, size, and radical product innovation." *Journal of marketing* **64(3)**: pp. 1–17.

Chaudhury, A., C. Duvoor, R. Dendi, V. Sena, S. Kraleti, A. Chada, R. Ravilla, A. Marco, N. S. Shekhawat, M. T. Montales *et al.* (2017): "Clinical review of antidiabetic drugs: implications for type 2 diabetes mellitus management." *Frontiers in endocrinology* **8**: p. 6.

Chemmanur, T. J., K. Krishnan, & D. K. Nandy (2011): "How does venture capital financing improve efficiency in private firms? A look beneath the surface." *The Review of Financial Studies* **24(12)**: pp. 4037–4090.

Christensen, C. M. (2013): *The innovator's dilemma: when new technologies cause great firms to fail.* Harvard Business Review Press.

Cloodt, M., J. Hagedoorn, & H. Van Kranenburg (2006): "Mergers and acquisitions: Their effect on the innovative performance of companies in high-tech industries." *Research policy* **35(5)**: pp. 642–654.

Colombo, M. G. & L. Rabbiosi (2014): "Technological similarity, post-acquisition R&D reorganization, and innovation performance in horizontal acquisitions." *Research Policy* **43(6)**: pp. 1039–1054.

Comanor, W. S. & F. M. Scherer (2013): "Mergers and innovation in the pharmaceutical industry." *Journal of health economics* **32(1)**: pp. 106–113.

Commission, E. *et al.* (1997): "Commission notice on the definition of relevant market for the purposes of community competition law." *Official Journal of the European Communities* **100**: pp. 5–13.

Cotropia, C. A., M. A. Lemley, & B. Sampat (2013): "Do applicant patent citations matter?" *Research Policy* **42(4)**: pp. 844–854.

Cunningham, C. (2017): "When does novelty pay?" *Academy of Management Proceedings* **2017(1)**: p. 16957.

Cunningham, C., F. Ederer, & S. Ma (2021): "Killer acquisitions." *Journal of Political Economy* **129(3)**: pp. 649–702.

Czarnitzki, D., K. Hussinger, & B. Leten (2020): "How valuable are patent blocking strategies?" *Review of Industrial Organization* **56(3)**: pp. 409–434.

Czarnitzki, D. & K. Kraft (2010): "Which firms buy licenses? Market positions and license expenditures." *Scandinavian Journal of Economics* **112(3)**: pp. 471–488.

Da Rin, M., T. Hellmann, & M. Puri (2013): "A survey of venture capital research." In "Handbook of the Economics of Finance," volume 2, pp. 573–648. Elsevier.

Danzon, P. M., A. Epstein, & S. Nicholson (2007): "Mergers and acquisitions in the pharmaceutical and biotech industries." *Managerial and Decision Economics* **28(4-5)**: pp. 307–328.

Deloitte (2020): "2019 Global life sciences outlook." *Technical report*, Deloitte.

Denicolò, V. & M. Polo (2018): "The innovation theory of harm: an appraisal." *Antitrust LJ* **82**: p. 921.

Desyllas, P. & A. Hughes (2010): "Do high technology acquirers become more innovative?" *Research Policy* **39(8)**: pp. 1105–1121.

DG Competition (2021): "The European Commission forms a Multilateral Working Group with leading competition authorities to exchange best practices on pharmaceutical mergers ."

Dranove, D., C. Garthwaite, & M. I. Hermosilla (2020): "Expected profits and the scientific novelty of innovation." *Technical report*, National Bureau of Economic Research.

Engel, D. & M. Keilbach (2007): "Firm-level implications of early stage venture capital investment—an empirical investigation." *Journal of Empirical Finance* **14(2)**: pp. 150–167.

Esteva Mosso, C. (2018): "Innovation in eu merger control." Remarks prepared for the 66th ABA Section of Antitrust Law Spring Meeting, Washington, 12 April.

