Essays in Regional and Innovation Economics

Knarik Poghosyan

Veröffentlichung als Dissertation in der Wirtschaftswissenschaftlichen Fakultät der Technischen Universität Dortmund

Dortmund 2024

Acknowledgments

I extend my heartfelt appreciation to all who have supported me throughout the writing of this thesis. Firstly, I am deeply grateful to my doctoral advisors, Philip Jung, and Nadine Riedel, for their invaluable support and guidance. Their profound grasp of economics and exceptional mentorship have not only shaped the content of this dissertation but have also played a pivotal role in my development as a researcher. I want to particularly thank Nadine Riedel for her exceptional devotion. Her close supervision has been instrumental in my success, and I am truly grateful for her invaluable contributions.

I am also thankful for the enriching experience of being part of the Research Training Group "Regional Disparities and Economic Policy," which provided a stimulating research environment. My gratitude extends to all the principal investigators and our research fellows, Arnaud Chevalier, and Wilhelm Kohler, for their valuable feedback and guidance. I am appreciative of my colleagues both from RTG and TU Dortmund for their support, insightful comments, and the enjoyable moments we shared. Especially, I want to express my sincere thanks to my co-authors, Martin Kalthaus, for his invaluable discussions, and Tobias Büscher, for our fantastic teamwork and shared research efforts.

Finally, I want to express my deepest gratitude to my friends and family—my mother, Lusine, my grandmother, Nina, and my sister, Emma—for their unwavering support and encouragement throughout my academic journey and the process of researching and writing this thesis. Above all, I want to extend a special thank you to my husband, Herve, for his love and support. His companionship and endless encouragement have been my pillars of strength. Thank you for being my proofreader, sounding board, and, most importantly, my best friend.

Contents

List of Tables

List of Figures

Introduction

Over the past decades, the importance of research and development (R&D) for economic growth and innovation widely increased. As a result, governments have increasingly implemented policies aimed at funding R&D activities, sometimes targeting specific regions and industries. However, the effectiveness of these policies in improving firm-level outcomes and regional development remains subject to debate among scholars. Some argue that such policies promote innovation and enhance regional prosperity (Aschhoff 2009; Breschi and Lissoni 2009; Crass et al. 2017; Czarnitzki and Fier 2002; Czarnitzki and Hussinger 2017; Czarnitzki and Licht 2006; Czarnitzki and Lopes-Bento 2011; Engel et al. 2017; Kleine et al. 2022), while others suggest that they may primarily benefit already successful firms, potentially exacerbating regional disparities (Brown et al. 2017; Cantner and Kösters 2011; Duranton 2011; Lehmann and Menter 2017).

This thesis contributes to the ongoing discussion by exploring the effects of R&D funding on individual firms and regional development. Comprising three self-contained essays, it examines the impact of both specific regional policies and broader R&D initiatives in Germany, shedding light on their implications for economic growth and innovation.

In Chapter 1, the focus lies on investigating a particular form of R&D funding - cluster funding characterized by its concentration within specific industries and regions. The aim is to unravel the intricate dynamics and regional implications associated with this targeted approach to R&D investment. To achieve this goal, we analyze the causal effects of cluster policy on firm performance, innovation activity, and R&D collaboration, using the German Leading-Edge Cluster Competition (LECC) as a case study. Through a dynamic event study design covering 2000-2016 and utilizing multiple datasets, including Bureau van Dijk's Orbis data and the EPO Worldwide Patent Statistical Database (PATSTAT), the chapter evaluates both direct effects on funded firms and potential spillovers on non-funded neighboring firms. Neighboring firms are identified through a combination of industry classification, utilizing 4-digit NACE codes for industries targeted by the LECC, and geographical proximity, determined by the distance to cluster firms.

The outcomes reveal a favorable influence of the cluster policy on the innovation and firm-level outcomes of funded firms. However, the spillover effect on non-funded neighboring firms appears limited, predominantly observed in firm-level outcomes within a 5 km radius.

When we further examine which types of firms benefit from the LECC policy in our heterogeneity analysis, we find that the policy significantly boosts firm-level outcomes, especially for small and medium-sized enterprises (SMEs). For larger firms, the impact is primarily seen in innovation outcomes. Interestingly, both funded and non-funded cluster members show significant positive effects. This suggests that the mechanisms implemented by the Leading-Edge Cluster Competition (LECC) are more influential than the actual funding in driving these positive outcomes.

These findings align with the extensive body of literature that highlights the positive impacts of cluster policy on the innovation outcomes of funded firms (Engel et al. 2017; Rothgang et al. 2017a; Uyarra and Ramlogan 2016). My research extends the existing evidence by providing an analysis of individual firm-level outcomes. Moreover, in contrast to existing literature which often conducts analyses at the regional level and adopts static short-term approaches, this study offers a fine-grained analysis at the firm level within a dynamic event study framework, focusing on medium to long-term dynamics. Additionally, it pioneers the examination of the indirect effects of cluster policy by reaffirming the knowledge spillover theory, highlighting that positive spillovers are predominantly observed among geographically proximate firms operating within similar industries.

In Chapters 2 and 3, I offer a comprehensive analysis of R&D funding in Germany, spanning from 1970 to 2020, utilizing the extensive Förderkatalog dataset. Chapter 2 begins by examining R&D funding trends in Germany, identifying key targets and characteristics associated with successful funding acquisition. It then delves into the direct impact of R&D funding on firm performance and innovation outcomes for both SMEs and large firms. Employing various datasets and empirical methods, this chapter uncovers a positive influence of R&D funding on SME performance, yet finds negligible effects on innovation outcomes for both SMEs and large firms. Despite efforts to mitigate pre-existing trends, the results suggest that factors beyond R&D funding may be driving these outcomes. This study contributes to the existing literature by offering a comprehensive analysis of various characteristics across all R&D funding schemes in Germany, in contrast to earlier research focusing on specific cases. Moreover, it enriches the literature by providing evidence of the effects of such initiatives on firm-level outcomes. Finally, it introduces methodological innovations by not only employing dynamic event study techniques but also highlighting the importance of addressing pre-trend violations.

While in Chapter 2, I delve into the actual effects of R&D funding, Chapter 3 takes a predictive approach to examine the potential impacts of R&D funding on firm performance and innovation activity in Germany from 2000 to 2020. By applying different empirical techniques following Fong et al. (2018), Girma et al. (2014), and Hudgens and Halloran (2008) and leveraging multiple databases (i.e. Förderkatalog, DAFNE, PATSTAT), Chapter 3 estimates these impacts by considering the overall effect and decomposing it into direct effects on funded firms and indirect (spillover) effects on non-funded firms. To address SUTVA violations and competition effects, we impose the Partial Interference Assumption and categorize all funded firms into region-industry-specific clusters, where the interactions are allowed within but not between clusters. Further, we analyze the share of funded firms within each cluster to identify potential spillover effects associated with varying proportions of funded firms. The findings reveal a generally positive impact of R&D funding on both funded and non-funded firms, with effects varying based on the concentration of funded firms in the cluster. Interestingly, the study highlights that the indirect impact on innovation outcomes outweighs the direct impact, underscoring the potential for broad R&D funding to stimulate innovation spillovers despite limitations in driving pronounced direct effects.

Chapter 3 introduces a novel approach to the existing literature by dissecting the total effect of R&D funding into direct and spillover effects, while also utilizing the share of funded firms as a mechanism to identify spillovers. This methodological innovation adds depth to our understanding of the dynamics of R&D funding and its impacts on firm performance and innovation. Furthermore, the research contributes valuable insights to the policy discourse by suggesting that more generic forms of funding may be effective in stimulating spillover effects, whereas more concentrated types of funding (e.g. cluster policy) could be better suited to drive direct effects. This nuanced perspective underscores the importance of tailoring R&D funding policies to specific objectives and contexts to maximize their effectiveness in fostering innovation and economic growth.

CHAPTER 1

How effective are cluster policies? Evidence from Germany

*w*ith Nadine Riedel, Martin Kalthaus, Michael Rothgang, and Anne-Marie Scholz

Chapter Abstract

We analyze the causal impact of cluster policy on firm performance, innovation activity and R&D collaboration using the German Leading-Edge Cluster Competition (LECC) as a testing ground. The analysis does not only quantify the direct effects of the policy on funded firms but also tests for possible spillovers on non-funded neighboring firms. This allows us to evaluate whether the LECC came with broader benefits for the development of regions that host cluster firms. Firm performance, collaboration, and innovation outcomes are measured by different financial accounting variables, their patent applications, and patent-based indicators for firm collaboration (i.e. joint patents). For that purpose, we use multiple data sets (Bureau van Dijk's Orbis data, EPO Worldwide Patent Statistical Database (PATSTAT), and the LECC Survey data) and implement a dynamic event study design covering the period of 2000-2016. To investigate the impact on neighboring non-funded firms, we identify non-funded firms in geographic proximity to cluster firms (the analysis accounts for distance rings up to a 10 km radius). The results suggest that the cluster policy had a positive impact on the firm performance of funded firms and cluster-hosting regions in 5 km proximity. However, the impact on innovation is only positive for funded firms and no impact is found for innovation and collaboration outcomes for non-funded neighbor firms. These findings align with the results of previous studies.

1.1 Introduction

Over the last decades, considerable attention has been paid to clusters and cluster policies. Contrary to public R&D subsidies for individual firms and projects, cluster policies promise to be particularly successful in stimulating innovation as they support the systemic nature of the innovation process (Smits and Kuhlmann 2004). Despite the increasing popularity of cluster policies, the jury is still out regarding their effectiveness (Graf and Broekel 2020). Many studies have found that cluster policies, due to their geographically concentrated setup, stimulate more interactions between cluster members, which contribute to the growth of firms' performance and innovation activity on a regional level (Akcigit et al. 2018; Baptista and Swann 1998; Breschi and Lissoni 2009; Crass et al. 2017; Engel et al. 2017; Engel et al. 2013; Hinzmann et al. 2017; Nishimura and Okamuro 2010; Uyarra and Ramlogan 2016). The latter idea has been supported by Acemoglu et al. (2016) and Breschi and Lissoni (2005) who found that knowledge flows are localized and the geographic proximity of collaborating actors may lead to more innovations via knowledge spillover. In contrast, many scholars claim that cluster policies are very costly and usually have only a short-term positive effect (Duranton 2011; Engel et al. 2013; Martin et al. 2011).

To shed more light on the ongoing debate about cluster policy benefits and fill the lack of cluster policy impact analyses on the firm level, our principal objective is to analyze the causal effect of cluster policies on innovative activity, firm performance, and cooperation across actors. We use the implementation of the German Leading-edge cluster competition as a testing ground. The latter stands out as one of the largest cluster policy initiatives globally, making it inherently compelling for study. The main research questions are (1) whether the policy increased innovative output and firm performance of funded firms and induced more collaborations, and (2) whether there are knowledge spillovers of the policy, which affect non-funded neighboring firms. The latter question follows the ideas raised in the studies of Breschi and Lissoni (2005) and Acemoglu et al. (2016), where the authors claim that a close location to the cluster region might lead to a better performance of non-funded neighbors via knowledge spillovers.

To address the research questions, we combine multiple datasets. We collect firm-level data from Orbis for the period 2000-2016 and combine with patent data from PATSTAT. Furthermore, we use a novel, geo-referenced data set, which contains information on firms funded by the Leading-edge cluster competition, which allows us to exploit geographical information to measure spillover. Methodologically, we employ an event study design following Sun and Abraham (2021) and matched difference in difference analysis (Iacus et al. 2011; Iacus et al. 2012).

This paper contributes to the existing literature in different ways. First, existing research on cluster policies concentrates mostly on regional analysis and only in the short term and a static framework (Cantner 2013; Graf and Broekel 2020; Rothgang et al. 2017a). Our paper is the first to exploit firm-level data and observe firm-level effects of the German Leading-Edge Cluster Competition (LECC) policy in a medium/long term and a dynamic event study framework. Furthermore, our research also involves the analysis of spillover effects for neighboring firms. This is naturally infeasible when using regional data as in prior work. Audretsch et al. (2018), Crass et al. (2017), and Lehmann and Menter (2017) touch upon this question, however detailed analyses tackling the spillover effects for performance, innovation, and collaboration activity on the firm level, medium/long term and in a dynamic setup are, to the best of our knowledge, still missing. Our results offer valuable insights both, for the academic literature and policymakers.

The remainder of the paper is organized as follows. Section 1.2 presents the theoretical framework used to study the research questions and derives core testable predictions of this paper. Section 2.3.1 introduces the Leading-edge cluster competition by focusing on the application procedure, main goals, and targets of the policy. Section 1.4 presents the data and some summary statistics. In section 1.5, the empirical framework is explained, first by discussing the estimation strategy, and then the identification strategy with underlying assumptions. Empirical results are presented in section 3.5. Finally, section 3.6 concludes the paper with a summary and discussion of the main findings and limitations and put forwards policy recommendations and suggestions for future research.

1.2 Literature Review

1.2.1 Empirical Evidence on Cluster Policies Worldwide

Currently, advanced economies are significantly reliant on science, technology, and innovation, playing pivotal roles in their global competitiveness. In response, national and regional governments craft and implement "innovation policies" with the goal of bolstering overall economic competitiveness or specific sectors, thereby fostering social welfare growth (Kuhlmann and Edler 2003).

The notion of "innovation policy" is expansive, covering research, and technology, and intersecting with industrial, environmental, labor, and social policies. In essence, it refers to the collective set of governmental initiatives aimed at fostering science, education, research, technological development, and the modernization of industries (Kuhlmann and Edler 2003).

Over the past years, there has been a growing emphasis on innovation policies targeting the formation of networks and clusters. Cluster policies typically concentrate efforts on supporting specific industries and technological niches. This may involve subsidizing economic activities in disadvantaged areas or nurturing promising regions (Autio and Rannikko 2016; Müller and Korsgaard 2017). This strategic approach seeks to enhance both sectoral and regional innovation systems by facilitating knowledge spillover, improving skills and finances, creating better conditions for attracting foreign direct investments (FDI), and generating other public good effects. Moreover, these policies exert a significant influence on mitigating market failures, such as underinvestment in technology and knowledge due to information asymmetry, network effects, and externalities, as well as addressing system and governmental failures, including poor actor connections and inefficiencies in existing programs and institutional delays in certain regions (Uyarra and Ramlogan 2016). The cluster approach, with its strong regional embeddedness, and emphasis on nurturing networks, stimulating entrepreneurial opportunities, and establishing relevant institutions, proves to be a valuable tool for policymakers (Autant-Bernard et al. 2013; Chatterji et al. 2014; Delgado et al. 2015). It allows for effective control over knowledge and innovation essential for overcoming barriers (Broekel et al. 2015; Smits and Kuhlmann 2004).

Numerous studies have delved into the impacts of cluster policies, with a division into two distinct strains examining the direct and indirect effects of the policy. This research extends to evaluations conducted both in Germany (Rothgang et al. 2021) and on an international scale (Uyarra and Ramlogan 2016). However, it's noteworthy that the predominant focus in the analysis of cluster policy lies in its direct effects. These studies predominantly scrutinize the policy's impact from the perspective of increased innovation and collaboration outcomes.

The comprehensive review of various cluster policies worldwide by Uyarra and Ramlogan (2016) revealed that these policies increased the probability of innovating and effectively promoted and facilitated collaborations among different entities. For instance, the NEC program in Norway serves as a prime example, demonstrating how it spurred an uptick in innovation projects by encouraging greater experimentation and fostering new collaborations. This trend was mirrored in other cluster initiatives such as the UK's Yorkshire and West Midlands programs, as well as Sweden's Vinnväxt initiative. Research indicates that 30-35 % of firms involved in these cluster programs cited establishing new business contacts as a key advantage (Uyarra and Ramlogan 2016). The latter results found support in studies for Spanish and Japanese clusters conducted by Aranguren et al. (2013) and Nishimura and Okamuro (2011). In Japan, for example, the evaluation of the Industrial Cluster Policy (ICP) by Nishimura and Okamuro (2011) demonstrated how close collaboration with national universities contributed to increased innovation outputs. This partnership not only provided access to new knowledge but also facilitated the exchange of tacit knowledge and reduced uncertainty by integrating into local collaborative networks. Further, research on the Basque paper cluster conducted by Aranguren et al. (2013) highlighted the positive impact of cluster membership on innovation activities and labor productivity. Their study uncovered a direct positive correlation between participation in clusters and engagement in innovative endeavors. Moreover, it revealed an indirect positive effect of cluster involvement on the growth of productivity in innovative activities. Limited short-term productivity effect is also found in the study on French clusters (Martin et al. 2011).

Conversely, another group of studies suggests that clusters may yield adverse effects, particularly due to their practice of "picking winners" (Brown et al. 2017; Cantner and Kösters 2011; Lehmann and Menter 2017). Duranton (2011) argue that this strategy may lead to drawbacks such as crowding out private investments or exacerbating inequality across regions. However, when cluster policies concentrate on lagging regions, their impact may weaken, as observed in the study by Martin et al. (2011). The authors noted that while the policy positively influenced firm-level productivity, the effect was extremely weak and short-term. They speculated that the policy might have aimed at protecting some large firms, and the absence of any significant impact could be attributed to the small amount of funding, insufficient to trigger substantial change.

1.2.2 Empirical Evidence on Cluster Policies in Germany

In Germany, a couple of studies addressed the input effects concerning R&D expenditure of business firms. Engel et al. (2017) and Rothgang et al. (2017a) analyze the Leading Edge Cluster Competition (LECC) and find that it stimulated the growth of R&D expenditures of funded firms. The study on Bavarian High-Tech Cluster by Falck et al. (2010) adds that the cluster policy also enhances the probability of becoming an innovator in the target industries. In another study, Engel et al. (2013) discuss the impact of the German biotech contest on firm-level R&D activity. Their analysis shows that participation in cluster policies had a positive short-term effect on R&D activities.

Some studies also find a substantial trigger effect of cluster policies on the development of cooperation during the funding period. This has been shown by some qualitative studies (Fromhold-Eisebith and Eisebith 2008; Kiese and Hundt 2014). By looking at local patent co-applications, Graf and Broekel (2020) find that the BioRegio contest in Germany had increased the cooperation network (as measured by co-invention networks) but that the network effect disappeared after the funding period. This is an indication that the network and cooperation effects of cluster funding could be transitory. Also, the Leading-Edge Cluster Competition has increased the cooperation networks of the clusters involved in the programs (Cantner 2013; Hinzmann et al. 2017; Rothgang et al. 2017a; Töpfer et al. 2017).

All of these studies on the impact of the Leading-Edge Cluster Competition address short-term funding effects. The same is true for studies on other cluster policies like the Bavarian cluster program (Falck et al. 2010). In fact, not much is known about the medium- to long-term effects of cluster promotion of the Leading-Edge Cluster Competition and other German Programs (Rothgang et al. 2021). The only German programs that have been analyzed from a more long-term perspective are the Biotech programs (e.g., Dohse and Staehler (2008)). Another notable example is Graf and Broekel (2020) who look at the influence of cluster promotion at innovation networks of the Bioregio contest. At the same time, it is still unclear, how far a medium- to long-term impact on the innovation outcomes and survival of the extended networks and new cooperation activities can be observed. Overall, the clarity surrounding the potential innovation additionality and network effects in the medium term is yet to be fully ascertained.

Further, according to the recent evaluations by Rothgang et al. (2021) the information on the effect of cluster funding on firm output and performance is still missing. Especially intriguing could be the examination of behavioral impacts on firms' performance, assessed through firm financial indicators. In the sole exploration by Lehmann and Menter (2017), the evaluation of cluster policy effects on regional productivity in Germany unveils a significant upswing in productivity within funded regions as a consequence of the German Leading-Edge Cluster Competition (LECC). It is essential to note, however, that this study is confined to a regional scope. In contrast, our micro-level analysis enhances the identification strategy by effectively controlling unobserved heterogeneity. It also facilitates heterogeneity and spillover analysis, offering a clearer insight into the mechanisms at work and identifying the main beneficiaries of the policy.

Finally, the existing studies have predominantly approached the evaluation of cluster policies within a static framework, limiting our ability to discern the dynamic effects of the policy. By analyzing the dynamic effects of cluster policies, we are able to provide deeper insights into their long-term impacts and sustainability, revealing how benefits evolve and persist over time. This approach highlights the stages of peak effectiveness and identifies behavioral adjustments by firms, which static analyses may miss.

In light of these considerations, we aim to explore whether the policy increased innovative output, collaboration, and firm performance of funded firms, and we set the following hypothesis:

Hypothesis 1: *Cluster policy has a positive impact on the performance, innovation outcomes, and R&D collaboration of funded firms.*

By looking at the causal effects of the Leading-Edge Cluster Policy on innovative activities, firm performance, and cooperation patterns in a medium-long-term dynamic framework, our analysis delves into new ground and enriches the existing literature by providing fresh insights into the medium-long-term dynamic impacts of cluster policies on various dimensions of economic activity.

Pioneering research conducted by Glaeser et al. (1992) suggests that the spatial clustering of firms fosters the circulation and exchange of knowledge, both within specific industries and across different sectors thus promoting innovation and driving overall economic growth.

This idea has been further developed by Breschi and Lissoni (2005) and Acemoglu et al. (2016), who discuss that close location to the cluster region might lead to a better performance not only for the funded cluster firms but also for the non-funded neighbors via knowledge spillover. The study by Crass et al. (2017) provides indirect evidence on the possible effects of clustering on both inputs (R&D intensity) and output (sales with a new product). The study looks at data from the German innovation program “Innovative SMEs” and asks, whether geographical clustering influences the effectiveness of the funding scheme. This would not mean that there are in fact policy spillovers, but it would show that positive cluster effects exist that influence the effect of policies directed at cluster firms. However, the authors find no evidence of input or output additionality of geographical clustering.

However, dissenting viewpoints such as those expressed by Burger et al. (2014) raise concerns about the exclusive emphasis on the positive externalities of regional agglomerations. These critics argue that it is crucial to also consider the concept of "borrowed size", as "agglomeration shadows" [1] may hinder collective benefits and result in unintended adverse outcomes.

In this context, Lehmann and Menter (2017) investigate the effects of the LECC on regional productivity, documenting a positive effect limited to non-funded regions with sufficient absorptive capacity. Conversely, non-funded neighbors faced challenges attributed to increased competition. A more recent study by Audretsch et al. (2018) explores the spillover effects of the LECC on non-funded firms outside the targeted industries. The findings suggest an indirect negative impact, creating what can be termed as "agglomeration shadows".

Nevertheless, there are no clear indications that firms that are in the cluster regions or even members of the cluster organizations but not funded in the Leading-edge cluster competition do benefit from the funded cooperation projects (Rothgang et al. 2017a). Thus, we ask whether there are knowledge spillovers of the policy, which affect non-funded neighboring firms in close geographic proximity. Following the results of the study by Audretsch et al. (2018), indicating the absence of positive spillovers for the neighbors in non-targeted industries, we focus on geographically close located neighbors in the LECC targeted industries and put forward Hypothesis 2:

Hypothesis 2*: Cluster policy has a positive impact on the performance, innovation outcome,s and R&D collaboration of non-funded neighbor firms due to geographical and technological spillovers.*

[1] The term "agglomeration shadows" refers to the phenomenon where regional agglomerations, such as cities, clusters, or urban centers, diminish the performance of nearby regions by absorbing economic resources (Dobkins and Ioannides 2001; Fujita et al. 1999; Partridge et al. 2009)

1.3 The Leading-Edge Cluster Competition

The Federal funding program "Leading-Edge Cluster Competition" (LECC) was initiated by the Federal Ministry of Education and Research in 2008. Public financing ended in 2017. Cluster initiatives were invited to apply for funding in the program by formulating a common strategy. These cluster initiatives consisted of more or less formalized consortia of actors (firms, universities, research organizations, and other organizations) mainly located within a cluster region. The program endowed successful applicants with financial support for different projects (mainly cooperative research projects). These projects contribute to the realization of a common cluster strategy (BMBF 2009). The program itself differs from many other cluster programs in respect to the endowment with substantial program funding and a rather rigorous selection process based on the strategy documents that were provided in the application process.