European Commission (2021): "Commission fines Sigma-Aldrich EUR 7.5 million for providing misleading information during Merck takeover investigation." `https://ec.europa.eu/commission/presscorner/detail/en/IP_21_2181`.

Federico, G., G. Langus, & T. Valletti (2017): "A simple model of mergers and innovation." *Economics Letters* **157**: pp. 136–140.

Federico, G., G. Langus, & T. Valletti (2018): "Reprint of: Horizontal mergers and product innovation." *International Journal of Industrial Organization* **61**: pp. 590–612.

Federico, G., F. S. Morton, & C. Shapiro (2020): "Antitrust and innovation: Welcoming and protecting disruption." *Innovation Policy and the Economy* **20(1)**: pp. 125–190.

Fernald, K., R. Hoeben, & E. Claassen (2015): "Venture capitalists as gatekeepers for biotechnological innovation." *Journal of Commercial Biotechnology* **21(3)**.

Fleming, L. (2001): "Recombinant uncertainty in technological search." *Management science* **47(1)**: pp. 117–132.

Foster, J. G., A. Rzhetsky, & J. A. Evans (2015): "Tradition and innovation in scientists' research strategies." *American Sociological Review* **80(5)**: pp. 875–908.

Frey, R. & K. Hussinger (2011): "European market integration through technology-driven M&As." *Applied Economics* **43(17)**: pp. 2143–2153.

Gans, J. S. & S. Stern (2003): "The product market and the market for "ideas": commercialization strategies for technology entrepreneurs." *Research policy* **32(2)**: pp. 333–350.

Gautier, A. & J. Lamesch (2021): "Mergers in the digital economy." *Information Economics and Policy* **54**: p. 100890.

Gilbert, R. (2018): "Mergers and R&D Diversity: How Much Competition is Enough?"

Gilbert, R. J. (2019): "Competition, Mergers, and R&D Diversity." *Review of Industrial Organization* pp. 1–20.

Giot, P. & A. Schwienbacher (2007): "Ipos, trade sales and liquidations: Modelling venture capital exits using survival analysis." *Journal of Banking & Finance* **31(3)**: pp. 679–702.

Gompers, P. A., W. Gornall, S. N. Kaplan, & I. A. Strebulaev (2020): "How do venture capitalists make decisions?" *Journal of Financial Economics* **135(1)**: pp. 169–190.

Gompers, P. A. & J. Lerner (2004): *The venture capital cycle*. MIT press.

Gorman, M. & W. A. Sahlman (1989): "What do venture capitalists do?" *Journal of business venturing* **4(4)**: pp. 231–248.

Grabowski, H. & M. Kyle (2008): "11. Mergers and alliances in pharmaceuticals: effects on innovation and R&D productivity." *The economics of corporate governance and mergers* p. 262.

Grassano, N., H. Hernandez, P. Fako, A. Tuebke, S. Amoroso, A. Georgakaki, L. Napolitano, F. Pasimeni, F. Rentocchini, R. Compano *et al.* (2021): "The 2021 EU Industrial R&D Investment Scoreboard." *Technical report*, Joint Research Centre (Seville site).

Gugler, K., D. C. Mueller, B. B. Yurtoglu, & C. Zulehner (2003): "The effects of

mergers: an international comparison." *International journal of industrial organization* **21(5)**: pp. 625–653.

Hall, B. H. & J. Lerner (2010): "The financing of R&D and innovation." In "Handbook of the Economics of Innovation," volume 1, pp. 609–639. Elsevier.

Harris, R. S., T. Jenkinson, & S. N. Kaplan (2014): "Private equity performance: What do we know?" *The Journal of Finance* **69(5)**: pp. 1851–1882.

Haucap, J., A. Rasch, & J. Stiebale (2019): "How mergers affect innovation: Theory and evidence." *International Journal of Industrial Organization* **63**: pp. 283–325.