The program was open to initiatives from all high-tech fields. According to the program targets, the Leading-Edge Clusters should evolve into the top group of globally leading clusters in their respective fields of technology. The program also aimed to stimulate better performance and increased R&D cooperation not only among funded actors but also expand the policy benefits for neighboring regions. The selection procedure extended over 3 consecutive rounds. In each round, 5 LECs were chosen by the Ministry based on the proposal of an expert jury. Funding for the three consecutive rounds of the competition amounted to 600 Mill. €. Each winner received approximately 40 Mill. € of government funding. According to the program regulations, the individual funding recipients had to provide another 600 Mio. € of project funding from their own research and development budget.

A total of 87 cluster initiatives from a broad range of industries applied for promotion in the 3 rounds of the competition. Some applicants who failed in the 1st or 2nd contest round applied once more in the consecutive round. A comparison of the applicants and winner clusters with around 350 technology-oriented cluster initiatives that are classified as "federal state" or "federal clusters" in Germany shows that the industry structure of the LECs is quite similar to the industry structures of all innovation clusters in high-tech and medium high-tech sectors of the economy (Rothgang et al. 2014). A more comprehensive overview of the LECC is provided in Appendix A.1, offering detailed insights into its structure, objectives, targets, and selection process.

1.4 Data

To estimate the impact of the Leading-Edge Cluster Competition on the performance and R&D collaboration of funded and non-funded neighbor firms, we rely on different types of data all covering the period of 2000-2016 and targeting specifically Germany. The first type of data is Bureau van Dijk's "Orbis" database (extracted in September 2020), which is a geo-referenced firm-level data containing business records for around 2 mln firms in Germany. This data contains information on firm finances, legal entity details, corporate structures, ownership, M&A activity, intellectual property, and other information, such as location, sector, year of incorporation, etc. This data is used to estimate the impact of cluster policy on firm performance via firm employment costs and fixed assets.

The second type of data is the EPO Worldwide Patent Statistical Database (PATSTAT) created by the European Patent Office (EPO) upon the request by a task force led by the Organisation for Economic Co-operation and Development (OECD). We retrieved the data from PATSTAT Spring Edition 2020 (extracted in June 2020), which contains bibliographical and legal event patent data with detailed information about the names and addresses of applicants and inventors, priority year, patent families, the title and abstract of patent applications, citation links, and patents classification by technology class. In this paper, we use PATSTAT data for Germany (which includes 69 898 firms and around 4 609 894 patent applications) to calculate the quality-adjusted number of patents (refer to Appendix A.2) for details) and identify R&D collaboration between firms via joint patenting.

The third type of data covers the Leading-Edge Cluster Competition, which we collected manually with the help of surveys of funded cluster initiative participants. The surveys were implemented from 2012 to 2013. The data involves information on application procedures such as application documents and the list of applicant clusters (67 in total) for each funding round (2008, 2010, 2012) by differentiating rejected, finalist, and successful cluster initiatives. The data also provides a complete list of actors involved in 15 funded clusters (2349 actors in total) and their location information by again differentiating between just cluster members (1528 actors) and those who got direct funding from the LECC (821 actors). Besides, we can distinguish between the types of cluster actors such as large firms, SMEs, research institutes and universities, and other actors. Additional details concerning the data generation process and descriptive statistics are available in the Appendix A.2.

1.5 Empirical Framework

1.5.1 Estimation Strategy

1.5.1.1 Exact Coarsened Matching

Since LECC policy targeted a specific range of firms, it makes the comparison between the LECC-funded and non-funded firms difficult due to existing differences between the outcomes. The most effective and intuitive empirical identification strategy would involve comparing cluster-funded firms to runner-ups. However, this is not feasible in our study because we only have data on the funded cluster members and not the runner-ups. Hence, to address these differences and selection issues, we conduct Coarsened Exact Matching (CEM) (Iacus et al. (2011) and Iacus et al. (2012)) to match each cluster firm with a suitable control firm. To avoid the matches with multiple control firms, we use the Nearest Neighbor matching algorithm inside CEM (specifying $k2k = TRUE$) with Mahalanobis distance. In this case, we get for each cluster firm only one control firm and have a balanced sample [2].

To reduce the matching sample size and calculation time, we reduce the full sample to the firms involved in similar industries funded by the LECC (based on the NACE 4-digit code). In this sample, for each cluster firm, the pool of potential control firms includes all non-cluster firms, which have never been funded by the LECC policy. By implementing CEM, we match exactly on firm age (calculated as $Year_{2007} - Year_{Incorporation}$), firm size ($ln(Total\ assets)$), and industry (Nace 2-digit code)), all measured in $t-1$ year (2007)

[2]This matching procedure allows us to successfully match 96.14% of the LECC firms to control firms.

(before the LECC implementation). It is very important to match on industry because the funded-cluster cohorts vary with respect to industry and industry affiliation significantly affects firm's selection into the cluster. Moreover, matching on firm age and size ensures that each LECC firm is matched to a control firm with a similar lifespan and market potential.

Table 1.1 and 1.2 present summary statistics for the LECC-funded and non-funded firms before and after the matching (see Appendix A.3.1 for further details on the unmatched sample). The first three columns report summary statistics for the LECC-funded and non-funded firms before matching and columns three, four, and five depict summary statistics after matching. Before matching, LECC-funded firms are on average about 2.9 years older than non-funded firms. They slightly differ with respect to firm size. There is a large difference between the patent output and hence patent-related collaboration of LECC-funded and non-funded firms, where LECC-funded firms produced about 15 times more patents in comparison to non-funded firms.

After matching (columns four and five), we observe only very small differences between LECC-funded and non-funded firms, which indicates that matching has been successfully implemented. By construction, the LECC-funded and non-funded firms are now identical with respect to coarsened covariates explicitly matched on. However, the LECC-funded and non-funded firms became also more identical in terms of covariates that were not used in the matching procedure.

Table 1.1: Descriptive statistics: Before vs after matching

	Before Matching			After Matching		
	LECC (n= 751)	Non-LECC (n=584296)	Overall (n=585047)	LECC (n= 722)	Non-LECC (n= 722)	Overall (n=1444)
Firm age	18.9 (27.0)	16.0 (21.0)	16.0 (21.0)	16.7 (20.3)	16.7 (20.1)	16.7 (20.2)
Firm size	15.5 (2.84)	12.9 (2.19)	12.9 (2.19)	15.2 (2.72)	15.2 (2.72)	15.2 (2.72)
Patent (simple)	79.6 (389)	5.41 (67.2)	7.48 (93.5)	10.8 (29.9)	7.12 (11.6)	9.56 (25.4)
Patent (adjusted)	226 (1120)	16.0 (203)	21.9 (276)	7.93 (53.4)	2.38 (14.2)	5.15 (39.2)
Joint patents	20.6 (143)	0.501 (13.9)	0.773 (21.7)	3.02 (25.2)	1.31 (4.21)	2.44 (20.7)
Co-partners	32.4 (242)	0.877 (23.5)	1.30 (36.7)	3.90 (18.4)	2.88 (9.21)	3.55 (15.9)
Fixed assets	709 (6810)	5.65 (184)	6.41 (290)	75.4 (717)	72.7 (588)	74.0 (655)
Total assets	1060 (8620)	9.48 (284)	10.6 (402)	137 (994)	154 (1090)	145 (1040)
Turnover	1670 (8610)	33.4 (494)	37.3 (652)	271 (840)	268 (885)	270 (860)
Sales	1680 (8680)	43.8 (502)	49.8 (731)	257 (789)	267 (863)	262 (824)
Employee costs	282 (1470)	8.72 (119)	10.2 (161)	43.5 (155)	31.9 (76.3)	38.2 (125)

Notes: Fixed & total assets, turnover, sales and labor costs values are in Mln.

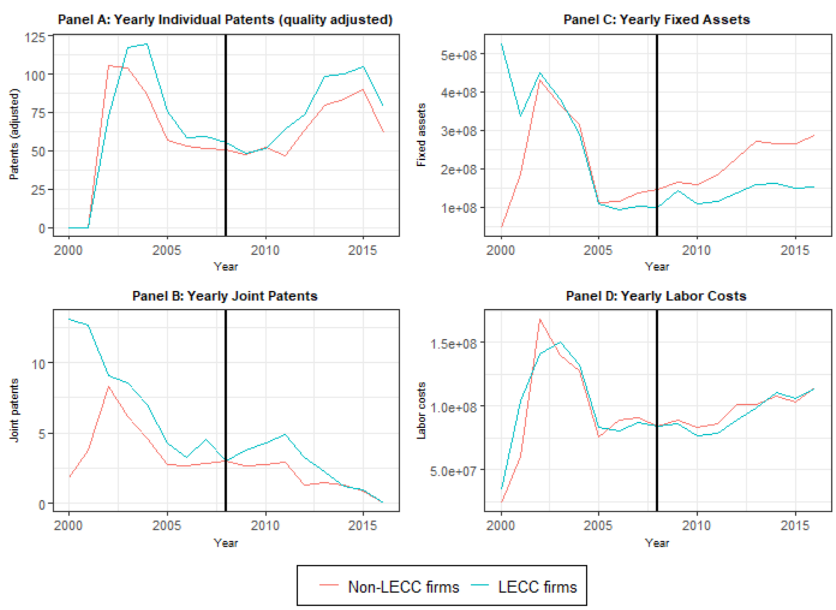
Figure 1 shows the means of innovation, collaboration, and firm-level outcomes of the LECC-funded firms and their matched control firms during 2000-2016. In Figure 1 we observe that the pre-funding trends of LECC firms and matched control firms are very similar, which verifies that the matching procedure for selecting a suitable control group worked well. In general, Figure 1 depicts that the LECC policy funding leads to the growth of the number of quality-adjusted individual patents (panel A), fixed assets (panel C), and labor costs (panel D) confirming existing evidence (Engel et al. 2017; Lehmann and Menter 2017). In contrast, the number of joint patents (panel B) drops for the LECC firms one year before the policy funding and then has a growing trend up to 2012, which declines again from 2013 onward.

Table 1.2: Descriptive statistics: Before vs after matching

Industry	Before Matching LECC (n= 751)	Before Matching Non-LECC (n=584296)	Before Matching Overall (n=585047)	After Matching LECC (n= 722)	After Matching Non-LECC (n= 722)	After Matching Overall (n=1444)
Administrative and Support Service Activites	7 (0.9%)	8275 (1.0%)	8282 (1.0%)	5 (0.7%)	4 (0.6%)	9 (0.6%)
Construction of Buildings	2 (0.3%)	27686 (3.4%)	27688 (3.4%)	3 (0.4%)	1 (0.1%)	4 (0.3%)
Education	4 (0.5%)	5518 (0.7%)	5522 (0.7%)	3 (0.4%)	3 (0.4%)	6 (0.4%)
Financial and Insurance Activities	2 (0.3%)	8129 (1.0%)	8131 (1.0%)	2 (0.3%)	0 (0%)	2 (0.1%)
Human Health and Social Work Activities	4 (0.5%)	4373 (0.5%)	4377 (0.5%)	4 (0.6%)	3 (0.4%)	7 (0.5%)
Information and Communication	19 (2.4%)	5833 (0.7%)	5852 (0.7%)	25 (3.5%)	19 (2.6%)	44 (3.0%)
Manufacturing	13 (1.7%)	4301 (0.5%)	4314 (0.5%)	10 (1.4%)	8 (1.1%)	18 (1.2%)
Other Service Activities	6 (0.8%)	6386 (0.8%)	6392 (0.8%)	4 (0.6%)	3 (0.4%)	7 (0.5%)
Professional, Scientific and Technical Activities	22 (2.8%)	25187 (3.1%)	25209 (3.1%)	25 (3.5%)	28 (3.9%)	53 (3.7%)
Real Estate Activities	5 (0.6%)	70568 (8.6%)	70573 (8.6%)	5 (0.7%)	5 (0.7%)	10 (0.7%)
Transportation and Storage	4 (0.5%)	5895 (0.7%)	5899 (0.7%)	4 (0.6%)	2 (0.3%)	6 (0.4%)
Wholesale and Retail Trade	21 (2.7%)	54526 (6.7%)	54547 (6.7%)	14 (1.9%)	16 (2.2%)	30 (2.1%)
Accommodation and Food Service Activites	0 (0%)	8363 (1.0%)	8363 (1.0%)	0	0	0
Activities of households as employers	0 (0%)	19 (0.0%)	19 (0.0%)	0	0	0
Arts, Entertainment and Recreation	0 (0%)	2404 (0.3%)	2404 (0.3%)	0	0	0
Public Administration and Defense: Compulsory Social Security	0 (0%)	1354 (0.2%)	1354 (0.2%)	0	0	0
Water Supply	0 (0%)	1256 (0.2%)	1256 (0.2%)	3 (0.4%)	3 (0.4%)	6 (0.4%)

This can be explained by the fact, that usually patenting requires time after a joint R&D project, and since individual patents start to have an upward trend only after 2010 (which is probably the outcome of earlier R&D activity not related to the LECC funding), the impact on collaboration triggered via LECC will take longer time to be visible.

Figure 1.1: Raw means after Coarsened-Exact Matching



Notes: This figure shows raw means for the LECC-funded firms and their matched control firms. The LECC firms are funded in three rounds in 2008, 2010 and 2012 respectively. The outcome data are measured in calendar years and observed from 2000-2016.

CEM is also applied to subsamples, distinguishing between SMEs and large firms, as well as funded and non-funded cluster members used in the heterogeneity analysis to assess the variations in the direct impact of the LECC. For more detailed information, please refer to Appendices A.3.2 and A.3.3.

1.5.1.2 Event Study

1.5.1.2.1 Own (direct) effect

To analyze the impact of the LECC on funded firms, we apply the event study design following Sun and Abraham (2021) (TWFE with adjusted interaction term). Here we define $D_{i,t}$ to be ever having been treated by the LECC and we estimate the impact of binary treatment $D_{i,t}$ on the outcome $y_{i,t}$ in a panel of units i and periods t ($E[\, y_{i,t}|D_{i,t}]$ is estimated parametrically) . We concentrate on a "staggered treatment" design, where for each unit i, $D_{i,t} = 1$ if it is treated at a specific event date E_i. We also allow for never treated units, $D_{i,t} = 0$, where we set $E_i = \infty$ (also referred as the "baseline outcome") for never treated unit i for all time periods t.

In our terminology, we group firms into cohorts e for $e \in \{0, ..., T, \infty\}$ depending on the time of receiving the treatment (E_i is defined as the survey wave of their initial treatment) where all firms in each cohort e get the first treatment at the same time $\{i : E_i = e\}$. Since the LECC policy was implemented in three rounds, hence, we end up with three cohorts E_i for $e \in \{1, 2, 3\}$ which get policy treatment in 2008, 2010, and 2012 respectively.

We estimate the firm-level impact of the LECC in the TWFE model a la Sun and Abraham (2021) with an adjusted interaction term (i.e. $first\, treatment\, period * year$). Here we estimate two types of effects: firm-level treatment effect and cohort-specific average treatment effect.

The firm-level treatment effect is estimated as the difference between the observed outcome relative to the never-treated counterfactual outcome: $Y_{i,t} - Y^{\infty}_{i,t}$ (the latter one is the "baseline outcome", i.e. potential outcome if firm i is never funded by the LECC).

The cohort-specific average treatment effect on the treated ($CATT_{e,l}$) is the average treatment effect l periods from the initial treatment for the cohort of firms first treated by the LECC at time e.

$$CATT_{e,\ell} = E\left[Y_{i,e+\ell} - Y^{\infty}_{i,e+\ell} \| E_i = e\right]. \tag{1.1}$$

Here, instead of calendar time index t we use relative period index l (indicates the periods since treatment for cohort e), where $l \in \{-e, T-e\}$ because we observe e periods before and $T-e$ periods after the initial treatment. Since our study covers the period of 2000-2016, we account for 8 pre and 8 post-treatment relative periods (leads and lags), which enable us the comparison of three cohorts by keeping their treatment exposure constant. Our dynamic specification has the following form:
General binned specification

$$Y_{i,t} = \alpha_i + \lambda_t + \beta \cdot \sum_{\ell<-K} D^{\ell}_{i,t} + \sum_{\ell=-K}^{-2} \mu_\ell D^{\ell}_{i,t} + \sum_{\ell=0}^{L} \mu_\ell D^{\ell}_{i,t} + \gamma \cdot \sum_{\ell>L} D^{\ell}_{i,t} + \varepsilon_{i,t}, \tag{1.2}$$

where $y_{i,t}$ is the outcome of interest (quality-adjusted number of patents, joint patents, fixed assets, employment costs) for cluster firm i at time t, α_i is firm fixed effects, λ_t is the time fixed effects, μ_l is the average effect for being treated in periods l (period since treatment for cohort e), $D^{l}_{i,t}$ is the relative period indicator (lags/leads of treatment), and $\varepsilon_{i,t}$ is the error term (robust standard errors are adjusted for both heteroscedasticity and clustering of observations (clustered at the firm level)).

Following the common practice in the literature discussed by Sun and Abraham (2021), we implement two steps to avoid multi-collinearity. Firstly, we exclude the relative period before the initial treatment (2007) by dropping $D_{i,t}^{-1}$ from our specification. Then, we bin distant relative periods into [2000:2004), and (2014:2016] and end up with 4 pre and 4 post-treatment periods. In the further section, we will discuss in detail whether identifying assumptions proposed by Sun and Abraham (2021) are likely to hold in our context.

1.5.1.2.2 Spillover (indirect) effect

To observe the spillover effects of the LECC on neighboring firms, we construct geographic rings indicating the distance between funded cluster firms and non-funded neighbor firms. First, by following the results by Audretsch et al. (2018) claiming the absence of the positive spillover effect for non-targeted industry firms, we constrain the sample of non-funded firms to cluster targeted industries to assure technological proximity required for the knowledge transfer and absorption. Then, to account for geographic proximity, we calculate the distance for each non-funded neighbor to all cluster firms. The distance is calculated based on the longitude and latitude information between two actors by finding all cluster neighbor firms for each non-funded firm. However, one should keep in mind, that the calculated distance only represents the air distance between two firms and does not account for any specificity related to the existing road network structure. With the help of the obtained distance matrix, we construct in total of 2 geographic rings indicating neighbors in a 5km, and from a 5-10km distance. Then, we count for each non-funded firm how many cluster neighbors are present in within a 5km and 5-10 km distance. Further, to analyze our second hypothesis, we use the TWFE model for staggered treatment by Sun and Abraham (2021) and we restrict our sample to firms that are not funded by the policy by dropping the cluster firms. In contrast to direct effect analysis, where the treatment group consists of cluster firms and the control group comprises non-cluster firms, in spillover analysis our treatment group represents the non-funded neighbors which are within a 10 km distance to cluster firms and the control group consists of all other non-funded firms located beyond a 10 km distance. We estimate the TWFE model for both subgroups of neighbors (within 5km and 5-10 km distance) separately by applying Equation 1.2 discussed earlier.

In addition to the primary analysis, we conduct a series of robustness checks by employing various specifications with additional fixed effects. The details of this analysis can be found in Appendix A.6.

1.5.2 Identification Strategy

According to Sun and Abraham (2021), the implementation of the event study design requires the following three identification assumptions:

Assumption 1 (Parallel Trends in Baseline Outcome): Since the LECC's main goal is to support excellence clusters with a strong focus on innovative projects, the probability to be funded by the LECC is higher among firms that are already involved in R&D activity and had higher performance earlier. Consequently, it is not probable that the baseline outcome Y^{∞} is mean independent of the timing of the policy funding. Here parallel trends assumption is more reasonable since it permits the dependence of the funding period on unobserved time-invariant characteristics (e.g. change in opportunities, trends, technological growth, etc.). Besides, clusters supported earlier by policy funding (particularly R&D) have an upward trend for total performance and R&D activity in contrast to clusters that have never been funded by the policy. However, this issue is avoided in our setup since all the cluster initiatives supported by the LECC have been established after the policy funding, which means we have no confounding bias coming from this channel.

Assumption 2 (No Anticipatory Behavior Prior to Treatment): We assume that there is no anticipatory behavior preceding the treatment under the LECC policy. This is supported by the fact that the tender began in September 2007, shortly before the funding in 2008, providing little time for firms to adapt their behavior. Moreover, the policy targets innovative activities, which are inherently challenging to adjust within a short timeframe. However, if firms possess private information regarding the likelihood of the LECC funding in advance and modify their behavior accordingly before the policy implementation, this assumption could be violated.

Assumption 3 (Treatment Effect Heterogeneity): The impact of the LECC on firm performance and R&D activity is potentially heterogeneous across firms funded in different time periods because it can be affected by different market conditions (e.g. financial crisis, the change in patent regulations, etc.). The LECC effect on R&D collaboration is also heterogeneous across different cohorts treated in different funding waves. Firms treated in later waves may have a longer time to accumulate cooperation partners beforehand, which can reduce the number of potential collaborations arising after the funding. However, the LECC impact on firm performance does not vary across cohorts since all 3 cohorts are treated equally by getting the same amount of funding over the same time length.

1.6 Results

1.6.1 Main Results

1.6.1.1 Direct effect: The impact of the LECC on funded firms

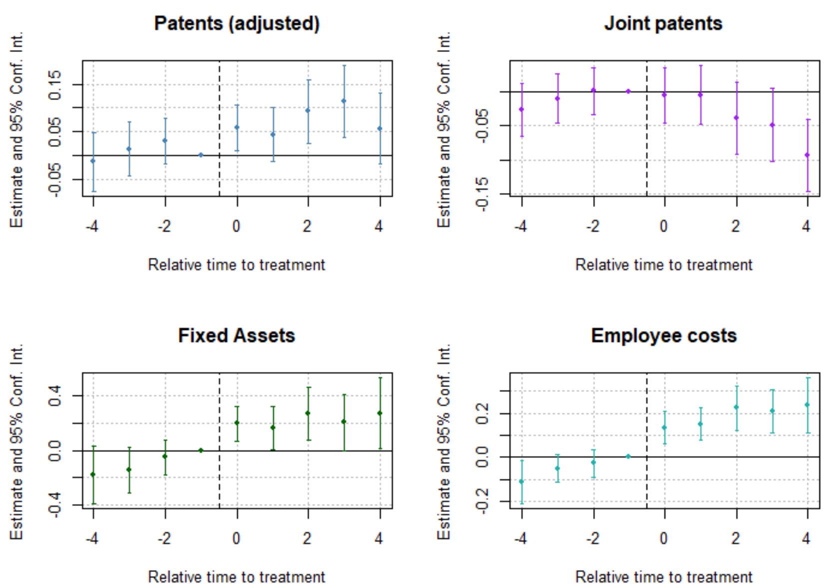
To bring the discussed methodology to the data, we firstly concentrate on the analysis of the direct effect of the LECC by using event study design and looking at the cohort effects separately for three cohorts funded in 2008, 2010, and 2013 (see Appendix A.4). Afterward, we pool all cohorts together to determine the total treatment effect of the LECC policy.

The results of the Sun and Abraham (2021) TWFE model show an interesting picture. In Table 1.3, we see that the LECC funding is positively correlated with the number of quality-adjusted patents filled by funded firms from the start of the funding, which might still be the outcome of firm innovation before the LECC policy. In contrast, when we look at the number of joint patents (patent applications filled in collaboration with 2 or more firms), we see rather a negative impact. This can be explained by the fact that first of all patent collaborations require time (up to 15 years) to be visible (patent application to be filed) and the trust to be established from long-term informal relationships, which do not exist yet. Secondly, the negative impact can be related to the fact, that the cluster policy-induced collaborations with SMEs, which rarely apply for a patent, and hence lead to more non-patent collaborations taking place. Additionally, the findings by Hagedoorn et al. (2003)

indicate that a higher number of inter-firm R&D alliances does not necessarily lead to an increase in joint patenting activities with those partners. In general, joint patenting presents greater challenges in terms of intellectual property rights (IPR), leading R&D managers to typically favor individual patents to mitigate potential IPR conflicts (Briggs and Wade 2014).