He, J. & X. Tian (2018): "Finance and corporate innovation: A survey." *Asia-Pacific Journal of Financial Studies* **47(2)**: pp. 165–212.

Hellmann, T. (1998): "The allocation of control rights in venture capital contracts." *The Rand Journal of Economics* pp. 57–76.

Hellmann, T. & M. Puri (2000): "The interaction between product market and financing strategy: The role of venture capital." *The review of financial studies* **13(4)**: pp. 959–984.

Hellmann, T. & M. Puri (2002): "Venture capital and the professionalization of start-up firms: Empirical evidence." *The journal of finance* **57(1)**: pp. 169–197.

Henkel, J., T. Rønde, & M. Wagner (2015): "And the winner is — Acquired. Entrepreneurship as a contest yielding radical innovations." *Research Policy* **44(2)**: pp. 295–310.

Higgins, M. J. & D. Rodriguez (2006): "The outsourcing of r&d through acquisitions in the pharmaceutical industry." *Journal of financial economics* **80(2)**: pp. 351–383.

Hoberg, G. & G. Phillips (2010): "Product market synergies and competition in mergers and acquisitions: A text-based analysis." *The Review of Financial Studies* **23(10)**: pp. 3773–3811.

Hong, H., J. D. Kubik, & J. C. Stein (2005): "Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers." *The Journal of Finance* **60(6)**: pp. 2801–2824.

Howell, S. T., J. Lerner, R. Nanda, & R. R. Townsend (2020): "Financial distancing: How venture capital follows the economy down and curtails innovation." *Technical report*, National Bureau of Economic Research.

Hsu, D. H. (2004): "What do entrepreneurs pay for venture capital affiliation?" *The journal of finance* **59(4)**: pp. 1805–1844.

Hsu, D. H. (2006): "Venture capitalists and cooperative start-up commercialization strategy." *Management Science* **52(2)**: pp. 204–219.

Hsu, P.-H., P. Huang, M. Humphery-Jenner, & R. Powell (2021): "Cross-border mergers and acquisitions for innovation." *Journal of International Money and Finance* **112**: p. 102320.

IFPMA (2017): "The Pharmaceutical Industry and Global Health. Facts and Figures

2017."

International Diabetes Federation (2021): "IDF diabetes atlas: Diabetes facts & figures." Available: https://idf.org/aboutdiabetes/what-is-diabetes/facts-figures.html.

Jaffe, A. B. (1986): "Technological opportunity and spillovers of r&d: evidence from firms' patents, profits and market value."

Jaffe, A. B., M. Trajtenberg, & R. Henderson (1993): "Geographic localization of knowledge spillovers as evidenced by patent citations." *the Quarterly journal of Economics* **108(3)**: pp. 577–598.

Jensen, K. & F. Murray (2005): "Intellectual property landscape of the human genome." *Science* **310(5746)**: pp. 239–240.

Jullien, B. & Y. Lefouili (2018): "Horizontal mergers and innovation." *Journal of Competition Law & Economics* **14(3)**: pp. 364–392.

Kamepalli, S. K., R. Rajan, & L. Zingales (2020): "Kill zone." *Technical report*, National Bureau of Economic Research.

Kang, J.-K., Y. Li, & S. Oh (2019): "Geographic concentration of venture capital investors, corporate monitoring, and firm performance." *Corporate Monitoring, and Firm Performance (April* **8**: p. 2019.

Kaplan, S. N. & P. E. Strömberg (2004): "Characteristics, contracts, and actions: Evidence from venture capitalist analyses." *The Journal of Finance* **59(5)**: pp. 2177–2210.

Kerr, W. R., J. Lerner, & A. Schoar (2014): "The consequences of entrepreneurial finance: Evidence from angel financings." *The Review of Financial Studies* **27(1)**: pp. 20–55.

Kortum, S. & J. Lerner (2001): *Does venture capital spur innovation?* Emerald Group Publishing Limited.

Koski, H., O. Kässi, & F. Braesemann (2020): "Killers on the road of emerging start-ups–implications for market entry and venture capital financing." *Technical report*, ETLA Working Papers.

Krieger, J., D. Li, & D. Papanikolaou (2017): "Developing novel drugs." *Available at SSRN 3095246* .

Krieger, J. L., D. Li, & D. Papanikolaou (2018): "Missing novelty in drug development." *Working Paper 24595*, National Bureau of Economic Research.

Kwon, D. & O. Sorenson (2021): "The Silicon Valley Syndrome." *Entrepreneurship Theory and Practice* p. 10422587211050892.

Kyle, M. K. (2007): "Pharmaceutical price controls and entry strategies." *The Review of Economics and Statistics* **89(1)**: pp. 88–99.

Lemley, M. A. & A. McCreary (2021): "Exit strategy." *BUL Rev.* **101**: p. 1.

Lerner, J. (1995): "Venture capitalists and the oversight of private firms." *the Journal of Finance* **50(1)**: pp. 301–318.

Lerner, J. (2009): *Boulevard of broken dreams.* Princeton University Press.

Letina, I., A. Schmutzler, R. Seibel *et al.* (2020): *Start-up acquisitions and innovation strategies.* Universität Bern.

Li, X., T. Liu, & L. A. Taylor (2020): "Do Venture Capitalists Stifle Competition?" *Available at SSRN 3479439* .

Maksimovic, V. & G. Phillips (2001): "The market for corporate assets: Who engages in mergers and asset sales and are there efficiency gains?" *The journal of finance* **56(6)**: pp. 2019–2065.

Malek, J., M. Newham, J. Seldeslachts, & R. Veugelers (2021): "Acquiring innovation: who, when, and what? Evidence from antidiabetic drug development."

Marco, A. C. & G. C. Rausser (2011): "Complementarities and spillovers in mergers: an empirical investigation using patent data." *Economics of Innovation and New Technology* **20(3)**: pp. 207–231.

Meder, H. (2015): "The dynamics of market-targeted drug development in post-ma environments." *Working paper*, Management and Strategy Institute, KU Leuven.

Meder, H. (2016): "The Incentives for M&As in the Pharmaceutical Industry." *Working paper*, Management and Strategy Institute, KU Leuven.

Moraga-González, J. L. & E. Motchenkova (2021): "How do start-up acquisitions affect the direction of innovation?" *Tinbergen Institute DP 2021-065/VII* .

Motta, M. & M. Peitz (2020): "Big tech mergers." *CEPR Discussion Paper No. DP14353* .

Motta, M. & S. Shelegia (2021): "Copying, acquiring and the "kill zone"." *Universitat Pompeu Fabra Working Paper Series* .

Motta, M. & E. Tarantino (2021): "The effect of horizontal mergers, when firms compete in prices and investments." *International Journal of Industrial Organization* **78**: p. 102774.

Mullard, A. (2014): "BMS sells diabetes business." *Nature Biotechnology* **32(3)**: pp. 209–210.

Nat. Med (2010): "Mechanism matters." *Nat. Med* **16(4)**: pp. 347–347.

Norbäck, P.-J., C. Olofsson, & L. Persson (2020): "Acquisitions for sleep." *The BE Journal of Economic Analysis & Policy* **20(2)**.

Norbäck, P.-J. & L. Persson (2009): "The organization of the innovation industry: Entrepreneurs, venture capitalists, and oligopolists." *Journal of the European Economic Association* **7(6)**: pp. 1261–1290.

NVCA (n.a.): "What is Venture Capital?" Available at`https://nvca.org/about-us/what-is-vc/#toggle-id-3`.

Office of The Assistant Secretary for Planning and Evaluation of the US Department of Health (2014): "Examination of clinical trial costs and barriers for drug development." *Technical report*, US Department of Health.