Regarding firm-level outcomes, we observe a positive and statistically significant trend for fixed assets and employment costs starting from the period of LECC funding. This trend is depicted in Figure 1.2 and detailed in Table 1.3, where fixed assets exhibit an increase from 19% to 27% until the last relative period. In parallel, the LECC leads to an increase in employment costs by up to 24% for funded firms. However, it is important to note that on average, the majority of firms in our dataset are SMEs, which may account for the substantial effect size observed in firm-level outcomes.

Figure 1.2: Event study: Staggered treatment for funded firms



The average treatment effect on treated (ATT) remains positive and significant for quality-adjusted patents and fixed assets and employment costs leading to around 7% increase for innovation outcomes and from 20% to 23% increase for respective firm-level outcomes. However, ATT on collaboration outcomes (joint patents) is negative. We also estimate cohort-specific average treatment effect on treated (CATT). For all 3 cohorts, the LECC is positively correlated with the capital and labor of firms (see Table 1.3).

Table 1.3: Sun&Abraham TWFE estimate for all cohorts

Dependent Variables:	Patent (adjusted)	Joint patents	Fixed assets	Employee costs
Model:	(1)	(2)	(3)	(4)
Variables				
Year = -4	-0.0141	-0.0270	-0.1782**	-0.1124**
	(0.0299)	(0.0202)	(0.0891)	(0.0472)
Year = -3	0.0125	-0.0101	-0.1447	-0.0502
	(0.0312)	(0.0215)	(0.0903)	(0.0360)
Year = -2	0.0303	0.0010	-0.0495	-0.0270
	(0.0320)	(0.0216)	(0.0921)	(0.0415)
Year = 0	0.0578*	-0.0053	0.1973**	0.1351***
	(0.0305)	(0.0217)	(0.0871)	(0.0369)
Year = 1	0.0434	-0.0045	0.1663**	0.1503***
	(0.0298)	(0.0220)	(0.0789)	(0.0315)
Year = 2	0.0919***	-0.0384	0.2698***	0.2250***
	(0.0320)	(0.0237)	(0.0875)	(0.0401)
Year = 3	0.1133***	-0.0488**	0.2047**	0.2088***
	(0.0337)	(0.0231)	(0.0881)	(0.0366)
Year = 4	0.0564*	-0.0940***	0.2730***	0.2372***
	(0.0308)	(0.0200)	(0.0956)	(0.0419)
ATT	0.0678**	-0.0515**	0.2342**	0.2017***
	(0.0284)	(0.0211)	(0.0952)	(0.0465)
CATT (cohort = 2008)	0.0730	-0.0273	0.7193***	0.2531***
	(0.0745)	(0.0450)	(0.2485)	(0.0689)
CATT (cohort = 2010)	0.0573*	-0.0346	0.0598	0.1661***
	(0.0308)	(0.0237)	(0.1000)	(0.0525)
CATT (cohort = 2012)	0.1085*	-0.1753***	0.1978	0.2234**
	(0.0627)	(0.0649)	(0.1387)	(0.0932)
Fixed-effects				
Firm	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	15,654	15,654	15,611	6,592
R^2	0.81800	0.61432	0.84057	0.93237
Within R^2	0.00392	0.01254	0.00729	0.02403

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

1.6.1.2 Spillover effect: The impact of the LECC on non-funded neighbor firms

In the second part of the analysis, we observe the spillover effects of the policy in dynamic TWFE setup a la Sun and Abraham (2021) for non-funded neighbors within 2 geographic rings (0-5 km; 5-10 km) separately. In the first geographic ring (0-5km), we see a significant positive impact on fixed assets appearing from the second relative period which reaches up to 4% in the last relative period. Nonetheless, the effect is rather small across the board. In contrast, we observe a negative trend after the policy for collaboration (joint patents) variable, which becomes significant from the second relative period onward for joint patents (see Figure 1.3 and Table 1.4). No significant impact of the LECC is found for the number of quality-adjusted patents and employment costs.

Figure 1.3: Dynamic spillover effects of the LECC for non-funded neighbor firms within 0-5 km

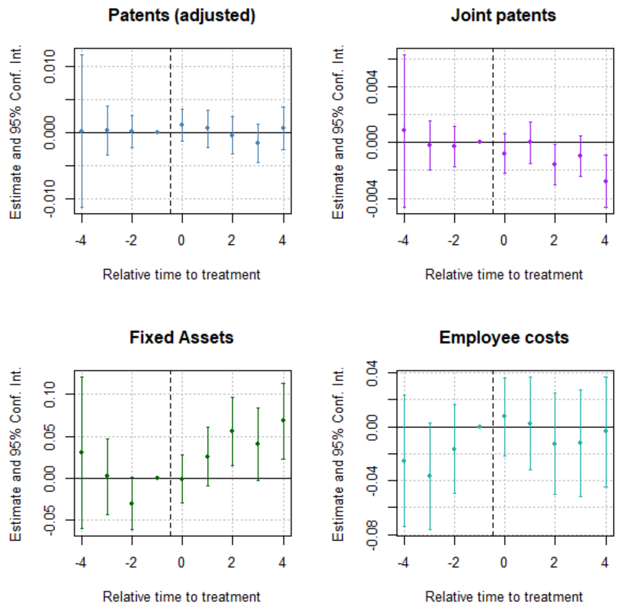


Table 1.4: Dynamic spillover effects of the LECC for non-funded neighbor firms within 0-5 km

Dependent Variables:	Patent (adjusted)	Joint patents	Fixed assets	Employee costs
Model:	(1)	(2)	(3)	(4)
Variables				
Year = -4	0.0026	0.0009	0.0466	-0.0142
	(0.0049)	(0.0023)	(0.0350)	(0.0220)
Year = -3	-1.58×10^{-5}	-0.0002	0.0103	-0.0368*
	(0.0019)	(0.0009)	(0.0256)	(0.0222)
Year = -2	8.74×10^{-5}	-0.0001	-0.0188	-0.0157
	(0.0015)	(0.0008)	(0.0221)	(0.0199)
Year = 0	0.0011	-0.0005	-0.0060	-0.0019
	(0.0014)	(0.0008)	(0.0195)	(0.0174)
Year = 1	0.0008	7.58×10^{-5}	0.0139	0.0065
	(0.0014)	(0.0008)	(0.0188)	(0.0178)
Year = 2	-0.0004	-0.0014*	0.0326*	-0.0104
	(0.0014)	(0.0008)	(0.0185)	(0.0174)
Year = 3	-0.0013	-0.0008	0.0105	-0.0170
	(0.0014)	(0.0007)	(0.0186)	(0.0178)
Year = 4	0.0019	-0.0014**	0.0425***	-0.0007
	(0.0013)	(0.0007)	(0.0164)	(0.0160)
ATT	0.0009	-0.0010	0.0273	-0.0033
	(0.0013)	(0.0007)	(0.0179)	(0.0180)
Fixed-effects				
Firm	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	1,352,023	1,352,023	1,351,840	161,222
R^2	0.69075	0.55170	0.87388	0.94579
Within R^2	8.6×10^{-6}	1.22×10^{-5}	2.3×10^{-5}	7.19×10^{-5}

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The TWFE regression results for the second cohort (neighbors within 5-10 km) show that in most cases, the LECC had no significant impact (see Figure 1.4 and Table 1.5).

Figure 1.4: Dynamic spillover effects of the LECC for non-funded neighbor firms within 5-10 km

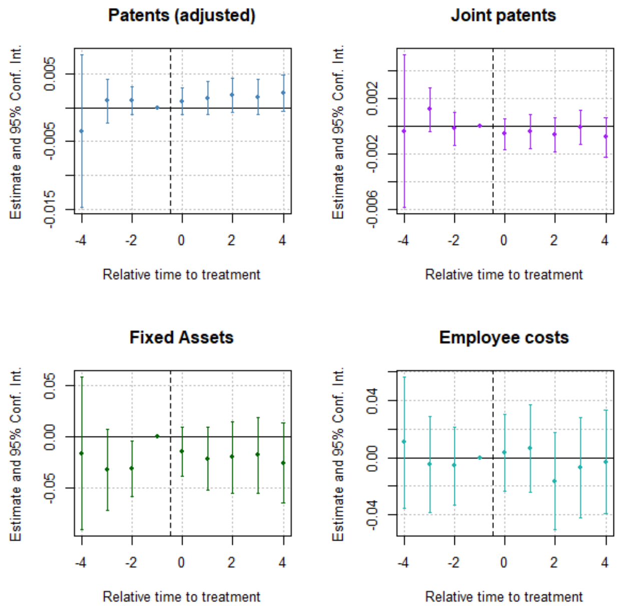


Table 1.5: Dynamic spillover effects of the LECC for non-funded neighbor firms within 5-10 km

Dependent Variables: Model:	Patent (adjusted) (1)	Joint patents (2)	Fixed assets (3)	Employee costs (4)
Variables				
Year = -4	-0.0018	0.0005	0.0117	0.0244
	(0.0044)	(0.0022)	(0.0292)	(0.0201)
Year = -3	2.12×10^{-5}	0.0012	-0.0346	0.0045
	(0.0017)	(0.0008)	(0.0225)	(0.0191)
Year = -2	0.0002	1.41×10^{-5}	-0.0335*	0.0028
	(0.0013)	(0.0007)	(0.0191)	(0.0175)
Year = 0	0.0003	-0.0005	-0.0174	0.0072
	(0.0013)	(0.0006)	(0.0167)	(0.0160)
Year = 1	0.0006	-0.0003	-0.0213	0.0204
	(0.0012)	(0.0006)	(0.0163)	(0.0166)
Year = 2	0.0008	-0.0005	-0.0156	-0.0048
	(0.0012)	(0.0006)	(0.0161)	(0.0166)
Year = 3	0.0008	6.03×10^{-5}	-0.0115	0.0002
	(0.0012)	(0.0006)	(0.0161)	(0.0165)
Year = 4	0.0019*	-4.49×10^{-5}	-0.0276*	0.0080
	(0.0011)	(0.0005)	(0.0141)	(0.0151)
ATT	0.0013	-0.0002	-0.0219	0.0068
	(0.0011)	(0.0005)	(0.0153)	(0.0163)
Fixed-effects				
Firm	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	1,480,529	1,480,529	1,480,373	175,493
R^2	0.68696	0.54223	0.87297	0.94174
Within R^2	5.79×10^{-6}	5.28×10^{-6}	7.05×10^{-6}	5.01×10^{-5}

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

1.6.2 Treatment Heterogeneity

The main results show that the LECC policy significantly increased the firm-level and innovation outcomes of funded firms, however, modest spillover effects were only present for the fixed assets of close-located non-funded neighbors (within 5 km). In this section, we investigate treatment heterogeneity in the direct effect of different types of firms and cluster members in order to identify more and less responsive groups.

1.6.2.1 SMEs vs large firms

The treatment effect may vary depending on the firm type. Particularly, the innovation process varies across firm size (Deschryvere 2014; Freel 2000; Rogers 2004). Besides, large firms are more able to afford to own a patent and start a patent-based collaboration in contrast to SMEs, which according to Hughes and Mina (2010) rarely file for patents due to their expensive nature. This implies that there might be almost no impact of the LECC on the innovative and collaborative outcomes of SMEs. Simultaneously, since large firms usually own a considerable amount of capital and labor, additional funding through LECC might be still too small (relative to generated revenue) to create any impact on firm-level outcomes of large firms. The same might work reversed for SMEs, which usually own less capital and labor, and additional funding might be a trigger to accumulate and hire more. To test the aforementioned, we identify SMEs and large firms by using the median total assets threshold, where a firm below or equal to the median of total assets is defined as an SME, otherwise, it is considered as a large firm.

Figure 1.5: Direct impact of the LECC policy on funded SMEs and large firms

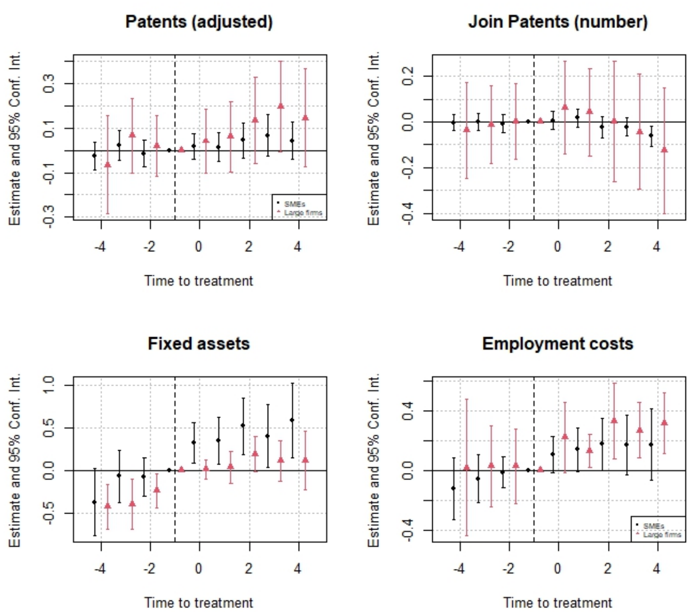


Figure 1.5 shows that the direct response of various outcomes to the LECC funding is different for SMEs and large firms. There is a positive trend for innovation outcomes for both SMEs and large firms, however, as anticipated, it is only significant for large firms. At the same time, we observe from Table 1.6, that SMEs experience significant growth of 35% for

fixed assets and 14 % for employment costs after the LECC implementation, which increases by around 1,5 times until the last relative period for the fixed assets and 1,2 times for the employment costs. Further, for SMEs, the ATT (average treatment effect on treated) estimate shows a positive significant correlation only with the firm-level outcomes (i.e. fixed assets, employment costs) while for CATT (cohort-specific average treatment effect on treated) the results become insignificant.

Table 1.6: Sun&Abraham TWFE estimate for all cohorts of funded SMEs and large firms

	SMEs				Large firms			
Dependent Variables:	Patent (adjusted)	Joint patents	Fixed assets	Employee costs	Patent (adjusted)	Joint patents	Fixed assets	Employee costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Year = -4	-0.0278	-0.0036	-0.3727**	-0.1213	-0.0655	-0.0367	-0.4253***	0.0174
	(0.0318)	(0.0167)	(0.1488)	(0.0890)	(0.1017)	(0.1005)	(0.1158)	(0.2265)
Year = -3	0.0214	0.0004	-0.0677	-0.0556	0.0648	-0.0127	-0.3951***	0.0281
	(0.0344)	(0.0209)	(0.1519)	(0.0849)	(0.0941)	(0.0975)	(0.1511)	(0.1378)
Year = -2	-0.0142	-0.0068	-0.0736	-0.0142	0.0191	0.0040	-0.2407**	0.0270
	(0.0366)	(0.0214)	(0.1592)	(0.0768)	(0.0986)	(0.1029)	(0.1215)	(0.1353)
Year = 0	0.0177	0.0073	0.3198**	0.1065	0.0402	0.0647	0.0130	0.2220*
	(0.0344)	(0.0212)	(0.1560)	(0.0728)	(0.0954)	(0.1081)	(0.0892)	(0.1220)
Year = 1	0.0140	0.0182	0.3512***	0.1382**	0.0598	0.0426	0.0392	0.1278**
	(0.0338)	(0.0210)	(0.1314)	(0.0663)	(0.0842)	(0.0918)	(0.1064)	(0.0539)
Year = 2	0.0436	-0.0229	0.5188***	0.1737**	0.1338	0.0017	0.1904*	0.3328***
	(0.0381)	(0.0218)	(0.1499)	(0.0785)	(0.0946)	(0.1110)	(0.1033)	(0.1221)
Year = 3	0.0672	-0.0213	0.4052***	0.1704**	0.1976**	-0.0435	0.1152	0.2677***
	(0.0411)	(0.0206)	(0.1526)	(0.0809)	(0.0920)	(0.0992)	(0.0973)	(0.0858)
Year = 4	0.0431	-0.0622***	0.5894***	0.1730*	0.1450	-0.1265	0.1188	0.3171***
	(0.0359)	(0.0198)	(0.1586)	(0.0905)	(0.0882)	(0.0968)	(0.1208)	(0.0728)
ATT	0.0379	-0.0270	0.4728***	0.1571*	0.1221	-0.0403	0.1005	0.2688***
	(0.0328)	(0.0177)	(0.1647)	(0.0881)	(0.0853)	(0.1091)	(0.1127)	(0.0885)
Cohort = 2008	0.0876	-0.0133	1.253***	0.2749**	0.0706	0.0610	0.2855	0.2339**
	(0.1069)	(0.0604)	(0.4527)	(0.1288)	(0.1444)	(0.1359)	(0.1879)	(0.1118)
Cohort = 2010	0.0229	-0.0295*	0.2651	0.1203	0.1610	-0.1552	0.0207	0.1828**
	(0.0311)	(0.0168)	(0.1698)	(0.1100)	(0.1230)	(0.1672)	(0.1206)	(0.0875)
Cohort = 2012	0.0325	-0.0404	0.1869	-0.0488	0.1121	0.0247	-0.0038	0.5231
	(0.0744)	(0.0344)	(0.1646)	(0.1282)	(0.1616)	(0.1769)	(0.1833)	(0.3854)
Fixed-effects								
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Observations	7,437	7,437	7,430	1,386	3,276	2,162	3,276	2,852
R^2	0.63813	0.34488	0.71222	0.94625	0.78255	0.56969	0.79500	0.82030
Within R^2	0.00417	0.01133	0.01554	0.05226	0.01136	0.00749	0.02411	0.02415

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The analysis of spillover effects on neighboring SMEs and large firms is provided in Appendix A.5.2.

1.6.2.2 Funded vs non-funded cluster members

The impact of LECC can vary between funded cluster members (those who received direct funding) and non-funded cluster members (those who are only cluster partners), a distinction that is observable in our cluster data. In general, we find in both cases a positive significant direct impact of the LECC policy on the innovation and firm-level outcomes (see Table 1.7 and Figure 1.6). However, in the case of the presence of direct funding (funded cluster members), the effect is larger. For instance, after the launch of the LECC policy, the funded cluster members increased their fixed assets up to 43% and employment costs up to 15% until the end of the observed period, while the non-funded cluster members had only up to 28% and 20% growth in parallel.

Figure 1.6: Direct impact of the LECC policy on funded & non-funded cluster members

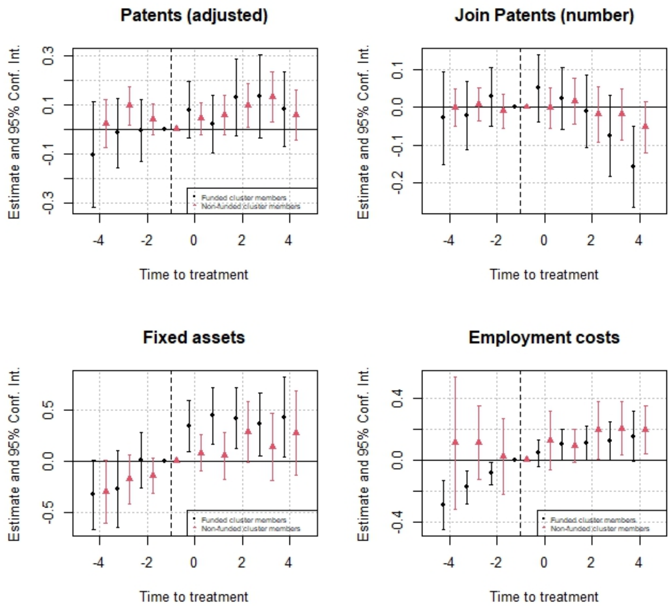


Table 1.7: Sun&Abraham TWFE estimate for all cohorts of funded & non-funded cluster members

	Funded Cluster Members				Non-Funded Cluster Members			
Dependent Variables:	Patent (adjusted)	Joint patents	Fixed assets	Employee costs	Patent (adjusted)	Joint patents	Fixed assets	Employee costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Year = -4	-0.1032	-0.0287	-0.3260**	-0.2928***	0.0219	-0.0002	-0.2936**	0.1123
	(0.1007)	(0.0526)	(0.1310)	(0.0719)	(0.0459)	(0.0255)	(0.1181)	(0.2185)
Year = -3	-0.0152	-0.0218	-0.2691	-0.1768***	0.0939**	0.0075	-0.1752	0.1129
	(0.0760)	(0.0496)	(0.1841)	(0.0653)	(0.0429)	(0.0281)	(0.1300)	(0.1264)
Year = -2	-0.0061	0.0273	0.0133	-0.0860	0.0408	-0.0107	-0.1365	0.0214
	(0.0751)	(0.0464)	(0.1626)	(0.0614)	(0.0443)	(0.0285)	(0.1354)	(0.1320)
Year = 0	0.0787	0.0505	0.3454**	0.0456	0.0443	-0.0021	0.0820	0.1283
	(0.0721)	(0.0507)	(0.1447)	(0.0557)	(0.0423)	(0.0283)	(0.1262)	(0.1027)
Year = 1	0.0218	0.0218	0.4426***	0.1004*	0.0581	0.0164	0.0564	0.0930
	(0.0666)	(0.0459)	(0.1313)	(0.0515)	(0.0402)	(0.0296)	(0.1144)	(0.0575)
Year = 2	0.1297*	-0.0111	0.4190***	0.1105*	0.0976**	-0.0187	0.2862**	0.1943**
	(0.0756)	(0.0498)	(0.1401)	(0.0585)	(0.0437)	(0.0307)	(0.1336)	(0.0962)
Year = 3	0.1333*	-0.0747	0.3632***	0.1225**	0.1310***	-0.0191	0.1437	0.2046**
	(0.0779)	(0.0492)	(0.1350)	(0.0584)	(0.0450)	(0.0291)	(0.1379)	(0.0869)
Year = 4	0.0812	-0.1582***	0.4314***	0.1544**	0.0586	-0.0533**	0.2781*	0.1962***
	(0.0645)	(0.0436)	(0.1612)	(0.0611)	(0.0411)	(0.0264)	(0.1439)	(0.0661)
ATT	0.0860	-0.0669*	0.4088***	0.1195**	0.0722*	-0.0240	0.1933	0.1697**
	(0.0621)	(0.0406)	(0.1548)	(0.0576)	(0.0392)	(0.0277)	(0.1445)	(0.0703)
CATT (cohort = 2008)	0.0220	-0.0335	0.7324**	0.2198**	0.1344	0.0318	0.7931	0.1809*
	(0.1186)	(0.0700)	(0.3040)	(0.0980)	(0.1145)	(0.0662)	(0.4852)	(0.1077)
CATT (cohort = 2010)	0.0981*	-0.0509	0.2917	0.0499	0.0375	-0.0340	0.0615	0.0755
	(0.0574)	(0.0498)	(0.1967)	(0.0780)	(0.0432)	(0.0321)	(0.1406)	(0.0637)
CATT (cohort = 2012)	0.1978	-0.1947*	0.0189	0.0664	0.1813*	-0.0525	0.0369	0.4943
	(0.2082)	(0.1115)	(0.1345)	(0.1147)	(0.1021)	(0.0570)	(0.1442)	(0.4171)
Fixed-effects								
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Observations	4,998	4,998	4,998	2,634	8,663	8,663	8,649	3,160
R^2	0.82062	0.75964	0.85532	0.93749	0.77177	0.55473	0.82360	0.91351
Within R^2	0.01131	0.02116	0.01589	0.05021	0.00616	0.00528	0.00704	0.01421

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Simultaneously, the LECC effects on innovation outcomes look quite similar for both funded and non-funded cluster members resulting in around a 13% increase in quality-adjusted number of patents in both cases. This suggests that while direct funding does have an additional impact, it is the collaborative institutional structure of cluster policy that primarily drives long-term impacts. Additional heterogeneity analysis concerning the firm size structure of both funded and non-funded cluster members is elaborated upon in the Appendix A.5.1.