Ornaghi, C. (2009a): "Mergers and innovation in big pharma." *International journal of industrial organization* **27(1)**: pp. 70–79.

Ornaghi, C. (2009b): "Positive assortive merging." *Journal of Economics & Management Strategy* **18(2)**: pp. 323–346.

Pammolli, F., L. Magazzini, & M. Riccaboni (2011): "The productivity crisis in pharmaceutical R&D." *Nature reviews Drug discovery* **10(6)**: p. 428.

Pang, T., B. Folwell, A. Osborne, H. Tung, C. Vasiliou, N. Boliter, G. Murphy, & F. Szücs (2020): *Study on the impact of mergers and acquisitions on innovation in the pharmaceutical sector*. Publications Office of the European Union.

Paul, S. M., D. S. Mytelka, C. T. Dunwiddie, C. C. Persinger, B. H. Munos, S. R. Lindborg, & A. L. Schacht (2010): "How to improve R&D productivity: the pharmaceutical industry's grand challenge." *Nature reviews Drug discovery* **9(3)**: pp. 203–214.

Petty, J. S. & M. Gruber (2011): "In pursuit of the real deal: A longitudinal study of VC decision making." *Journal of Business Venturing* **26(2)**: pp. 172–188.

Phillips, G. M. & A. Zhdanov (2013): "R&D and the Incentives from Merger and Acquisition Activity." *The Review of Financial Studies* **26(1)**: pp. 34–78.

Popov, A. A. & P. Roosenboom (2009): "Does private equity investment spur innovation? Evidence from Europe." *ECB working paper* .

Prencipe, D. (2017): "The European venture capital landscape: an EIF perspective. Volume III: Liquidity events and returns of EIF-backed VC investments." *Technical report*, EIF Working Paper.

Puri, M. & R. Zarutskie (2012): "On the life cycle dynamics of venture-capital-and non-venture-capital-financed firms." *The Journal of Finance* **67(6)**: pp. 2247–2293.

Ragozzino, R. & D. P. Blevins (2016): "Venture–backed firms: How does venture capital involvement affect their likelihood of going public or being acquired?" *Entrepreneurship Theory and Practice* **40(5)**: pp. 991–1016.

Rainey, C. & K. McCaskey (2021): "Estimating logit models with small samples." *Political Science Research and Methods* **9(3)**: pp. 549–564.

Sahlman, W. A. (1990): "The structure and governance of venture-capital organizations." *Journal of financial economics* **27(2)**: pp. 473–521.

Samila, S. & O. Sorenson (2011): "Venture capital, entrepreneurship, and economic growth." *The Review of Economics and Statistics* **93(1)**: pp. 338–349.

Sampat, B. & H. L. Williams (2019): "How do patents affect follow-on innovation? Evidence from the human genome." *American Economic Review* **109(1)**: pp. 203–36.

Schechter, A. (2018): "Google and Facebook's "Kill Zone": "We've Taken the Focus Off of Rewarding Genius and Innovation to Rewarding Capital and Scale"." `https://tinyurl.com/dy6xmxbe`. Accessed: 2020-07-02.

Schmoch, U. (2008): "Concept of a technology classification for country comparisons." *Final report to the world intellectual property organisation (wipo), WIPO* .

Schnitzer, M. & M. Watzinger (2017): "Measuring the spillovers of venture capital." *Review of Economics and Statistics* pp. 1–48.

Schuhmacher, A., O. Gassmann, & M. Hinder (2016): "Changing R&D models in research-based pharmaceutical companies." *Journal of translational medicine* **14(1)**: pp. 1–11.

Scott Morton, F., P. Bouvier, A. Ezrachi, B. Jullien, R. Katz, G. Kimmelman, A. D. Melamed, & J. Morgenstern (2019): "Market structure and antitrust subcommittee report." *George J. Stigler Center for the Study of the Economy and the State, The University of Chicago Booth School of Business, Committee for the Study of Digital Platforms* .