1.7 Conclusions

In this study, we investigate the impact of the German cluster policy, particularly the Leading-edge cluster competition, on firm performance, innovative activity, and R&D collaboration. Utilizing firm-level, patent, and cluster data spanning from 2000 to 2016, our analysis aims to assess how the cluster policy influences the outcomes of both funded firms (direct effect) and non-funded neighbor firms (spillover effect). We formulate two hypotheses to examine whether the implementation of cluster policy has significantly enhanced the performance, R&D output, and collaboration of successful cluster firms compared to non-funded firms in Germany. Additionally, we explore whether these positive effects extend to non-funded neighbor firms through localized spillovers or if they instead experience crowding out from competitive cluster neighbors.

In general, the results show that the LECC policy has a positive impact on the performance of both funded and non-funded neighbor firms. Hypothesis 1 assuming supported firms by cluster policy to have a better performance, more patents, and collaboration than non-supported firms partially holds. The results of our analysis show a positive significant impact on firm-level and innovation outcomes, but no impact on collaboration outcomes. The latter result is intuitive and expected since the impact on collaboration (joint-patents) usually takes a long time to occur and in general, the co-patenting process is quite time-consuming and requires many years. At the same time, the process of fostering collaboration requires a long time too, since it includes the component of trust between actors, which takes a longer time to build even if the actors belong to the same cluster.

In the heterogeneity analysis, we observe that the policy has a significantly positive impact, particularly on firm-level outcomes for SMEs. Conversely, for large firms, the effect is predominantly pronounced on innovation outcomes. Interestingly, when comparing funded and non-funded cluster members, the positive significant impact is evident for both groups. This suggests that the efficiency of mechanisms induced by the Leading-edge cluster competition (LECC) plays a more substantial role than the actual funding itself in driving these positive effects.

According to Hypothesis 2, the cluster policy can positively influence the outcomes of non-funded neighbor firms in targeted industries as well via the spillover channel. This argument particularly finds support for the neighboring firms within a 5 km distance, which experience small positive spillover effects on the firm outcome. Hypothesis 2 is rejected for the neighbor firms above 5km distance. These results provide the intuition that first of all, spillovers are localized and secondly, they might be very regional and take a longer time to occur based on the arguments mentioned above.

These results are in line with the previous research on the impact of cluster policy on firm performance (Uyarra and Ramlogan 2016). However, in contrast to (Cantner 2013; Rothgang et al. 2017a), we find no impact on collaboration, which can be the reason for some natural limitations that our study has. One explanation of not detected impact in contrast to the aforementioned studies is that the latter ones used surveys to identify linkages between actors in contrast to our research using joint patents as a measure of collaboration. As long as patent applications are quite expensive, hard to file, and sometimes even not essential depending on the industry specialization, it is clear that not all collaborations in the network necessarily lead to patenting. This idea has been discussed in a few studies (Cantner and Graf 2006; Graf and Henning 2008). They state that not all inventions can be patented, which makes the use of joint patents as a collaboration measure quite problematic. Graf and Henning (2008) also add that the problem is even more severe when it comes to the collaboration analyses of clusters specialized in different industries with dissimilar patenting tendencies. In other words, the analyses will automatically show a smaller number of joint patents and consequently, weak or no collaboration for the clusters specialized in industries with a low proportion of patents. Furthermore, research by Hagedoorn et al. (2003) and Briggs and Wade (2014) suggests that inter-firm R&D collaborations often do not lead to joint patents due to conflicts arising from intellectual property rights (IPR).

While survey-based methods remain valuable, they are susceptible to bias, notably due to the potential presence of the Hawthorne effect. This phenomenon suggests that participants may alter their behavior in response to being observed, potentially inflating the reported levels of collaboration. In contrast, patent collaborations offer a more tangible and objective measure, providing a "hard" metric of actual collaborative activities, thus offering a more reliable assessment of the impact of funding initiatives on collaboration.

Another plausible explanation for the obtained results could be the relatively short period of analysis (2000-2016). As discussed earlier, the Leading-edge cluster competition (LECC) was implemented in three rounds (2008, 2010, 2012), supporting five clusters in each round for approximately five years. Since our analysis covers only the period from 2000 to 2016 due to the unavailability of more recent data, it is evident that especially for clusters supported in later cohorts, we may not have sufficient after-treatment periods to observe post-treatment effects accurately. Moreover, according to Rothgang et al. (2017a), cluster policy impacts on supported actors usually do not occur during the period of policy support, rather they appear after a longer period and might require decades to be completely visible. A similar conclusion was reached by Cantner et al. (2019), who claim that cluster policies do not induce an immediate effect but rather create structures to maintain constant success after the funding period.

In conclusion, our analysis reveals that while funding clusters do indeed have a positive impact on funded firms, the cluster policy approach, particularly the "the winners" concept exemplified by the Leading-edge cluster competition (LECC), tends to benefit SMEs within the cluster more significantly. This is primarily because large firms within the LECC were already performing well even before the policy implementation due to their excellence. The impact of the policy might have been more substantial if it had focused on supporting clusters that are not already leading edge. Moreover, funding excellence clusters can result in the over-concentration of highly competitive firms in a particular area, potentially leading to the crowding out of less competitive neighbors and exacerbating regional disparities. This observation is consistent with studies by Lehmann and Menter (2017) and Audretsch et al. (2018). Therefore, future policy implications may be more effective if they prioritize stimulating new and less-developed cluster initiatives, thereby fostering broader economic development and reducing regional disparities.

Nonetheless, given that policy impacts typically manifest in the long run, it is crucial to conduct ex-post evaluations after the end of policy support. Therefore, we highly advocate for policymakers to base their decisions regarding future cluster initiatives not only on concurrent evaluations but also consider evaluations conducted 5-10 years after the end of policy support. Such assessments provide valuable insights into the lasting effects and overall success of cluster policies, enabling policymakers to make informed decisions for the future.

Future research endeavors should engage in more extensive analyses of cluster policies and strive to address the limitations observed in the current study over the long term. Exploring the impact of these policies on the public sector, including universities and research institutes, would be particularly insightful. Additionally, investigating inventor-based patent applications as a measure of collaboration could offer more precise information regarding the location and connection of patents to specific firms. Finally, delving into non-patented collaborations, which are more common for SMEs and may involve surveys or other methods, presents a rich and challenging avenue for further studies. By exploring these areas, researchers can deepen our understanding of the dynamics and outcomes of cluster policies, contributing to more informed policymaking in the future.

CHAPTER 2

Funding the Future: Evolution and Impact of R&D Funding in Germany

Chapter Abstract

This study examines R&D funding trends in Germany and focuses on assessing the direct impact of R&D funding on both firm performance and innovation outcomes of SMEs and large firms. Leveraging various datasets including Bureau van Dijk's Orbis database, the EPO Worldwide Patent Statistical Database (PATSTAT), and Förderkatalog data, I employ different empirical methods (OLS regressions, Propensity Score Matching, and event study approach) over the period of 1970 to 2020. While initial results suggest a positive influence of R&D funding on SME performance in terms of firm-level outcomes, there is no discernible effect on innovation outcomes for either SMEs or large firms. However, these outcomes are influenced by pre-existing trends, which persist even after sensitivity analysis and detrending, indicating that the effects might be potentially driven by other factors beyond R&D funding.

2.1 Introduction

Over the past decades, innovation has been recognized as the main driver of economic growth by prominent economists such as Romer (1990) and Solow (1957), which was followed by the growing effort to promote innovation worldwide through the provision of R&D subsidies. They are particularly important for commercial R&D projects. The latter ones offer large anticipated social benefits but insufficient expected returns to private investors due to the incomplete appropriability of innovation and knowledge spillovers thus reducing private incentives to invest in R&D. This is the classical market failure argument developed by Arrow (1962). Another justification for public R&D subsidies is connected to financial market failure, which states that due to the high cost of external capital (particularly for SMEs), some innovations will never be developed (Hall 2002a; Hall 2002b). Hence, R&D subsidies induce firms to implement R&D projects which would have not been profitable without a subsidy and thus stimulate private R&D expenditures (Jaffe 2002; Wallsten 2000). But are R&D subsidies effective in stimulating R&D investment? Despite huge efforts during 60 years of research, empirical evidence is ambiguous and there is no final answer to the question of how public R&D funding impacts recipients' performance and innovation outcomes. This inconsistency of results arises due to the differences in the selected variables, empirical methods, and the selected population (e.g. industries, countries, time periods) of these studies.

While existing studies primarily focus on the input (Almus and Czarnitzki 2003; Clausen 2009; Hussinger 2008; Wallsten 2000) and output effects (Aschhoff 2009; Bronzini and Piselli 2016; Czarnitzki and Fier 2002; Czarnitzki and Hussinger 2004; Czarnitzki and Licht 2006; Kleine et al. 2022) of R&D subsidies, such as their influence on private R&D expenditures and innovation output, less attention has been given to their effects on other outcomes. Understanding the impact of R&D funding on outcome measures is crucial. As the primary goal of R&D funding is to address innovation-related externalities, given that firms may not invest adequately in this area independently, assessing whether this policy instrument effectively achieves its intended objectives is of paramount importance. A few studies have explored the impact of R&D funding on factors like firms' employment and sales, revealing a positive association (Hall et al. 2009; Lerner 1999). However, there is a lack of in-depth analysis regarding the impact of R&D funding on firm-level outcomes.

To address this gap, this study investigates the intricate relationship between R&D funding, innovative activity, and firm performance in Germany, with a specific focus on heterogeneous firm groups based on size (SMEs and large firms). Utilizing a comprehensive dataset spanning from 1970 to 2020, which includes firm-level information from DAFNE, patent data from PATSTAT, and funding records from Förderkatalog (a unique, comprehensive database containing detailed information on all funded R&D projects in Germany from 1970 up to 2020), I conduct rigorous empirical analysis to uncover the nuanced effects of R&D funding on funded firms' outcomes. To address the selection issue, I begin by applying Nearest Neighbor Propensity Score Matching. Subsequently, I utilize dynamic event study techniques, taking into account pre-trend violations and implementing necessary corrections through detrending. Additionally, I aim to provide insights into the evolving patterns of R&D funding within the German context, examining the targeted actors, industries, regions, and research types. This analysis will facilitate a deeper exploration of the various dimensions of heterogeneity in R&D funding, guiding us to investigate whether factors such as firm size, previous patenting experience, or industry/regional concentration influenced the likelihood of receiving funding.

This paper contributes to the existing literature in several ways. Firstly, it leverages the rich Förderkatalog database to provide a comprehensive overview of all funding initiatives in Germany, offering insights into funding evolution patterns based on different characteristics, unlike previous studies focusing on a specific funding initiative (Almus and Czarnitzki 2003; Cantner 2013; Crass et al. 2017; Engel et al. 2013; Falck et al. 2010; Rothgang et al. 2017a; Töpfer et al. 2017). Secondly, by utilizing this extensive dataset, I conduct fine-grained firm-level analysis, examining the effects of R&D funding not only on innovation outcomes but also on firm-level indicators such as capital and labor. While previous studies by Hall et al. (2009) and Lerner (1999) have touched on these aspects, comprehensive analysis observing firm-level dynamics is missing. This analysis yields a significant implication for policymakers concerning the effectiveness of R&D funding and its potential to address shortages of labor and capital within firms. Finally, with this paper, I introduce methodological novelties by implementing dynamic effects analysis within a comprehensive event study framework, accounting for pre-trend violations, and proposing corrective approaches. To the best of my knowledge, this aspect is lacking in the existing innovation literature, which has primarily focused on static analyses over short-term periods.

The structure of the paper is as follows: Section 2.2 presents a systematic review of the empirical literature, discussing the heterogeneous effects of funding on different outcomes, findings across different countries, and the specific case of Germany. Section 2.3 provides an overview of the data and descriptive analysis on R&D funding evolution in Germany. I detail the empirical methodologies used in Section 2.4, while Section 2.5 discusses the various sets of results. Finally, Section 3.6 concludes with remarks summarizing findings.

2.2 Literature Review

2.2.1 Analyzing R&D Policy Effects

Most of the research concluded that R&D subsidies indeed trigger private R&D expenditures (Almus and Czarnitzki 2003; Clausen 2009; Czarnitzki and Fier 2002; Czarnitzki and Hussinger 2017; Czarnitzki and Licht 2006; Hussinger 2008; Wallsten 2000). However, besides input additionality effects, R&D subsidies can influence firm output as well. For instance, the study by Bérubé and Mohnen (2009) asserts that there is a positive correlation between a firm's R&D grant acquisition and being more innovative. This finding is further supported by various studies claiming that R&D subsidies positively impact a firm's innovation output (Aschhoff 2009; Czarnitzki and Fier 2002; Czarnitzki and Hussinger 2017; Czarnitzki and Licht 2006; Czarnitzki and Lopes-Bento 2011; Kleine et al. 2022). Other scholars (Bronzini and Piselli 2016; Czarnitzki and Hussinger 2017; Czarnitzki and Licht 2006) also add that R&D subsidies positively impact a firm's patenting activities.

Regarding firm performance, the evaluation of the SIBR program in the USA shows that participant firms experienced an increase in sales and employment relative to the control group (Lerner 1999). Similar findings were reported in the studies of Falck et al. (2010) and Hall et al. (2009) who find a positive impact of R&D subsidies on employment growth. R&D subsidies also induce positive impacts on collaboration between firms (Czarnitzki et al. 2007;

Kleine et al. 2022). However, the question still remains which are the drivers of these positive impacts among the actors, and in which kind of projects/industries/firm-types R&D funding will bring higher returns?

It has been proven that small and medium-sized enterprises (SMEs) play an irreplaceable role in innovation and economic growth (Audretsch et al. 2006; Howell 2017) due to their particular effectiveness in producing radical innovations (Hottenrott and Lopes-Bento 2016). Although, SMEs are usually facing a shortage of financial and human capital because of their size and hence are more targeted by policymakers for R&D support than large firms. The findings by Busom (2000) verify that SMEs indeed have a higher probability of getting R&D funding in contrast to large firms. Further studies reveal that R&D funding has a direct positive effect on innovation outcomes of SMEs by leading to an increased number of patents and new products (Czarnitzki and Delanote 2015; Czarnitzki and Hottenrott 2011; Howell 2017; Kleine et al. 2022; Lerner 1999). Simultaneously, R&D subsidies indirectly enhance the likelihood of future collaborations (Bianchini et al. 2019) and the availability of long-term debts (Meuleman and Maeseneire 2012). Subsidizing R&D of SMEs also leads to positive input effects by stimulating private R&D investments (Lach 2002). Moreover, Czarnitzki and Delanote (2012) assert that young innovative firms have higher growth rates for sales and employment in contrast to the other firms. Hence, in their further study Czarnitzki and Delanote (2015) suggest that R&D funding should be rather directed to young and independent SMEs in high-tech sectors. Contradicting results are found by Wallsten (2000) who studies R&D grant's effect on private R&D for firms involved in the Small Business Innovation Research (SBIR) program. He finds that R&D grants lead to crowding out of private R&D expenditures.

Not only the right choice of the subsidy recipient but also the program is important to reach maximum efficiency. According to Arrow (1962), private R&D is rarely invested in basic research because the latter has less commercial value and is far from the market. Following this idea, Clausen (2009) claims that 'far from the market' R&D projects positively influence the quantity and the quality of firm-level R&D and hence should be given preferential treatment by the policymakers. Akcigit et al. (2020) confirms that R&D subsidies targeting basic research and stimulating its collaboration with the firms lead to considerable welfare gains. The same does not hold true for applied research. Overfunding of the latter usually leads to dynamic misallocation in the economy (Akcigit et al. 2020).

The share of R&D subsidies might vary in different industries, which Svensson (1998) explains is influenced by the political (e.g. labor intensity explaining political power, the decline in the industry measured as the four-year change in value-added) and economic factors (e.g. technological level describing the ratio of scientists and engineers in total employees, returns to labor). He further emphasizes that political preferences are directed to the industries that have a low degree of concentration, low returns to capital, and experience industrial decline, while economic preferences are rather aimed at high-tech industries, which are more likely to stimulate positive externalities and welfare effects. Blanes and Busom (2004) verify the findings regarding the preferential funding of high-tech industries. The study also reveals that firms with previous experience in R&D have a higher probability to receive an R&D subsidy in contrast to non-R&D doers.

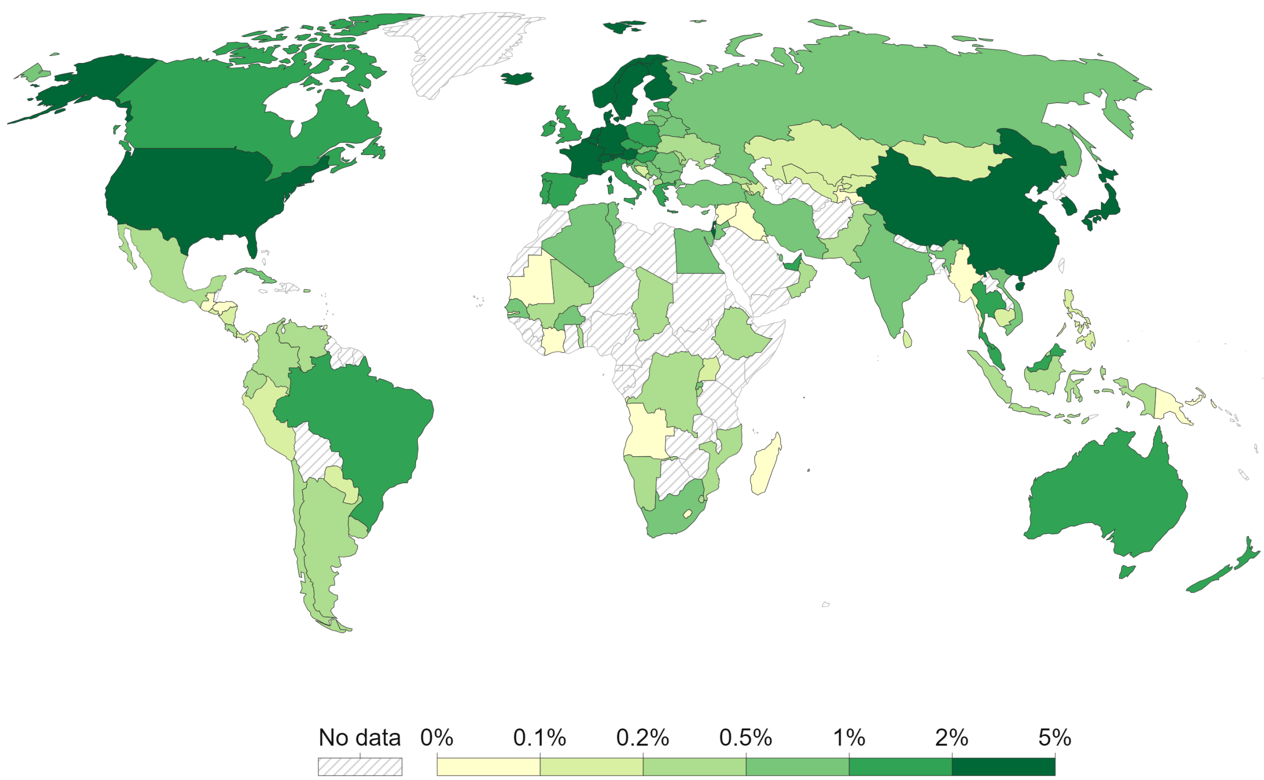
Finally, R&D funding decisions (both from the perspective of a firm and policy makers) can be influenced by the institutional and other differences between countries (David et al.

2000). More insight into the country-specific R&D funding and research will be given in the next sub-chapter.

2.2.2 Comparative Study of R&D Subsidies

Public funding of R&D is a well-established practice worldwide. According to Eurostat (2023), in 2021 the total government budget allocations for R&D (GBARD) across the EU were €111 393 million (0.77 % of GDP in contrast to 0.72 % of GDP in 2011) with Germany leading the EU countries with the GBARD reaching 1.12 % of GDP. Despite the slight increase, the GBARD of the EU was lower than the respective ratios of leaders such as Japan (1.51%), South Korea (1.32%), and Iceland (1.20%). However, the EU's GBARD ratio to GDP was still higher than in the US (0.71%).

Figure 2.1: R&D spending as a share of GDP, 2019



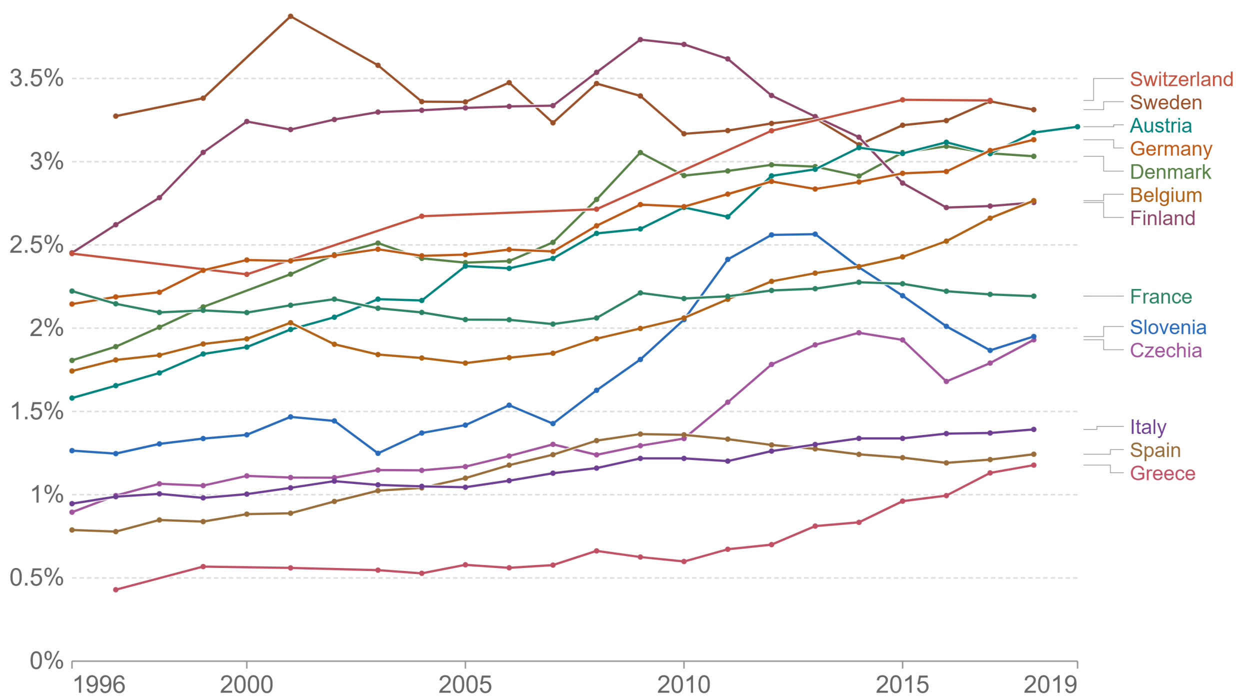
Source: UNESCO 2022
Note: Spending includes current and capital expenditures (public and private) on research (basic, applied, and experimental development).

In contrast, the pattern was different for gross domestic expenditure on R&D (GERD). Figure 2.1 shows that in 2018, the highest share of GERD to GDP was reported in Israel (4.9 % of GDP), South Korea (4.5 % of GDP), Japan (3.3 % of GDP), the US (2.8 % of GDP), and China (2.1 % of GDP) (UNESCO 2022). In the EU, the GERD ratio to GDP was 2.02 % over the period of 2011 to 2018, which slightly increased to 2.23 % in 2019. This is still lower than the targeted policy goal of the EU to increase total private and public R&D expenditure to at least 3% of the gross domestic product (GDP). This leaves the EU behind the R&D intensive counterparts such as Israel, South Korea, Japan, and the United States (Eurostat 2022).

However, when zooming into Europe for country-wise comparison in Figure 2.2, we observe that the highest share of GERD to GDP was performed by Switzerland (3.4 % of GDP) and Sweden (3.3 % of GDP), which positions them among the top 5 leaders worldwide by leaving Japan and the US behind. Further leaders in the EU for R&D expenditures are Austria (3.2 % of GDP) and Germany (3.2 % of GDP) which show a similar growing pattern from 1996 in contrast to Finland which was leading in the EU until 2012 and started a downward trend from 2009 (UNESCO 2022).

Regarding the heterogeneities concerning the source of funding, in the majority of the countries, the public sector support for private R&D is organized through the public agencies. For instance, there are 17 supranational, national, and regional R&D funding programs in Europe, while in the US, the R&D funding programs are conducted at the federal and state level (Zúñiga-Vicente et al. 2012).

Figure 2.2: R&D spending as a share of GDP



Source: UNESCO 2022

Note: Spending includes current and capital expenditures (public and private) on research (basic research, applied research, and experimental development).

The significant share of research performed up to this date targets mainly European countries and the US. The studies for the US were widely implemented until 2000 indicating mixed results both in favor of created additionality (Diamond 1999; Rosenberg 1976; Scott 1984) and substitution (Wallsten 2000) of private R&D. In contrast, European evaluations of

R&D policies became popular from early 2000. Most of the studies supported the hypothesis of both input and output additionality effects in France (Duguet 2004; Hassine and Mathieu 2020), Belgium (Aerts and Czarnitzki 2004), Spain (González and Pazó 2008; González et al. 2005), Italy (Bronzini and Piselli 2016), Austria (Falk 2004; Streicher et al. 2004) and Nordic countries (Clausen 2009; Hyytinen and Toivanen 2005; Lööf and Heshmati 2004; Sörensen et al. 2003).

2.2.3 R&D Funding in Germany

To increase national competitiveness and stimulate technological progress and long-term growth, from the 1950s Germany promotes R&D activities in the private sector. There are three policy instruments to promote innovation activities: publicly conducted research (government or university), public subsidy for business performed R&D, and fiscal incentives (Czarnitzki and Fier 2002). Among three policy instruments to foster innovation, the most important and preferred one in Germany is the public funding of business R&D projects performed by the private sector [1]. It allows flexible adjustment to new challenges, initiation of public-private collaborations, and enhances the quality due to the competitive ("pick the winner") approach by only supporting the firms with relevant innovative capabilities and financial resources. To be eligible for the R&D subsidy, firms have to apply with a targeted project proposal. This is followed by the peer review process, after which the grants are given and obliges applicants to contribute a minimum of 50% to the funded projects (Czarnitzki and Fier 2002). Since the reunification of Western and Eastern Germany in the 1990s, the federal government has expanded its "technology portfolio" and begun to more actively support "novel" technologies (e.g. biotechnology, microsystems, and chemical technology) and close-to-the-market projects, in their early stages of development within the corporate sector. Additionally, efforts to facilitate technology transfer between academia and industry through collaborative projects have intensified. As a result, the number of companies, particularly SMEs, receiving grants through direct project funding has notably risen since then (Czarnitzki et al. 2003).

Regarding the particular case of Germany, the majority of studies indicate input additionality effects and claim that firms that got R&D subsidy invest more into R&D in contrast to non-funded firms (Aerts and Schmidt 2008; Almus and Czarnitzki 2003; Aschhoff 2009; Czarnitzki and Fier 2002; Czarnitzki and Hussinger 2004; Czarnitzki and Hussinger 2017; Czarnitzki and Licht 2006; Czarnitzki and Lopes-Bento 2014; Czarnitzki and Toole 2007; Hud and Hussinger 2015; Hussinger 2008). Moreover, Kaiser (2002) provide evidence that collaborative firms in the service sector invest more in R&D than non-collaborative firms. Other scholars (Aschhoff 2009; Becker and Dietz 2004; Czarnitzki and Fier 2002; Czarnitzki and Hussinger 2017; Czarnitzki and Licht 2006; Czarnitzki and Lopes-Bento 2011) also claim that R&D subsidies led to the increase of innovation output of funded firms in Germany. Finally, a recent study by Koehler (2018) shows that R&D subsidy programs lead to higher welfare gains. To conclude, all the studies evaluating R&D subsidy effects in Germany concentrate on direct effect analysis by mostly tackling input and output additionality effects. However, less is known about the effects of R&D subsidies on other firm-level outcomes.

[1] Note: R&D tax credits were introduced in Germany in 2020 and have not gained widespread adoption yet.

2.3 Data

I employ a combination of three datasets to assess the impact of R&D funding on both firm performance and innovation output, focusing on German firms from 1970 to 2020.

The first dataset utilized is the DAFNE database, collected in September 2020. This extensive database covers approximately 2 million firms in Germany and includes detailed information on firm characteristics such as location, sector, financial data, legal entity details, corporate structures, and ownership. We leverage this dataset to derive firm-level metrics such as employment costs, fixed assets, sales, and other relevant indicators.

The second dataset used is the EPO Worldwide Patent Statistical Database (PATSTAT), developed by the European Patent Office (EPO) at the behest of a task force led by the Organisation for Economic Co-operation and Development (OECD). We obtained the data from the PATSTAT Spring Edition 2020, extracted in June 2020. This dataset provides comprehensive bibliographical and legal event patent data, including detailed information on applicant and inventor names and addresses, priority years, patent families, titles and abstracts of patent applications, citation links, and patent classifications by technology class. Specifically for Germany, the PATSTAT data comprises records for approximately 69,898 firms and around 4,609,894 patent applications. I use this data to identify and analyze patent counts, employing two types of measures: the simple number of patents and the quality-adjusted number of patents. The simple count refers to the number of patent applications submitted by each firm annually. For the quality-adjusted patent number, I follow the methodology of Boeing and Mueller (2016), which takes into account factors such as the number of technology classes, family size (the number of countries where a patent application was filed), and the number of forward citations received by the patent.

DAFNE and PATSTAT data are linked to the Förderkatalog database (extracted in October 2022), which is used to identify firms that got R&D funding. The latter database contains information on funding recipients, their location, funding, and project type, responsible authority, etc., which will be described in detail in the next sub-chapter.

Table 2.1: Descriptive statistics

	Funded (n=17211)	Non-funded (n=975671)	Overall (n=992882)
Average firm age	27.8 (31.2)	20.5 (23.0)	20.7 (23.2)
Firm size	15.6 (2.14)	13.4 (1.97)	13.4 (2.01)
Patent (simple)	20.0 (172)	5.72 (87.7)	9.49 (116)
Patent (adjusted)	52.2 (462)	14.5 (220)	24.5 (303)
Fixed assets	110 (1580)	6.56 (318)	8.98 (397)
Total assets	175 (2360)	10.8 (467)	14.7 (587)
Turnover	277 (2460)	23.8 (612)	31.5 (740)
Sales	266 (2460)	25.7 (485)	37.4 (721)
Labor costs	47 (393)	7.95 (341)	11.4 (346)

Notes: Fixed and total assets, turnover, sales, and labor costs values are in Mln.

Descriptive statistics are depicted in Table 2.1 and Table 2.2. Table 2.1 shows that R&D funded firms, on average, have been established for approximately 7.3 years longer than non-funded firms. There is a slight disparity in firm size, with funded firms being, on average, larger. Nevertheless, in terms of other outcomes, there is a significant divergence, with funded firms having higher levels of innovation (approximately 4 times more patents), greater size (on average, 16 times larger fixed and total assets), and hence, generating at least 10 times more turnover and sales, as well as allocating at least 5 times more resources to labor. Furthermore, both funded and non-funded firms are represented across various industries. Despite both being prevalent in sectors like Wholesale and Retail Trade, and Financial and Insurance Activities, funded firms show a higher presence in R&D-intensive industries such as Education, Professional, Scientific, and Technical activities, Human Health and Social Work Activities, and Information and Communication (see Table 2.2).

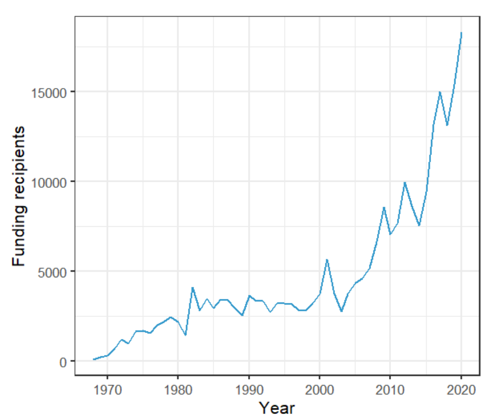
Table 2.2: Descriptive statistics on industry

Industry	Funded (n=17211)	Non-funded (n=975671)	Overall (n=992882)
Accommodation and Food Service Activities	782 (0.4%)	100147 (1.0%)	100929 (1.0%)
Administrative and Support Service Activities	846 (0.4%)	97963 (1.0%)	98809 (1.0%)
Arts, Entertainment and Recreation	603 (0.3%)	29229 (0.3%)	29832 (0.3%)
Construction of Buildings	1866 (1.0%)	355793 (3.6%)	357659 (3.6%)
Education	4569 (2.4%)	67483 (0.7%)	72052 (0.7%)
Financial and Insurance Activities	459 (0.2%)	120207 (1.2%)	120666 (1.2%)
Human Health and Social Work Activities	2475 (1.3%)	65061 (0.7%)	67536 (0.7%)
Information and Communication	2727 (1.4%)	66665 (0.7%)	69392 (0.7%)
Manufacturing	3013 (1.6%)	53810 (0.6%)	56823 (0.6%)
Other Service Activities	2138 (1.1%)	75732 (0.8%)	77870 (0.8%)
Professional, Scientific and Technical Activities	4255 (2.2%)	262423 (2.7%)	266678 (2.7%)
Public Administration and Defense: Compulsory Social Security	1370 (0.7%)	13664 (0.1%)	15034 (0.2%)
Real Estate Activities	3264 (1.7%)	793354 (8.1%)	796618 (8.0%)
Transportation and Storage	1551 (0.8%)	69116 (0.7%)	70667 (0.7%)
Water Supply	835 (0.4%)	16178 (0.2%)	17013 (0.2%)
Wholesale and Retail Trade	5474 (2.9%)	671482 (6.9%)	676956 (6.8%)
Activities of households as employers	0 (0%)	165 (0.0%)	165 (0.0%)

2.3.1 The Förderkatalog Database

All R&D funding projects in Germany are reported in the elaborate dataset called Förderkatalog (funding catalog). The Förderkatalog is a publicly available database, which is jointly initiated by the Federal Ministry of Education and Research (BMBF) and the Federal Ministry of Economics and Technology (BMWi) in Germany. It started in 1970 and is regularly updated. New projects are accepted 60 days after approval.

Figure 2.3: R&D funding recipients in Germany



The database contains data on 90,000 individual projects. The total budget represented in Förderkatalog comprised 44 bn €. It is apparent from Figure 2.3, that the number of funding recipients tremendously increased from 1970 to 2020 leading to 28013 firms in total.

The Förderkatalog database includes the following ministries with respective projects:

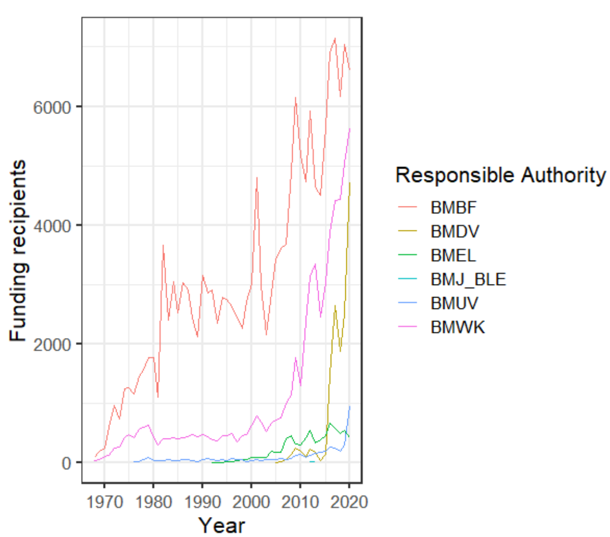
- Federal Ministry of Education and Research (BMBF): Information on project funding measures.
- Federal Ministry for the Environment, Nature Conservation, Nuclear Safety and Consumer Protection (BMUV): Information on project funding measures.
- Federal Ministry for Digital Affairs and Transport (BMDV): Information on project funding measures.
- Federal Ministry of Economics and Climate Protection (BMWK): Information on direct project funding in the fields of energy, aeronautics research, multimedia, space travel, and InnoNet.
- Federal Ministry of Food and Agriculture (BMEL): Information on direct project funding from the Federal Agency for Agriculture, Food and Renewable Resources.
- Federal Ministry of Justice (BMJ): Information on R&D projects in the field of innovations for consumer protection.

2.3.2 Descriptive Analysis of R&D Funding in Germany

2.3.2.1 Funding Targets

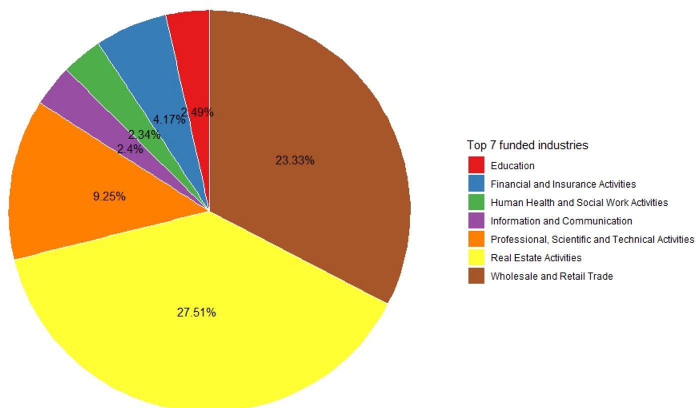
A descriptive analysis of R&D funding in Germany based on the Förderkatalog database allows observing that the largest share of funding recipients are under the responsibility of the Federal Ministry of Education and Research (BMBF), followed by the Federal Ministry of Economics and Climate Protection (BMWK) and the Federal Ministry for Digital Affairs and Transport (BMDV) (see Figure 2.4).

Figure 2.4: Funding recipients distribution per responsible authority in Germany



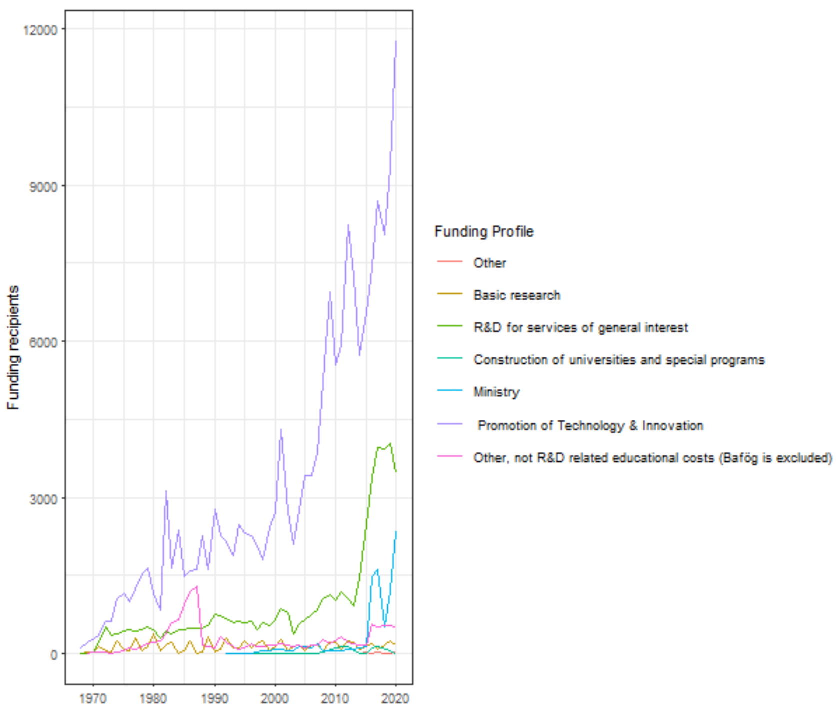
Despite there being no industry focus for R&D funding, Figure 2.5 shows that there is a preference for funding R&D projects connected to real estate activities, wholesale and retail trade, professional scientific and technical activities, financial and insurance activities, education, information and communication, and finally, human health and social work activities.

Figure 2.5: Total share of top 7 funded industries (per funding recipient) in 2000-2020



Regarding the funding profile, as expected, the largest proportion of funding is given to the applied research (promotion of technologies and innovation) (see Figure 2.6). Figure 2.6 further confirms the increasing pattern of R&D funding recipients over time reported earlier.

Figure 2.6: Funding recipients per funding profile in Germany

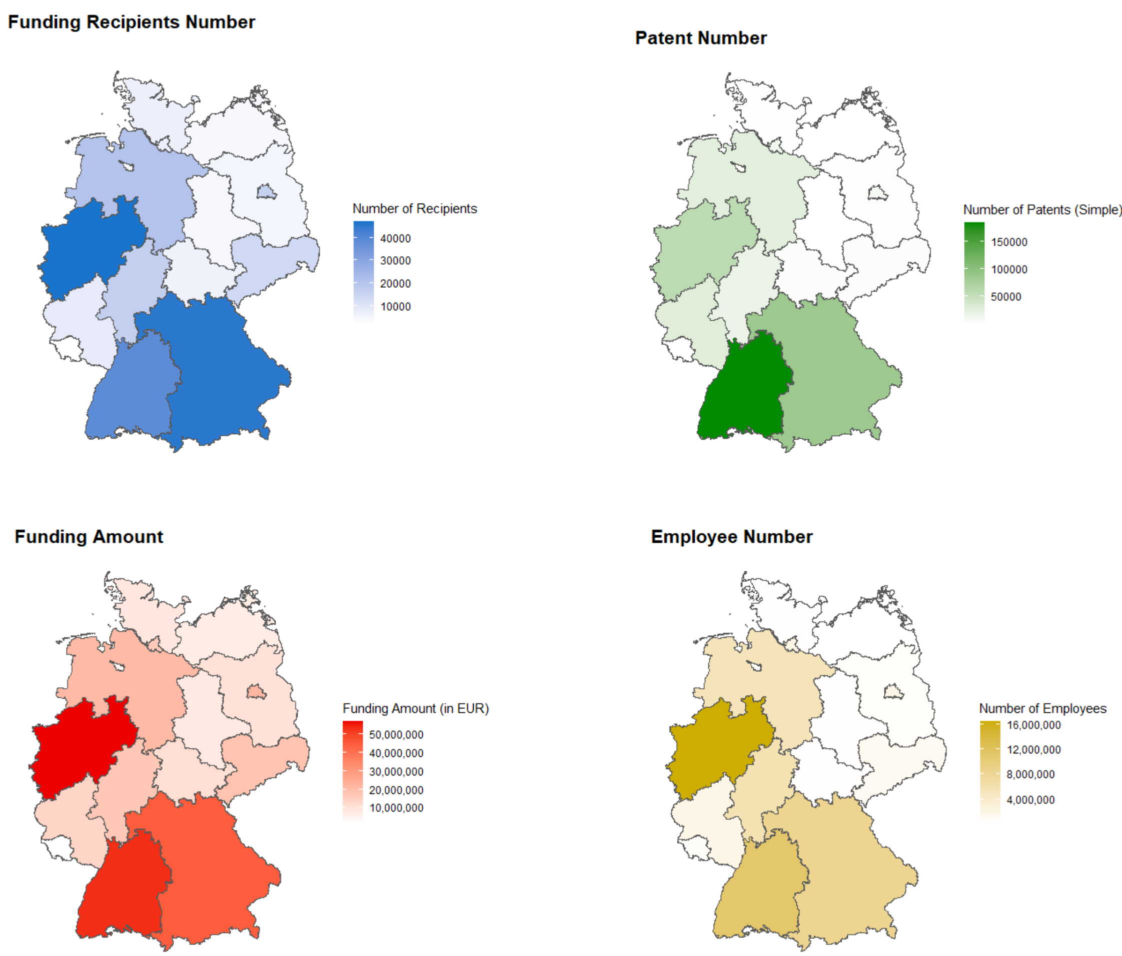


2.3.2.2 Funding concentration

Förderkatalog allows observing the distribution of funding recipients and funding amount (in EUR) in German regions (both on NUTS1 (Länder) and NUTS2 levels).

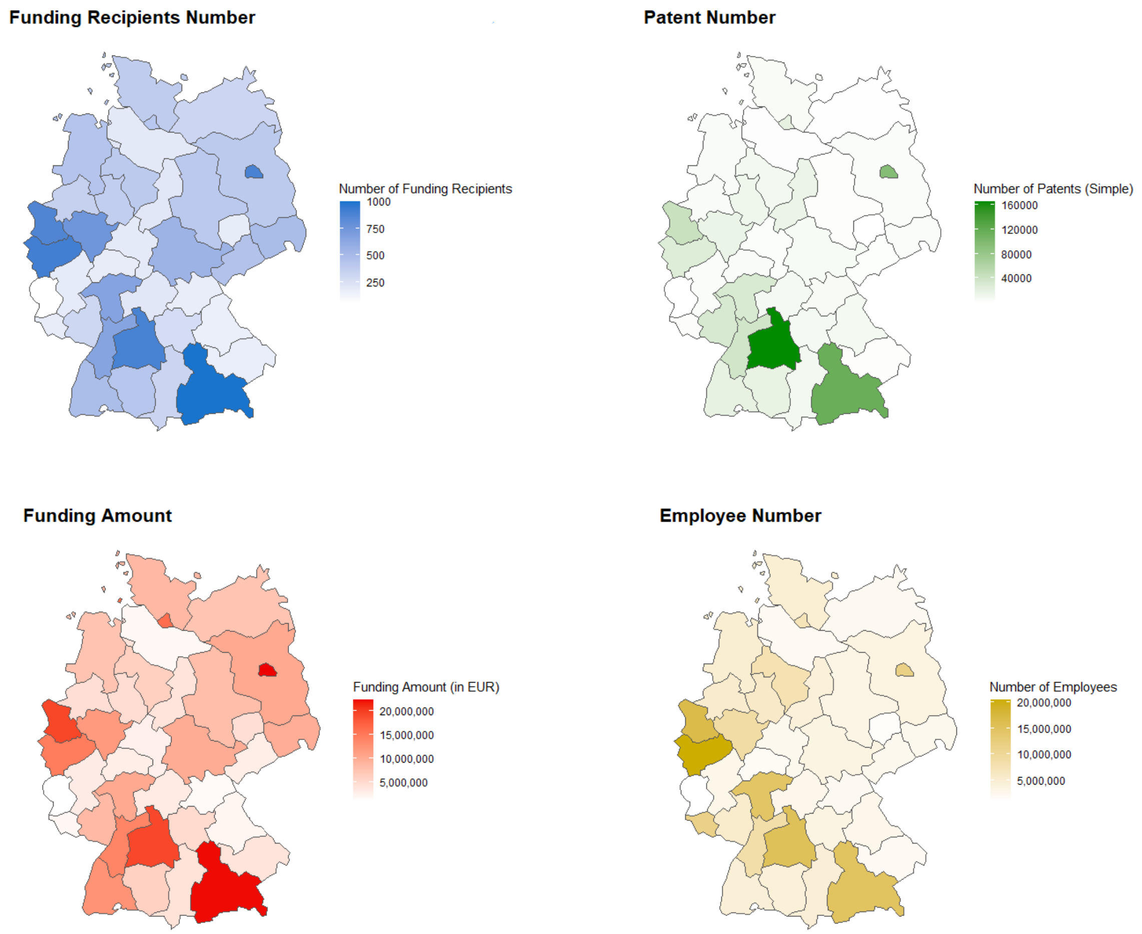
Figure 2.7 represents the distribution of funding recipients, funding amount (in EUR), patents (simple number), and employee numbers across German states. Upon initial examination of the figures, it becomes evident that Bavaria, North Rhine Westphalia, and Baden-Württemberg stand out as the primary recipients of funding, receiving the largest share of funding amounts. Remarkably, these regions also exhibit high levels of patenting activity and boast a larger number of employees. This suggests that the funding has been directed towards regions known for their innovation and abundant human capital.