Seldeslachts, J., J. Malek, & M. Newham (2021): "Killer acquisitions en innovatie." *KVS Preadviezen 2020* .

Shane, S. (2001): "Technological opportunities and new firm creation." *Management science* **47(2)**: pp. 205–220.

Shapiro, C. (2011): "Competition and innovation: did arrow hit the bull's eye?" In "The rate and direction of inventive activity revisited," pp. 361–404. University of Chicago Press.

Siebert, R. & Z. Tian (2020): "Dynamic Mergers Effects on R&D Investments and Drug Development across Research Phases in the Pharmaceutical Industry." *CESifo Working Paper* .

Sørensen, M. (2007): "How smart is smart money? A two-sided matching model of venture capital." *The Journal of Finance* **62(6)**: pp. 2725–2762.

Stiebale, J. & F. Szücs (2019): "Mergers and market power: Evidence from rivals' responses in European markets." *Technical report*, DICE Discussion Paper.

Szücs, F. (2014): "M&A and R&D: Asymmetric effects on acquirers and targets?" *Research Policy* **43(7)**: pp. 1264–1273.

Takebe, T., R. Imai, & S. Ono (2018): "The current status of drug discovery and development as originated in united states academia: the influence of industrial and academic collaboration on drug discovery and development." *Clinical and translational science* **11(6)**: pp. 597–606.

The Wall Steet Journal (2014): "Competition Is for Losers."

Trajtenberg, M., R. Henderson, & A. Jaffe (1997): "University versus corporate patents: A window on the basicness of invention." *Economics of Innovation and new technology* **5(1)**: pp. 19–50.

Ueda, M. & M. Hirukawa (2008): "Venture capital and industrial'innovation'." *Available at SSRN 1242693* .

Verhoeven, D., J. Bakker, & R. Veugelers (2016): "Measuring technological novelty with patent-based indicators." *Research Policy* **45(3)**: pp. 707–723.

Veugelers, R. (2012): "Innovation in EU Merger control: walking the talk." *Bruegel*

policy contribution **2012**.

Veugelers, R., M. Cincera, R. Frietsch, C. Rammer, T. Schubert, A. Pelle, A. Renda, C. Montalvo, & J. Leijten (2015): "The impact of horizon 2020 on innovation in Europe." *Intereconomics* **50(1)**: pp. 4–30.

Veugelers, R., G. Zachmann *et al.* (2020): *Racing against COVID-19: a vaccines strategy for Europe*. Bruegel.

Wagner, M. (2011): "To explore or to exploit? An empirical investigation of acquisitions by large incumbents." *Research Policy* **40(9)**: pp. 1217–1225.

Welch, X., S. Pavićević, T. Keil, & T. Laamanen (2020): "The pre-deal phase of mergers and acquisitions: A review and research agenda." *Journal of Management* **46(6)**: pp. 843–878.

Wollmann, T. G. (2020): "How to get away with merger: Stealth consolidation and its real effects on US healthcare." *Technical report*, National Bureau of Economic Research.

Yeo, H.-j. (2013): "Geography of mergers and acquisitions in the container shipping industry." *The Asian Journal of Shipping and Logistics* **29(3)**: pp. 291–314.

Yu, Y., N. Umashankar, & V. R. Rao (2016): "Choosing the right target: Relative preferences for resource similarity and complementarity in acquisition choice." *Strategic management journal* **37(8)**: pp. 1808–1825.

Zider, B. (1998): "How venture capital works." *Harvard business review* **76(6)**: pp. 131–139.

Faculty of Economics and Business
Naamsestraat 69
3000 LEUVEN, België
tel. + 32 16 00 00 00
fax + 32 16 00 00 00
feb@kuleuven.be
www.kuleuven.be

Doctoral dissertations from the Faculty of Economics and Business,
http://www.kuleuven.ac.be/doctoraatsverdediging/archief.htm