Figure 2.7: The concentration of funding, innovation, and labor in German NUTS1 regions



When delving deeper into the NUTS2 level (see Figure 2.8), it becomes apparent that the bulk of funding recipients, innovation, and labor are centralized in major cities like Munich, Stuttgart, Cologne, Düsseldorf, and Berlin. This underscores the significant role urban centers play in driving innovation and economic growth within their respective regions.

Figure 2.8: The concentration of funding, innovation, and labor in German NUTS2 regions



Additionally, an examination of the dynamics of funding amounts, as depicted in Figure 2.9, reveals a consistent increase in R&D funding over the years. Notably, the top regions receiving the highest funding amounts are Baden-Württemberg, Bavaria, and Berlin, each receiving approximately 150-250 million EUR funding in 2020. Furthermore, since 1990, there has been a discernible positive trend in funding for less-supported regions such as Thuringia, Schleswig-Holstein, Saxony, and Saxony-Anhalt.

I also conducted an analysis on the distribution of R&D funding recipients across Germany, utilizing the Gini coefficient [2] and Lorenz curve. The Lorenz curve shows the cumulative share of funding plotted against the cumulative share of firms, where the 45-degree red line represents perfect equality and the blue curve represents the actual distribution. The Gini coefficient, indicating inequality, is calculated as the difference between the Lorenz curve and the line of perfect equality. These metrics provide insights into the concentration of funding among firms in 2000, 2010, and 2020 (see Figure 2.10). In 2000, the Gini coefficient stood at 0.91, indicating a notably high level of inequality in R&D funding distribution. Subsequently, there was a slight decrease to 0.88 in 2010 before returning to 0.91 in 2020.

[2] Gini coefficient is a measure of inequality, which ranges from 0 (or 0%) to 1 (or 100%), with 0 representing perfect equality and 1 representing perfect inequality.

These fluctuations suggest that a minority of firms received a disproportionately large share of R&D funding, leaving the majority with minimal resources (see Figure 2.10).

Figure 2.9: Funding distribution per region in Germany (in EUR)

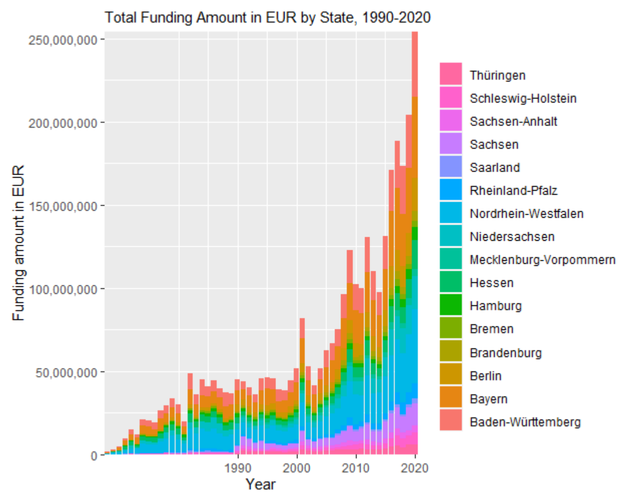
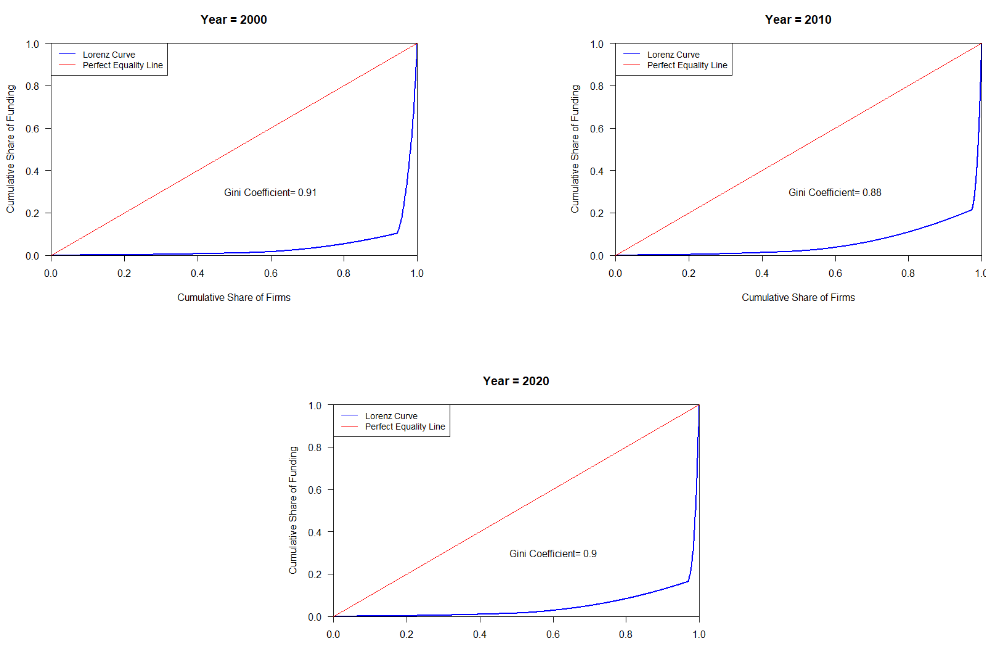


Figure 2.10: R&D funding concentration among firms in Germany



The Lorenz curve visually illustrates the cumulative distribution of R&D funding among firms. Notably, the curve consistently deviates from the line of perfect equality (45-degree line), indicating significant disparities in funding allocation. This further underscores the concentration of R&D funding among a select group of firms, potentially leading to imbalances in innovation and competitiveness (see Figure 2.10).

2.3.2.3 Funding probability

In this section, I examine the likelihood of receiving funding using a cross-sectional approach, focusing on the average year of 2010. I investigate the probability of receiving funding across various characteristics, including firm size, age, innovativeness, sector, multinational status, and research type, using an OLS regression model (see Appendix B.1.2 for the full list). Additionally, I conduct robustness checks using logistic regression (see Appendix B.2.1).

Table 2.3: Probability of being funded according to firm size (fixed assets)

Dependent Variable:	Treat
Model:	(1)
Variables	
(Intercept)	-0.0084***
	(0.0003)
Firm size (log(fixed assets))	0.0008***
	(2.97×10^{-5})
Fit statistics	
Observations	738,310
R^2	0.00185
Adjusted R^2	0.00185

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Tables 2.3 and 2.4 present the results of an OLS regression analysis indicating a statistically significant positive correlation between firm size (measured by log(fixed assets) and number of employees respectively) and the probability of being funded. These findings align with previous research conducted by Busom (2000).

Table 2.4: Probability of being funded according to firm size (number of employees)

Dependent Variable:	Treat
Model:	(1)
Variables	
(Intercept)	-0.0031***
	(0.0002)
Firm size (log(number of employees)	0.0030***
	(0.0001)
Fit statistics	
Observations	255,597
R^2	0.00728
Adjusted R^2	0.00728

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

In Table 2.5, the probability of receiving the funding is examined based on innovation activity. Models (1) and (2) reveal that each unit increase in the log of simple patents and quality-adjusted patents number is associated respectively with a 2.20% and 1.55% increase in the probability of receiving funding. Moreover, having prior patenting experience (patenting dummy = 1) increases the probability of being funded by 3.13% (Model 3), indicating that firms with patents are more favorably considered for funding. Similar findings are reported by Blanes and Busom (2004), who suggest that firms with prior experience in R&D are more likely to receive R&D subsidies compared to those without R&D experience.

Table 2.5: Probability of being funded according to innovation activity

Dependent Variable:		Treat	
Model:	(1)	(2)	(3)
Variables			
(Intercept)	0.0014***	0.0013***	0.0013***
	(3.97×10^{-5})	(3.94×10^{-5})	(3.86×10^{-5})
log(Patents (simple))	0.0220***		
	(0.0018)		
log(Patents (quality-adjusted))		0.0155***	
		(0.0012)	
Patenting dummy			0.0313***
			(0.0021)
Fit statistics			
Observations	877,934	877,934	877,934
R^2	0.00457	0.00507	0.00511
Adjusted R^2	0.00457	0.00507	0.00511

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

It's important to note that this relationship is merely a correlation, as is the case with other factors examined. In this context, it becomes apparent that R&D funding could potentially result in patents (suggesting reverse causality), rather than solely patents leading to funding. However, the inclusion of past patenting experience in the analysis helps mitigate this concern to some extent.

I further investigate the likelihood of receiving funding across various categories of firm size and patent output. Firm size is quantified using the natural logarithm of fixed assets (ln(fixed assets)), while patent output is assessed through the quality-adjusted number of patents. To categorize firms into different groups, I divide them into 20 bins based on the increasing order of their size or patent output. For instance, firms with the smallest size or patent output are assigned to bin 1, whereas those with the largest size or output are placed in bin 20. This segmentation enables a comprehensive examination of how funding propensity varies across different levels of firm size and patent output.

In Figure 2.11, we observe the distribution of funded and non-funded firms across various bins representing different firm sizes and patent outcomes. Notably, there is a lower proportion of funded firms in smaller-size bins, suggesting a preference for funding larger firms. Additionally, Figure 2.12 confirms this trend by illustrating a consistent increase in the propensity to be funded as the size of the bin increases, particularly favoring larger and more patent-intensive firms. These findings align with the expectation that larger firms may have greater resources, innovation capabilities, track records, and potential for commercialization, making them more attractive for R&D funding.

Figure 2.13 complements this by illustrating the distribution of funding recipients across different firm-size and patent bins. The red line represents the total number of funding recipients within each bin, while the blue line- the average number of funding recipients. The mean funding amount increases modestly from smaller to larger bins, which suggests a gradual increase in funding amounts as firm size and patent output increase.

Figure 2.11: The proportion of funded firms in firm size, and patent output groups

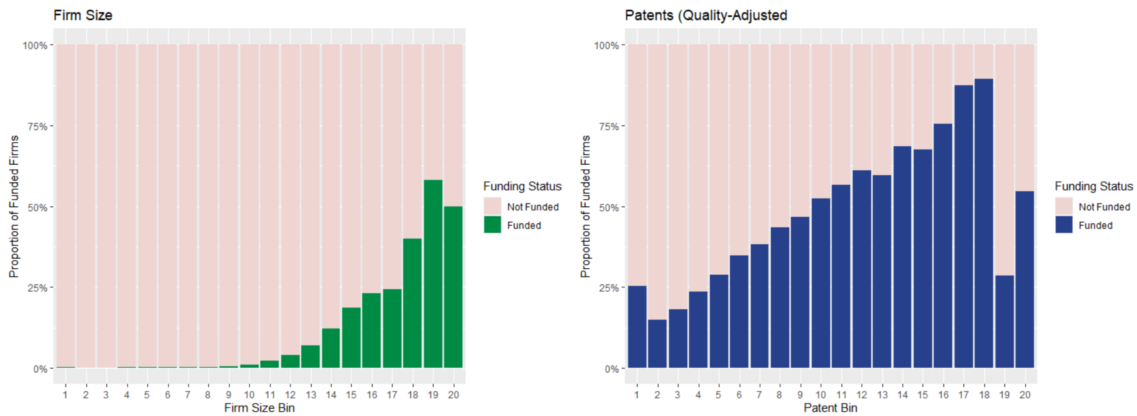


Figure 2.12: The propensity of being funded in firm size and patent output groups

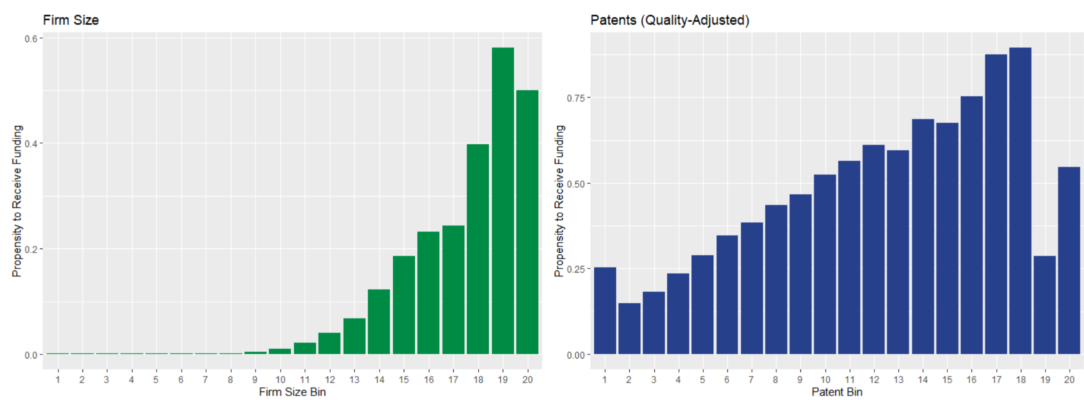
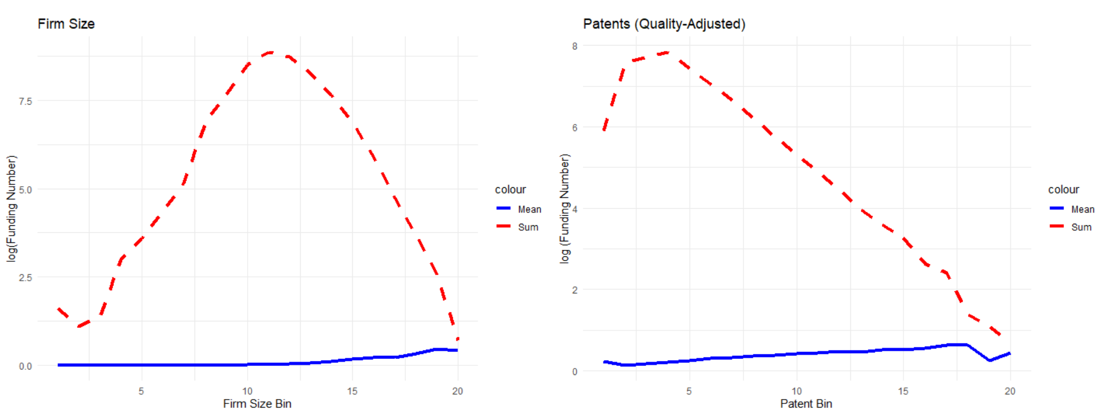


Figure 2.13: The average and total number of funded firms in firm size, and patent output groups



However, the pattern for the total number of funding recipients is different for firm size and patent bins (see red dashed line in Figure 2.13). For different firm-size bins, the total number of funding recipients exhibits an inverse U-shaped pattern. This implies that while there may be a steady increase in the number of funding recipients as firm size increases, there is an optimal firm size range (around bin 11) where the total number of funding recipients is maximized. Beyond this optimal range, the total number of funding recipients decreases as firm size continues to increase. This indicates that funders may prioritize firms within this optimal size range to distribute their resources more effectively.

Figure 2.14: Funding amount distribution in firm size, and patent output groups

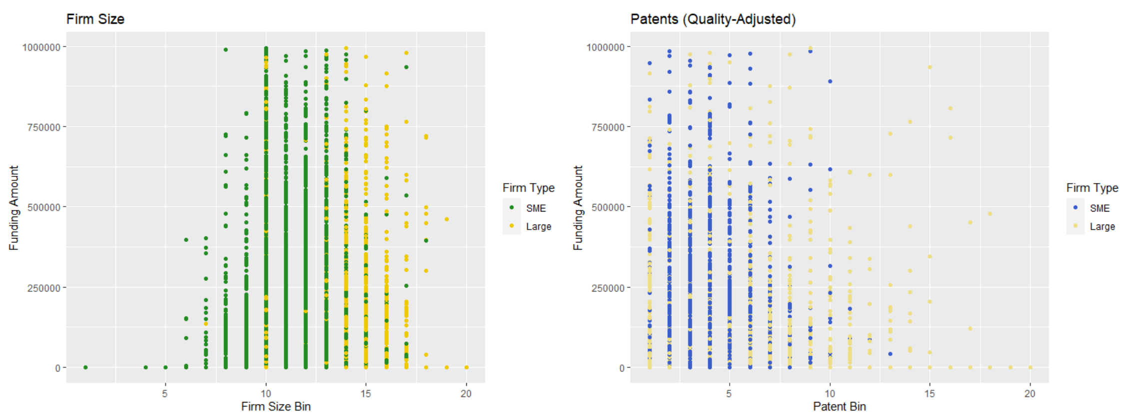
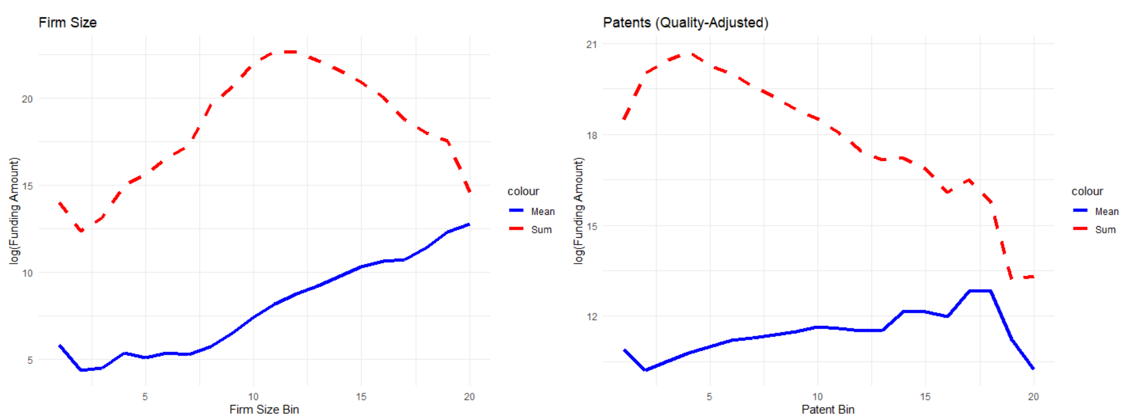


Figure 2.15: The average and total amount of funding in firm size, and patent output groups



For different patent bins, the total number of funding recipients increases up to patent bin 4 and then decreases (see Figure 2.13). This indicates that while there may be a relatively high number of funding recipients among firms with moderate levels of patent output (up to bin 4), the total number of funding recipients declines as patent output increases beyond this point. This implies that while funders are likely to favor companies with a proven track record of innovation, shown by their patent output, the drop in the total number of funding recipients after patent bin 4 suggests that funders might concentrate their resources on a smaller number of highly innovative firms with the potential for substantial impact, instead of distributing funds to a larger number of recipients.

In contrast, an examination of the distribution of funding amounts reveals a notable trend: a larger proportion of funding is allocated to SMEs and firms with lower numbers of patents, as depicted in Figure 2.14. Moreover, Figure 2.15 illustrates a decrease in both the average and total funding amounts as we move to higher bins, indicating a preference for allocating funds to firms with cash deficits and lower levels of patent intensity.

I expand the analysis of funding likelihood to include additional factors such as firm age, industry, and other relevant criteria. The findings displayed in Table 2.6 suggest a positive correlation between firm age [3] and the likelihood of securing R&D funding, indicating that older firms tend to have a higher probability of receiving funding.

Table 2.6: Probability of being funded according to firm age

Dependent Variable:	Treat
Model:	(1)
Variables	
(Intercept)	0.0004
	(0.0002)
log(Firm age)	0.0004***
	(7.49×10^{-5})
Fit statistics	
Observations	872,101
R^2	2.8×10^{-5}
Adjusted R^2	2.69×10^{-5}

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

In Table 2.7, a more general categorization of industries into primary, secondary, and tertiary sectors is used based on NACE 4-digit codes. However, due to the absence of firms in the primary industry (such as agriculture or raw material extraction) in the data, the analysis focuses on the secondary and tertiary sectors, where the tertiary sector is the reference industry. The relationship with secondary industries shows a positive and statistically significant coefficient (0.0017). This indicates that firms in secondary industries have a higher likelihood of funding compared to those in tertiary industries.

Table 2.7: Probability of being funded according to industry type

Dependent Variable:	Treat
Model:	(1)
Variables	
(Intercept)	0.0015***
	(4.21×10^{-5})
Secondary industry	0.0017**
	(0.0008)
Fit statistics	
Observations	877,934
R^2	1×10^{-5}
Adjusted R^2	8.9×10^{-6}

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The results in Table 2.8 present the probability of firms being funded based on more detailed industry categorization, as classified by NACE 1-digit code. Notably, several sectors

[3] Firm age is derived by subtracting the year of incorporation from the reference year (2010).

such as Education, Information and Communication, Manufacturing, Public Administration, and Defense demonstrate particularly strong positive associations with funding probability. Similar trends were observed by Svensson (1998) and Blanes and Busom (2004), suggesting that high-tech industries were the primary focus of R&D policies.

Table 2.8: Probability of being funded according to the industry type (NACE code)

Dependent Variable:	Treat
Model:	(1)
Variables	
(Intercept)	$-2.6 \times 10^{-18***}$
	(7.32×10^{-19})
Activities of households as employers	$-6.07 \times 10^{-18***}$
	(1.62×10^{-18})
Administrative and Support Service Activites	0.0009^{***}
	(0.0001)
Arts, Entertainment and Recreation	8.55×10^{-5}
	(8.55×10^{-5})
Construction of Buildings	0.0002^{***}
	(4.27×10^{-5})
Education	0.0053^{***}
	(0.0009)
Financial and Insurance Activities	0.0002^{***}
	(8.66×10^{-5})
Human Health and Social Work Activities	0.0017^{***}
	(0.0003)
Information and Communication	0.0046^{***}
	(0.0003)
Manufacturing	0.0045^{***}
	(0.0002)
Other Service Activities	0.0024^{***}
	(0.0003)
Professional, Scientific and Technical Activities	0.0016^{***}
	(9.18×10^{-5})
Public Administration and Defense: Compulsory Social Security	0.0051^{***}
	(0.0019)
Real Estate Activities	0.0002^{***}
	(5.55×10^{-5})
Transportation and Storage	0.0014^{***}
	(0.0002)
Water Supply	0.0025^{***}
	(0.0007)
Wholesale and Retail Trade	0.0006^{***}
	(5.62×10^{-5})
Fit statistics	
Observations	876,793
R^2	0.00159
Adjusted R^2	0.00157

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 2.9: Probability of being funded according to the type of research

Dependent Variable:	Treat	
Model:	(1)	(2)
Variables		
(Intercept)	0.0016***	5.08×10^{-5}***
	(4.21×10^{-5})	(7.66×10^{-6})
Basic research	-0.0016***	
	(4.21×10^{-5})	
Applied research		0.1107***
		(0.0029)
Fit statistics		
Observations	877,934	877,934
R^2	3.55×10^{-9}	0.10584
Adjusted R^2	-1.14×10^{-6}	0.10584

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Further, the impact of research type on funding probability is examined (see Table 2.9). The findings reveal that basic research is associated with a significant decrease in funding probability by approximately 0.16%, while firms engaged in applied research experience a substantial increase in funding probability by around 11.07 %, highlighting the preference towards applied research in R&D funding distribution. These findings are consistent with those of Arrow (1962), who argue that private R&D investment is typically not directed towards basic research due to its lower commercial potential.

Table 2.10: Probability of being funded according to being a multinational

Dependent Variable:	Treat
Model:	(1)
Variables	
(Intercept)	0.0013***
	(3.86×10^{-5})
MNE	0.0073***
	(0.0005)
Fit statistics	
Observations	877,934
R^2	0.00136
Adjusted R^2	0.00136

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Finally, multinational enterprises (MNEs) are significantly more likely to receive funding, with an approximate increase in the probability of 0.73%, emphasizing their heightened chances compared to non-multinational firms (see Table 2.10).

2.4 Method

2.4.1 Nearest Neighbor Matching

Next, I examine the effect of R&D funding on firm outcomes. To tackle the selection bias when comparing funded and non-funded firms, I employ Nearest Neighbor Matching (NNM). This approach matches each funded firm with an appropriate control firm, enabling a more robust comparison between the two groups, despite the inherent selectivity in R&D funding allocation. Given that our dataset spans from 1970 to 2020, with R&D funding initiated in 1970, a traditional pre-treatment period is unavailable. Therefore, I adopt a matching on-averages strategy.

To optimize the matching process, we focus on two subsets: small and medium-sized enterprises (SMEs) and large firms, as per the OECD classification [4]. For each funded firm in both subsets, we assemble a pool of potential control firms consisting solely of non-funded entities. Through NNM, we ensure precise matches based on average firm age (calculated as the mean of the difference between the current year and the year of incorporation), firm size (expressed as the mean of ln(total assets)), and industry (classified by NACE 2-digit codes). This rigorous matching strategy on industry addresses the significant variability in funded firms across industries, ensuring comparability between funded and control firms. Matching on average age and size further guarantees that each funded firm is paired with a control firm with a similar average lifespan and market potential.

Table 2.11, 2.12, 2.13, and 2.14 display summary statistics respectively for both funded and non-funded SMEs and large firms before and after the matching process. The initial three columns illustrate summary statistics for funded and non-funded firms pre-matching, while columns four, five, and six showcase summary statistics post-matching.

SMEs: Before matching, funded SMEs are, on average, approximately 5.5 years older than non-funded SMEs. There are slight differences in firm size, patents, and collaboration output between the two groups. Further, both funded and non-funded SMEs have a heterogeneous industry representation. However, after matching (columns four and five), we observe only minimal differences between funded and non-funded SMEs, suggesting successful implementation of the matching process. Funded and non-funded SMEs are now nearly identical in terms of explicitly matched covariates. Additionally, there is an increased similarity in financial outcomes not used in the matching procedure (see Table 2.11 and 2.12).

[4] According to the OECD definition, SMEs comprise firms with up to 249 employees, while large firms exceed this threshold.

Table 2.11: Descriptive statistics on SMEs before & after NN Matching

	Before Matching			After Matching		
	Funded (n=14206)	Non-Funded (n=959128)	Overall (n=973334)	Funded (n=14206)	Non-Funded (n=24489)	Overall (n=38695)
Average firm age	27.0 (28.5)	21.5 (24.5)	21.7 (24.6)	25.1 (26.7)	25.5 (32.7)	25.4 (30.5)
Firm size	15.8 (1.74)	13.9 (1.99)	14.0 (2.01)	15.3 (1.92)	14.3 (1.98)	14.7 (2.02)
Patent (simple)	3.42 (11.2)	2.87 (11.9)	3.00 (11.7)	4.22 (14.9)	3.71 (15.2)	4.06 (15.0)
Patent (adjusted)	8.07 (36.8)	8.60 (36.5)	12.1 (45.1)	10.9 (52.5)	11.7 (47.6)	22.6 (51)
Fixed assets	23.9 (279)	6.81 (150)	7.36 (156)	32.2 (606)	13.7 (363)	20.7 (470)
Total assets	39.8 (379)	11.5 (211)	12.4 (219)	50.9 (798)	22.2 (425)	33.1 (595)
Turnover	45.6 (436)	10.9 (212)	11.8 (220)	67.9 (527)	34.6 (372)	49.9 (450)
Sales	43.7 (138)	16.6 (194)	17.7 (192)	64.7 (399)	37.3 (391)	50.0 (395)
Labor costs	7.43 (23.1)	4.81 (28.1)	5.07 (27.7)	10.7 (64.2)	7.13 (80.3)	9.10 (71.9)

Notes: Fixed and total assets, turnover, sales, and labor costs values are in Mln.

Table 2.12: Descriptive statistics on NACE industry distribution of SMEs before & after NN Matching

Industry	Before Matching			After Matching		
	Funded (n=14206)	Non-Funded (n=959128)	Overall (n=973334)	Funded (n=14206)	Non-Funded (n=24489)	Overall (n=38695)
Accommodation and Food Service Activities	380 (0.4%)	50107 (1.0%)	50487 (1.0%)	770 (0.5%)	1162 (0.4%)	1932 (0.4%)
Administrative and Support Service Activities	415 (0.4%)	48041 (1.0%)	48456 (1.0%)	776 (0.5%)	1317 (0.5%)	2093 (0.5%)
Arts, Entertainment and Recreation	428 (0.4%)	16412 (0.3%)	16840 (0.3%)	586 (0.4%)	1186 (0.4%)	1772 (0.4%)
Construction of Buildings	908 (0.9%)	176461 (3.6%)	177369 (3.5%)	1729 (1.0%)	2002 (0.8%)	3731 (0.9%)
Education	2487 (2.6%)	35164 (0.7%)	37651 (0.8%)	4243 (2.5%)	7217 (2.7%)	11460 (2.7%)
Financial and Insurance Activities	228 (0.2%)	68506 (1.4%)	68734 (1.4%)	389 (0.2%)	586 (0.2%)	975 (0.2%)
Human Health and Social Work Activities	770 (0.8%)	30542 (0.6%)	31312 (0.6%)	1492 (0.9%)	2237 (0.8%)	3729 (0.9%)
Information and Communication	1273 (1.3%)	32672 (0.7%)	33945 (0.7%)	2565 (1.5%)	3664 (1.4%)	6229 (1.4%)
Manufacturing	1479 (1.5%)	26386 (0.5%)	27865 (0.6%)	2540 (1.5%)	4182 (1.6%)	6722 (1.6%)
Other Service Activities	1307 (1.4%)	43583 (0.9%)	44890 (0.9%)	1982 (1.2%)	3321 (1.3%)	5303 (1.2%)
Professional, Scientific and Technical Activities	2135 (2.2%)	118173 (2.4%)	120308 (2.4%)	3918 (2.3%)	6129 (2.3%)	10047 (2.3%)
Public Administration and Defense: Compulsory Social Security	765 (0.8%)	6309 (0.1%)	7074 (0.1%)	1281 (0.8%)	2125 (0.8%)	3406 (0.8%)
Real Estate Activities	1953 (2.0%)	359216 (7.3%)	361169 (7.2%)	3068 (1.8%)	5095 (1.9%)	8163 (1.9%)
Transportation and Storage	667 (0.7%)	33627 (0.7%)	34294 (0.7%)	1125 (0.7%)	1927 (0.7%)	3052 (0.7%)
Water Supply	448 (0.5%)	8082 (0.2%)	8530 (0.2%)	694 (0.4%)	1113 (0.4%)	1807 (0.4%)
Wholesale and Retail Trade	2909 (3.0%)	367590 (7.5%)	370499 (7.4%)	5078 (3.0%)	7545 (2.8%)	12623 (2.9%)
Activities of households as employers	0 (0%)	88 (0.0%)	88 (0.0%)	0	0	0

Large firms: Table 2.13 reveals that prior to matching, funded large firms demonstrate an average age of approximately 9.5 years greater than non-funded large firms. While there is an expected similarity in firm size (given the initial subset to large firms), substantial disparities persist in terms of industry focus and financial and innovation metrics. Funded large firms are predominantly concentrated in R&D-intensive sectors such as education, professional, scientific, and technical activities, and wholesale and retail trade. Conversely, non-funded large firms exhibit a significant presence in industries like wholesale and retail trade, real estate activities, and construction of buildings (see Table 2.14). Following the matching process, their characteristics become highly comparable concerning matched covariates (average firm size, age, and industry) (see Table 2.13 and 2.14).

Table 2.13: Descriptive statistics on large firms before & after NN Matching

	Before Matching			After Matching		
	Funded (n=3005)	Non-Funded (n=16543)	Overall (n=19548)	Funded (n=3005)	Non-Funded (n=4707)	Overall (n=7712)
	(N=27146)	(N=103844)	(N=130990)	(N=38047)	(N=54287)	(N=92334)
Average firm age	45.7 (47.0)	36.2 (39.2)	38.2 (41.1)	42.7 (44.5)	40.1 (43.4)	41.1 (43.9)
Firm size	18.9 (1.67)	17.4 (1.81)	17.8 (1.88)	18.4 (1.87)	17.4 (1.95)	17.8 (1.98)
Patent (simple)	46.9 (273)	22.3 (225)	33.7 (249)	40.1 (251)	22.6 (234)	32.2 (244)
Patent (adjusted)	125 (759)	54.1 (560)	86.9 (661)	103 (674)	54.8 (581)	81.4 (634)
Fixed assets	709 (4270)	150 (2050)	274 (2710)	532 (3590)	161 (1580)	313 (2600)
Total assets	1130 (6350)	264 (2590)	456 (3780)	843 (5340)	283 (2340)	512 (3870)
Turnover	878 (4570)	445 (3070)	546 (3490)	709 (4010)	304 (1660)	482 (2940)
Sales	835 (4590)	305 (2060)	439 (2930)	666 (4000)	284 (1610)	454 (2930)
Labor costs	150 (735)	53.3 (1150)	76.2 (1070)	114 (626)	51.6 (1520)	78.5 (1220)

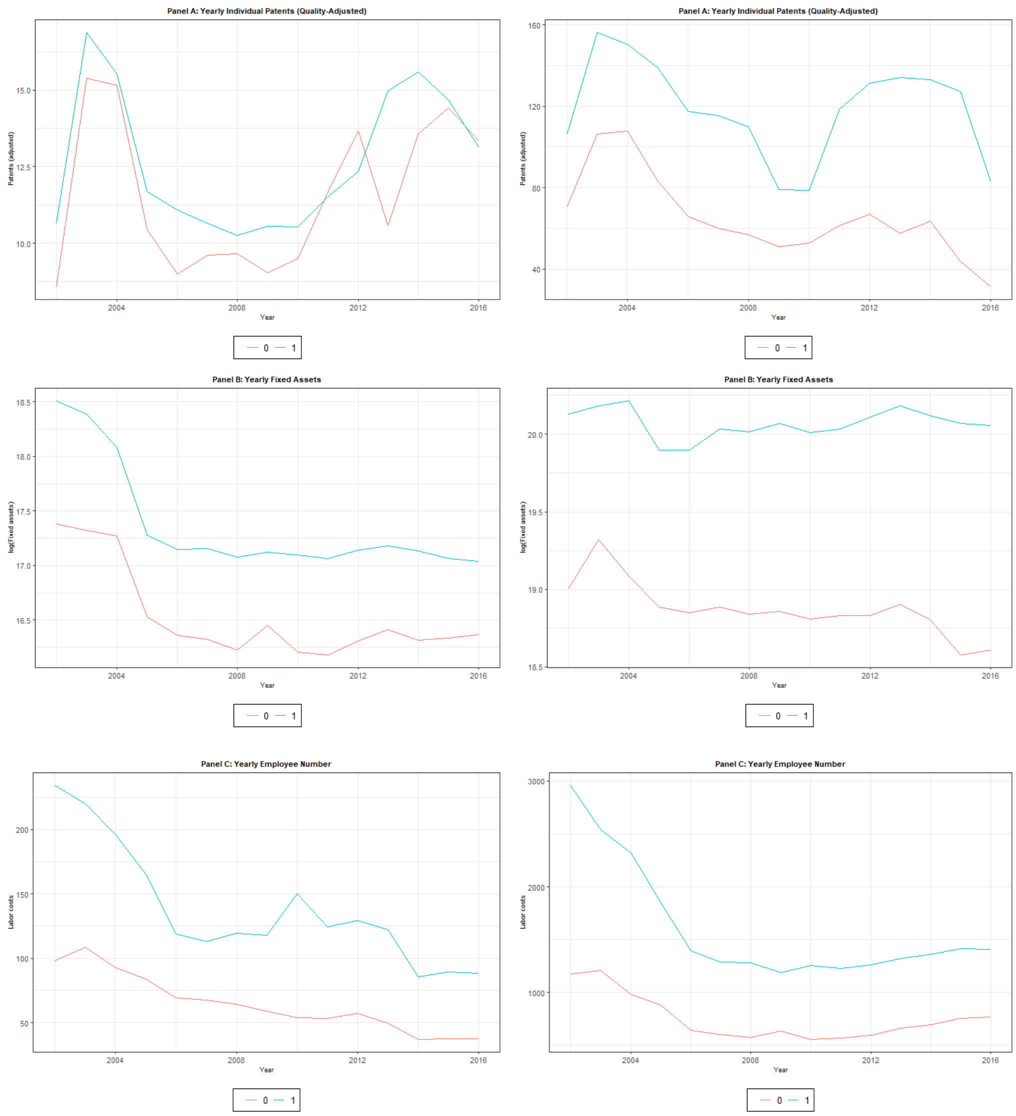
Notes: Fixed and total assets, turnover, sales, and labor costs values are in Mln.

Table 2.14: Descriptive statistics on NACE industry before & after NN Matching

	Before Matching			After Matching		
Industry	Funded (n=3005)	Non-Funded (n=16543)	Overall (n=19548)	Funded (n=3005)	Non-Funded (n=4707)	Overall (n=7712)
Accommodation and Food Service Activities	9 (0.0%)	927 (0.9%)	936 (0.7%)	18 (0.0%)	22 (0.0%)	40 (0.0%)
Administrative and Support Service Activities	93 (0.3%)	2024 (1.9%)	2117 (1.6%)	118 (0.3%)	207 (0.4%)	325 (0.4%)
Arts, Entertainment and Recreation	28 (0.1%)	275 (0.3%)	303 (0.2%)	44 (0.1%)	74 (0.1%)	118 (0.1%)
Construction of Buildings	120 (0.4%)	822 (0.8%)	942 (0.7%)	184 (0.5%)	265 (0.5%)	449 (0.5%)
Education	422 (1.6%)	814 (0.8%)	1236 (0.9%)	764 (2.0%)	1272 (2.3%)	2036 (2.2%)
Financial and Insurance Activities	89 (0.3%)	1875 (1.8%)	1964 (1.5%)	132 (0.3%)	180 (0.3%)	312 (0.3%)
Human Health and Social Work Activities	979 (3.6%)	5598 (5.4%)	6577 (5.0%)	1313 (3.5%)	2001 (3.7%)	3314 (3.6%)
Information and Communication	146 (0.5%)	580 (0.6%)	726 (0.6%)	221 (0.6%)	327 (0.6%)	548 (0.6%)
Manufacturing	587 (2.2%)	1023 (1.0%)	1610 (1.2%)	815 (2.1%)	1083 (2.0%)	1898 (2.1%)
Other Service Activities	171 (0.6%)	918 (0.9%)	1089 (0.8%)	249 (0.7%)	369 (0.7%)	618 (0.7%)
Professional, Scientific and Technical Activities	254 (0.9%)	786 (0.8%)	1040 (0.8%)	423 (1.1%)	675 (1.2%)	1098 (1.2%)
Public Administration and Defense: Compulsory Social Security	108 (0.4%)	393 (0.4%)	501 (0.4%)	183 (0.5%)	227 (0.4%)	410 (0.4%)
Real Estate Activities	130 (0.5%)	1433 (1.4%)	1563 (1.2%)	265 (0.7%)	568 (1.0%)	833 (0.9%)
Transportation and Storage	440 (1.6%)	922 (0.9%)	1362 (1.0%)	579 (1.5%)	816 (1.5%)	1395 (1.5%)
Water Supply	132 (0.5%)	166 (0.2%)	298 (0.2%)	188 (0.5%)	197 (0.4%)	385 (0.4%)
Wholesale and Retail Trade	413 (1.5%)	5122 (4.9%)	5535 (4.2%)	609 (1.6%)	936 (1.7%)	1545 (1.7%)

Figure 2.16 displays the mean values of innovation and firm-level outcomes for funded SMEs and large firms and their matched control firms from 2000 to 2016. Notably, the increased similarity in funding trends between funded and matched controls both for SMEs and large firms confirms the effectiveness of the matching process. The figure indicates that R&D funding contributes to an increase in the number of quality-adjusted patents (panel A) for SMEs, aligning with previous research findings (Bronzini and Piselli 2016; Czarnitzki and Hussinger 2017; Czarnitzki et al. 2007), however, it had no impact on innovation outcomes of large firms. Simultaneously, the impact of R&D funding on fixed assets (panel B) and employee numbers (panel C) for both SMEs and large firms appears negligible, likely due to the relatively small scale of the funding, which may not significantly affect firm financial outcomes.

Figure 2.16: Descriptive statistics for funded vs non-funded SMEs before and after matching



2.4.2 The Standard TWFE

In this analysis, I utilize the canonical Two-Way Fixed Effects (TWFE) model (following (Callaway and Sant'Anna 2021; Goodman-Bacon 2021)) to investigate the effects of R&D funding on firm performance and innovation activity in Germany from 1970 to 2020. In the following, i indexes the firm, $t \in \{1, \ldots T\}$ the time period, $y_{i,t}$ denotes the outcome variable, and $D_{i,t}$ is a binary treatment indicator equal to one if firm i receives R&D funding in period t. I assume that once a firm becomes treated, that unit remains treated in future time periods. This allows to define groups $g \in G = \{1, \ldots T, \infty\}$ by the time period when a unit first becomes treated, where $g = \infty$ refers to the group of never-treated firms.

The target parameters are the group-time average treatment effects

$$ATT(gt) \equiv \mathbb{E}\left[y_{i,t}(1) - y_{i,t}(0) \mid G = g\right]. \tag{2.1}$$

that report the average difference in potential outcomes under treatment ($y_{i,t}(1)$) and no treatment ($y_{i,t}(0)$) for group g at time period t. Notice that the group-time average treatment effects capture both leads ($t < g$), lags ($t > g$), and the instantaneous ($t = g$) effect of treatment.

The $ATT(g,t)$ are estimated via the following dynamic two-way fixed effects (TWFE) regression:

$$y_{i,t} = \alpha_i + \lambda_t + \sum_{\ell=-K}^{-2} \mu_\ell D_{i,t}^{\ell} + \sum_{\ell=0}^{L} \mu_\ell D_{i,t}^{\ell} + \varepsilon_{i,t}, \tag{2.2}$$

where $y_{i,t}$ is the outcome of interest (quality-adjusted number of patents, fixed assets, employment costs) for firm i at time t, α_i denotes firm fixed effects and λ_t denotes time fixed effects. D_{it}^{ℓ} is equal to 1 for unit i in period t if unit i has been exposed to treatment for exactly ℓ periods in period t. For example, D_{it}^{0} is equal to one for units that become treated in period t, D_{it}^{2} is equal to one for units that became treated in period $t-2$ and D_{it}^{-2} equals one for all units that become treated in period $t+2$. We refer to ℓ as event time. The interest lies in the coefficients μ_ℓ that identify (under appropriate assumptions) the group-time average treatment effects. Notice that the coefficients μ_ℓ should be interpreted as the effect of treatment relative to the omitted reference period $\ell = -1$. Finally, $\varepsilon_{i,t}$ is the error term (robust standard errors are adjusted for both heteroscedasticity and clustering at the firm level).

As is common in practice, I bin distant relative periods [-22:4), and (4:49] to end up with 4 pre and 4 post-treatment periods.

In this context, I discuss the following identification assumptions underlying the dynamic TWFE model (see Goodman-Bacon (2021) and Callaway and Sant'Anna (2021)).

Assumption 1 (Irreversibility of Treatment (Staggered Adoption)): This assumption stipulates that treatment assignment is irreversible.

Assumption 2 (Random Sampling): This assumption implies that the units in the sample are randomly drawn from the population of interest conditional on unit and time-fixed effects. It ensures that the estimated treatment effects are generalizable to the broader population and not biased due to non-random selection.

Assumption 3 (Parallel Trends): Formally, this assumption states that average potential outcomes under no treatment evolve in parallel over time for all groups and time periods.

Assumption 4 (No/Limited Anticipatory Behavior Prior to Treatment): This assumption rules out anticipatory responses of firms in all periods prior to the first time of treatment. Together with the parallel trend assumption it implies that the μ_ℓ should be zero for $\ell < 0$. Put differently, $\mu_\ell \neq 0, \ell < 0$ quantifies the violation of the parallel trends assumption.

2.4.3 Honest Difference-in Differences

Subsequent analysis reveals that the parallel trends assumption is violated (see 2.5.1), motivating the adoption of the Honest Difference-in-Differences (DiD) approach, as advocated by (Rambachan and Roth 2023). This approach enables the application of robust inference and sensitivity analysis in scenarios where the assumption of parallel trends may not be valid. Building on the traditional Difference-in-Difference (DID) framework and following the work of Manski and Pepper (2018), Rambachan and Roth (2023) show that causal parameters of interest are partially identified by imposing restrictions that ensure that deviations from parallel trends after treatment are not too dissimilar from pre-treatment trend differences. Inference accounts both for "statistical uncertainty" (due to imprecise estimation of the true pre-trend) and "identification uncertainty" (the challenge of accurately extrapolating the true pre-trend even if known).

The general idea behind the approach of Rambachan and Roth (2023) is to learn from pre-treatment parallel trend violations about parallel trend violations in the treatment period. To do so, they discuss two restrictions that link parallel trend violations in the pre-and post-treatment period. To formally define these restrictions, let δ_t denote the average time trend differential between groups for the period t (relative to a base period) and $\delta_{t+1} - \delta_t$ the time trend differential between periods $t+1$ and t (because the base period is always the same and thus cancels). The first restriction imposes that $\delta \in \Delta^{\mathrm{RM}}(\bar{M})$ for $\bar{M} \geq 0$, where

$$\Delta^{\mathrm{RM}}(\bar{M}) = \left\{\delta : \forall t \geq 0, |(\delta_{t+1} - \delta_t)| \leq \bar{M} \cdot \max_{s<0} |(\delta_{s+1} - \delta_s)|\right\}. \tag{2.3}$$

$\Delta^{\mathrm{RM}}(\bar{M})$[5] constrains the maximum post-treatment deviation of parallel trends between consecutive periods by $\bar{M}$ times the maximum pre-treatment deviation of parallel trends. Opting for $RM(\bar{M})$ could be justified if the researcher suspects that potential violations of parallel trends result from confounding economic shocks of comparable magnitude to those in the preperiod. When the number of pre-treatment and post-treatment periods is comparable, a natural reference point might be $\bar{M} = 1$, which restricts the worst-case post-treatment trend difference to the maximum equivalent in the pre-treatment period.

The second restriction imposes a bound on the *change* in the parallel trend violation across consecutive periods. Formally, it imposes that $\delta \in \Delta^{\mathrm{SD}}(M)$ for $M \geq 0$, where

$$\Delta^{\mathrm{SD}}(M) := \{\delta : |(\delta_{t+1} - \delta_t) - (\delta_t - \delta_{t-1})| \leq M, \forall t\}. \tag{2.4}$$

[5]RM stands for "relative magnitudes".

In the special case $M = 0, \Delta^{\mathrm{SD}}(0)$ mandates a precisely linear trend difference, aligning with the typical assumption in parametric linear specifications used in practice (Rambachan and Roth 2023).

2.4.4 Parametric Event Study- Detrending

In addition to employing the standard and Honest TWFE approaches, I control for a linear treatment group-specific time trend. Identification now relies on the weaker assumption of parallel trends after accounting for linear pre-trends.

I employ two different approaches for implementing such a parametric event study. In the first approach, I adjust the baseline model (see Equation 2.2) to allow for a linear trend in event time ℓ and estimate the resulting regression equation using all data:

$$y^1_{i,t} = \alpha'_i + \lambda'_t + \tau\ell + \sum_{\ell=-K}^{-2} \mu'_\ell D^\ell_{i,t} + \sum_{\ell=0}^{L} \mu'_\ell D^\ell_{i,t} + \varepsilon'_{i,t} \tag{2.5}$$

The second approach follows a two-step procedure. In the first step, I estimate the augmented regression equation above using only observations under no treatment, i.e. firms in the control group and treated firms up to the year of treatment (this, of course, requires dropping terms that refer to lagged treatment effect, i.e. $\ell > 0$). In a second step, I derive a de-trended outcome by deducting the linear pre-trend estimated in step 1 from the outcome for post-treatment observations.

2.5 Results

2.5.1 The Standard TWFE

The table 2.15 presents the results of the direct impact of R&D funding on firm-level and innovation outcomes for both SMEs and large firms during 1970-2020 using a standard two-way fixed effects (TWFE) model. For instance, innovation outcomes of SMEs display a notable negative impact post-funding, whereas large firms demonstrate a significant positive growth of approximately 6% in quality-adjusted patent numbers one year after R&D funding. Likewise, the coefficients for firm outcomes also exhibit disparities between SMEs and large firms across various treatment periods. While SMEs experience significant growth of 13% and 5% for fixed assets and employment costs, respectively, the impact of R&D funding on firm-level outcomes for large firms is deemed insignificant.

However, in Figure 2.17 I observe statistically significant effects in the pre-treatment period, suggesting potential pre-existing trends, which will be explored in the next subsection.

Figure 2.17: Direct impact of R&D funding on firm-level and innovation outcomes

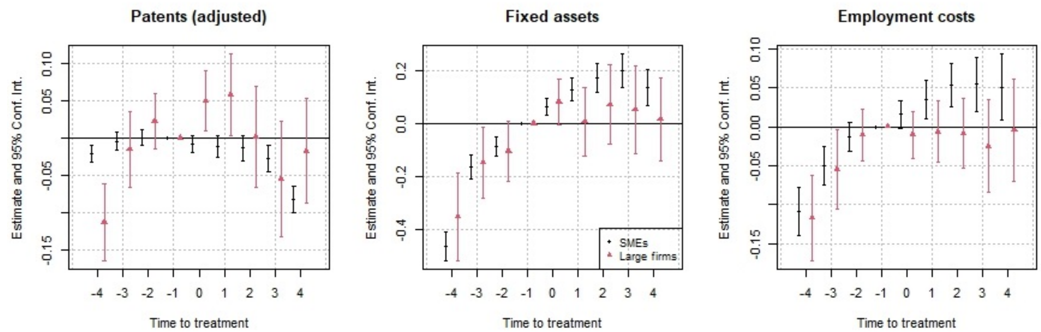


Table 2.15: Standard TWFE coefficients for the direct impact of R&D funding

	SMEs			Large firms		
Dependent Variables:	Patent (quality-adjusted)	Fixed assets	Employment costs	Patent (quality-adjusted)	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Treat × custom_bins = -4	-0.0211***	-0.4626***	-0.1090***	-0.1134***	-0.3515***	-0.1173***
	(0.0053)	(0.0212)	(0.0114)	(0.0219)	(0.0631)	(0.0198)
Treat × custom_bins = -3	-0.0039	-0.1639***	-0.0508***	-0.0154	-0.1459**	-0.0547**
	(0.0062)	(0.0246)	(0.0128)	(0.0258)	(0.0724)	(0.0236)
Treat × custom_bins = -2	0.0008	-0.0854***	-0.0126	0.0229	-0.1026	-0.0106
	(0.0064)	(0.0252)	(0.0122)	(0.0257)	(0.0772)	(0.0216)
Treat × custom_bins = 0	-0.0077	0.0651***	0.0163	0.0499*	0.0828	-0.0100
	(0.0070)	(0.0249)	(0.0133)	(0.0274)	(0.0722)	(0.0250)
Treat × custom_bins = 1	-0.0115	0.1297***	0.0347**	0.0583**	0.0074	-0.0063
	(0.0076)	(0.0255)	(0.0147)	(0.0294)	(0.0755)	(0.0263)
Treat × custom_bins = 2	-0.0135*	0.1727***	0.0536***	0.0019	0.0725	-0.0088
	(0.0081)	(0.0261)	(0.0139)	(0.0326)	(0.0751)	(0.0253)
Treat × custom_bins = 3	-0.0273***	0.2004***	0.0546***	-0.0547	0.0534	-0.0249
	(0.0085)	(0.0279)	(0.0157)	(0.0369)	(0.0782)	(0.0290)
Treat × custom_bins = 4	-0.0820***	0.1357***	0.0507***	-0.0171	0.0175	-0.0043
	(0.0069)	(0.0242)	(0.0140)	(0.0270)	(0.0660)	(0.0230)
Fixed-effects						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	428,073	385,055	108,104	87,822	72,964	64,366
R^2	0.58846	0.81653	0.90453	0.72840	0.81236	0.81464
Within R^2	0.00139	0.00352	0.00290	0.00141	0.00182	0.00132

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

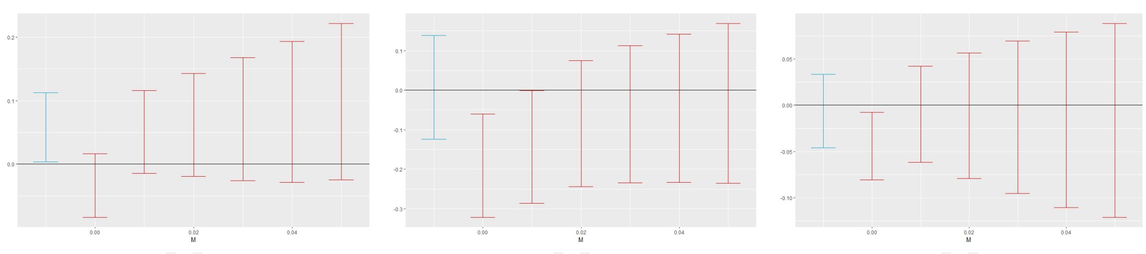
2.5.2 Honest DID

Given the potential presence of significant pre-trends in the standard TWFE model (see Section 2.5.1) and the consequent risk of violating the parallel trend assumption in this context, I adopt the Honest Difference-in-Differences (HDD) method by Rambachan and Roth (2023). Since the observations of Figure 2.17 show that the difference in trends seems to be linear, I implement sensitivity analysis using **the smoothness restriction**. Additional sensitivity analysis using the relative magnitude restriction is provided in Appendix B.2.2.

It is important to note that the smoothness restriction in the Honest DID framework specifically addresses the period immediately following treatment. This restriction involves constraining the extent to which the slope of the trend difference can change between consecutive periods, with the degree of change bounded by the parameter M.

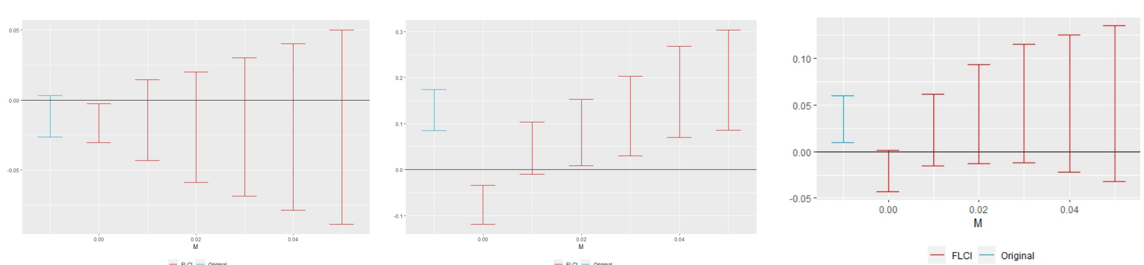
Large firms: Figure 2.18 presents the results of the R&D funding impact on firm performance and innovation outcomes in the first post-treatment period, along with robust confidence intervals assessing the non-linearity of trend differences (see Table 2.16 for more details on coefficients). In blue, the original TWFE confidence intervals for μ^1 are depicted from specification 2.2. In red, fixed length confidence intervals (FLCIs) are shown for different values of M, where M = 0 signifies linear violations of parallel trends, and larger M values allow for greater deviations from linearity. In the analysis for large firms (Figure 2.18), the original TWFE estimate shows a positive effect on the quality-adjusted number of patents, with the confidence interval excluding the 0 effect. However, for fixed assets and employment costs, the null effects cannot be ruled out, suggesting no significant impact of funding. Nonetheless, when allowing for linear violations of parallel trends (M = 0), the scenario changes markedly. Despite observing positive upward trends in all outcomes, the significance diminishes immediately (the threshold value for a significant effect is M=0), signaling a significant pre-trend in the analysis that impacts the outcome.

Figure 2.18: Sensitivity analysis using smoothness restriction



SMEs: Figure 2.19 illustrates the effect of R&D funding one year after treatment, revealing insignificant impacts on innovation outcomes but positive and significant effects on fixed assets and employment costs (depicted by the blue line). Nevertheless, permitting linear deviations from parallel trends (M = 0) results in the significant effect vanishing, with a critical value of M = 0 (depicted by the red FLCIs) (more details on coefficients are provided in Table 2.17).

Figure 2.19: Sensitivity analysis using smoothness restriction



In summary, the significant effect detected in the initial TWFE analysis diminishes after conducting a sensitivity analysis with Honest DID. This suggests that the initial finding may have been affected by potential violations of the parallel trends assumption. Consequently, it's advisable to interpret the initial result with caution, acknowledging the uncertainty surrounding the estimated effect of R&D funding on the outcome variable, especially considering the possibility of pre-existing trends. As a next step, I proceed with further analysis using the parametric event study approach and detrend the outcome variable.

Table 2.16: Honest DID: Sensitivity analysis using smoothness restriction for Large firms

lb	ub	method	Delta	M	Model
-0.0837325	0.0162888	FLCI	DeltaSD	0.00	Patents (adjusted)
-0.0145219	0.1161079	FLCI	DeltaSD	0.01	Patents (adjusted)
-0.0194132	0.1430725	FLCI	DeltaSD	0.02	Patents (adjusted)
-0.0258184	0.1675384	FLCI	DeltaSD	0.03	Patents (adjusted)
-0.0286330	0.1930255	FLCI	DeltaSD	0.04	Patents (adjusted)
-0.0245585	0.2218887	FLCI	DeltaSD	0.05	Patents (adjusted)
-0.3215141	-0.0604507	FLCI	DeltaSD	0.00	Fixed assets
-0.2859914	-0.0004535	FLCI	DeltaSD	0.01	Fixed assets
-0.2437178	0.0750096	FLCI	DeltaSD	0.02	Fixed assets
-0.2344220	0.1120610	FLCI	DeltaSD	0.03	Fixed assets
-0.2334369	0.1417263	FLCI	DeltaSD	0.04	Fixed assets
-0.2351716	0.1690722	FLCI	DeltaSD	0.05	Fixed assets
-0.0807193	-0.0077774	FLCI	DeltaSD	0.00	Employment costs
-0.0615718	0.0419712	FLCI	DeltaSD	0.01	Employment costs
-0.0790608	0.0565222	FLCI	DeltaSD	0.02	Employment costs
-0.0953191	0.0694249	FLCI	DeltaSD	0.03	Employment costs
-0.1104177	0.0789739	FLCI	DeltaSD	0.04	Employment costs
-0.1211028	0.0883888	FLCI	DeltaSD	0.05	Employment costs

Table 2.17: Honest DID: Sensitivity analysis using smoothness restriction for SMEs

lb	ub	method	Delta	M	Model
-0.0302948	-0.0025490	FLCI	DeltaSD	0.00	Patents (adjusted)
-0.0431313	0.0144104	FLCI	DeltaSD	0.01	Patents (adjusted)
-0.0586593	0.0200793	FLCI	DeltaSD	0.02	Patents (adjusted)
-0.0686593	0.0300793	FLCI	DeltaSD	0.03	Patents (adjusted)
-0.0786593	0.0400793	FLCI	DeltaSD	0.04	Patents (adjusted)
-0.0886593	0.0500793	FLCI	DeltaSD	0.05	Patents (adjusted)
-0.1189649	-0.0344228	FLCI	DeltaSD	0.00	Fixed assets
-0.0096242	0.1032494	FLCI	DeltaSD	0.01	Fixed assets
0.0088225	0.1530859	FLCI	DeltaSD	0.02	Fixed assets
0.0297369	0.2029838	FLCI	DeltaSD	0.03	Fixed assets
0.0698674	0.2681608	FLCI	DeltaSD	0.04	Fixed assets
0.0854400	0.3041405	FLCI	DeltaSD	0.05	Fixed assets
-0.0429791	0.0016012	FLCI	DeltaSD	0.00	Employment costs
-0.0155892	0.0613586	FLCI	DeltaSD	0.01	Employment costs
-0.0129183	0.0930599	FLCI	DeltaSD	0.02	Employment costs
-0.0123896	0.1144837	FLCI	DeltaSD	0.03	Employment costs
-0.0223895	0.1244835	FLCI	DeltaSD	0.04	Employment costs
-0.0323895	0.1344835	FLCI	DeltaSD	0.05	Employment costs

2.5.3 Parametric Event Study: De-trending Treatment Group

After detecting violations of pre-trends through sensitivity analysis using Honest DiD, I adjust for linear time trend differentials in the final step of my analysis.

2.5.3.1 Approach 1 - Adding linear time trend

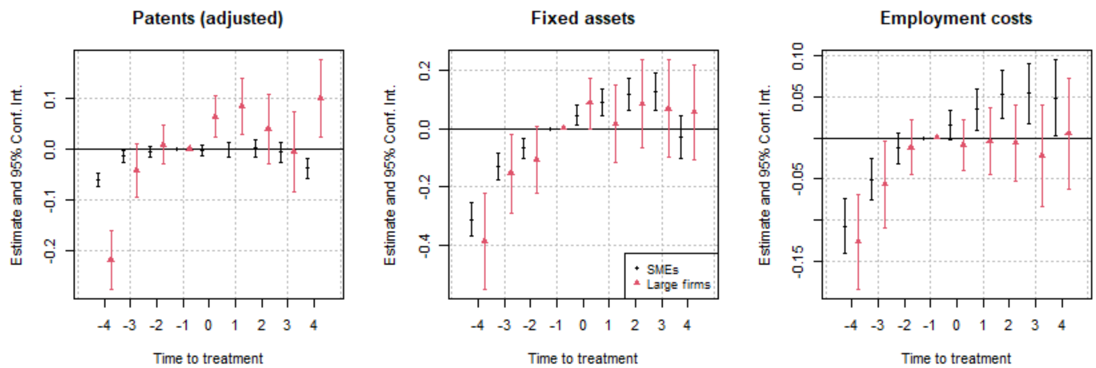
In approach 1, I incorporate a linear treatment group-specific time trend by including the interaction term of the relative year (Time to treatment) with the treatment indicator. Table 2.18 demonstrates that even after adjusting for linear trend differences, the results remain largely similar, albeit with smaller coefficients. Upon examining the interaction between treatment and the continuous variable representing the linear trend, I find contrasting effects of R&D funding. Specifically, while the impacts on firm-level outcomes such as fixed assets and employment costs for SMEs are positive, they exhibit negative effects on innovation outcomes. Conversely, for large firms, the analysis indicates no significant effects. Refer to Figure 2.20 for a visual representation of these findings.

Table 2.18: The direct impact of R&D funding on firm-level & innovation outcomes

	SMEs			Large firms		
Dependent Variables:	Patent (quality-adjusted)	Fixed assets	Employment costs	Patent (quality-adjusted)	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Treat × Relative period = -4	-0.0606***	-0.3111***	-0.1067***	-0.2176***	-0.3856***	-0.1258***
	(0.0058)	(0.0219)	(0.0121)	(0.0239)	(0.0639)	(0.0204)
Treat × Relative period = -3	-0.0139**	-0.1285***	-0.0502***	-0.0413	-0.1534**	-0.0566**
	(0.0063)	(0.0247)	(0.0129)	(0.0260)	(0.0724)	(0.0236)
Treat × Relative period = -2	-0.0044	-0.0669***	-0.0124	0.0088	-0.1066	-0.0117
	(0.0064)	(0.0253)	(0.0123)	(0.0258)	(0.0772)	(0.0216)
Treat × Relative period = 0	-0.0023	0.0459*	0.0160	0.0636**	0.0867	-0.0090
	(0.0070)	(0.0250)	(0.0133)	(0.0274)	(0.0722)	(0.0250)
Treat × Relative period = 1	-0.0007	0.0915***	0.0342**	0.0848***	0.0150	-0.0044
	(0.0076)	(0.0256)	(0.0147)	(0.0296)	(0.0756)	(0.0264)
Treat × Relative period = 2	0.0025	0.1160***	0.0528***	0.0403	0.0833	-0.0060
	(0.0081)	(0.0263)	(0.0140)	(0.0329)	(0.0751)	(0.0254)
Treat × Relative period = 3	-0.0061	0.1246***	0.0535***	-0.0044	0.0677	-0.0213
	(0.0087)	(0.0281)	(0.0159)	(0.0375)	(0.0784)	(0.0291)
Treat × Relative period = 4	-0.0375***	-0.0312	0.0483***	0.0997***	0.0546	0.0049
	(0.0074)	(0.0251)	(0.0146)	(0.0289)	(0.0668)	(0.0235)
Treat × Time to treatment	-0.0087***	0.0324***	0.0006	-0.0233***	-0.0073**	-0.0019
	(0.0004)	(0.0017)	(0.0013)	(0.0019)	(0.0030)	(0.0014)
Fixed-effects						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	428,073	385,055	108,104	87,822	72,964	64,366
R^2	0.58924	0.81672	0.90454	0.72946	0.81238	0.81465
Within R^2	0.00329	0.00454	0.00290	0.00529	0.00190	0.00137

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Figure 2.20: Direct impact of R&D funding on firm-level and innovation outcomes



2.5.3.2 Approach 2 - Pre-period analysis and de-trending

In **step 1**, I adjust for a linear pre-trend by including all control and treated firms up to the year of treatment. Table 2.19 confirms the existence of a positive and significant pre-trend impacting all outcomes among both SMEs and large firms. These findings imply systematic variations in the trends of outcome variables between the treatment and control groups.

Table 2.19: Parametric event study in pre-treatment period

	SMEs			Large firms		
Dependent Variables:	Patent (quality-adjusted)	Fixed assets	Employment costs	Patent (quality-adjusted)	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Treat × Time to treatment	0.0012***	0.0702***	0.0124***	0.0084***	0.0296***	0.0144***
	(0.0003)	(0.0019)	(0.0013)	(0.0015)	(0.0042)	(0.0017)
Fixed-effects						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	353,807	322,198	81,281	66,692	56,505	48,767
R^2	0.60605	0.81348	0.91296	0.72165	0.80384	0.78951
Within R^2	6.42×10^{-5}	0.00466	0.00234	0.00060	0.00101	0.00215

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Step 2: Table 2.20 presents the results of the parametric event study after detrending using the interaction term "Treat × Time to treatment" for both SMEs and large firms across different dependent variables. This exercise leaves the coefficients for innovation outcomes in both SMEs and large firms largely unaffected.

Table 2.20: Parametric event study: Detrending (Treat × Time to treatment)

	SMEs			Large firms		
Dependent Variables:	Patent (quality-adjusted)	Fixed assets	Employment costs	Patent (quality-adjusted)	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Treat × Relative period = -4	-0.0606***	-0.3111***	-0.1067***	-0.2176***	-0.3856***	-0.1258***
	(0.0058)	(0.0219)	(0.0121)	(0.0239)	(0.0639)	(0.0204)
Treat × Relative period = -3	-0.0139**	-0.1285***	-0.0502***	-0.0413	-0.1534**	-0.0566**
	(0.0063)	(0.0247)	(0.0129)	(0.0260)	(0.0724)	(0.0236)
Treat × Relative period = -2	-0.0044	-0.0669***	-0.0124	0.0088	-0.1066	-0.0117
	(0.0064)	(0.0253)	(0.0123)	(0.0258)	(0.0772)	(0.0216)
Treat × Relative period = 0	-0.0023	0.0459*	0.0160	0.0636**	0.0867	-0.0090
	(0.0070)	(0.0250)	(0.0133)	(0.0274)	(0.0722)	(0.0250)
Treat × Relative period = 1	-0.0007	0.0915***	0.0342**	0.0848***	0.0150	-0.0044
	(0.0076)	(0.0256)	(0.0147)	(0.0296)	(0.0756)	(0.0264)
Treat × Relative period = 2	0.0025	0.1160***	0.0528***	0.0403	0.0833	-0.0060
	(0.0081)	(0.0263)	(0.0140)	(0.0329)	(0.0751)	(0.0254)
Treat × Relative period = 3	-0.0061	0.1246***	0.0535***	-0.0044	0.0677	-0.0213
	(0.0087)	(0.0281)	(0.0159)	(0.0375)	(0.0784)	(0.0291)
Treat × Relative period = 4	-0.0375***	-0.0312	0.0483***	0.0997***	0.0546	0.0049
	(0.0074)	(0.0251)	(0.0146)	(0.0289)	(0.0668)	(0.0235)
blue!10 Treat × Time to treatment	-0.0099***	-0.0378***	-0.0118***	-0.0317***	-0.0369***	-0.0163***
	(0.0004)	(0.0017)	(0.0013)	(0.0019)	(0.0030)	(0.0014)
Fixed-effects						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	428,073	385,055	108,104	87,822	72,964	64,366
R^2	0.58811	0.81778	0.90280	0.72461	0.80937	0.80913
Within R^2	0.00402	0.00217	0.00283	0.00813	0.00289	0.00372

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

2.6 Conclusions

In this study, the focus is on examining the influence of R&D funding on both firm performance and innovative activity. By leveraging a comprehensive dataset encompassing firm-level information, patent data, and funding records spanning from 1970 to 2020, the analysis endeavors to shed light on the effects of R&D funding on the outcomes of funded firms. Additionally, the study aims to offer insights into the evolving patterns of R&D funding in Germany over time.

The initial phase of the analysis reveals that in Germany, large multinational corporations with a history of early patenting, focused on applied research, and operating in high-tech sectors have a greater likelihood of receiving funding. This aligns with previous research findings (Arrow 1962; Blanes and Busom 2004; Busom 2000). In the second part of the analysis, I delve into examining the direct impacts of R&D funding on firm performance and innovation outcomes. The findings from the standard Two-Way Fixed Effects (TWFE) analysis suggest that the policy had a positive impact only on the firm-level outcomes of SMEs, while no significant effects were observed for large firms. The findings from this study are consistent with prior research such as Czarnitzki and Delanote 2012, which also reported positive effects of R&D funding on firm-level outcomes for SMEs. However, they diverge from the conclusions of existing studies (Czarnitzki and Delanote 2015; Czarnitzki and Hottenrott 2011; Howell 2017; Kleine et al. 2022; Lerner 1999), which have typically found a positive association between R&D funding and innovation outcomes. However, these results are potentially biased due to the presence of pre-existing trends, as identified through sensitivity analysis using the smoothness restriction approach of the Honest DID method. Despite attempts to address this bias by detrending, the analysis did not completely eliminate the pre-existing linear trend, indicating that the estimated treatment effects may be influenced by unobserved factors (i.e. external economic conditions, industry-specific dynamics, or other policy interventions that coincide with the timing of the R&D funding) affecting the outcomes over time. This highlights the importance of accounting for and mitigating potential confounding variables or unobserved trends when estimating treatment effects. Further investigation into the nature and sources of these pre-existing trends is warranted to better understand the true impact of the treatment on the outcomes of interest.

CHAPTER 3

Subsidizing R&D in Germany: Direct and Spillover Effects

*w*ith Tobias Büscher

Chapter Abstract

We estimate the effects of R&D funding on firm performance and innovation activity in Germany for the period 2000-2020. Since R&D collaborations between funded and non-funded firms can lead to positive externalities, our study accounts not only for the direct effects of the policy on funded firms but also for indirect (spillover) effects on non-funded firms. We use multiple databases (DAFNE, PATSTAT, Förderkatalog) and conduct empirical estimations on both firm and cluster levels. Our findings suggest that, in general, R&D funding has a positive impact on both funded and non-funded firms, which varies with the share of funded firms in the cluster. Interestingly, the indirect impact on innovations is larger than the direct impact. The latter finding points out the advantages of this broad type of R&D funding (no sectoral or regional concentration) in enforcing innovation spillovers, however weakness in stimulating more pronounced direct effects.

3.1 Introduction

Recently, considerable effort has been made for the development of R&D and innovation policies worldwide. This growing trend is justified by the positive effects of private sector R&D on sustainable growth (Aghion and Howitt 2009; Doraszelski and Jaumandreu 2013; Griliches 1979; Grossman and Helpman 1994), innovation (increased number of patent applications and new products) (Czarnitzki and Delanote 2015; Czarnitzki and Hottenrott 2011; Czarnitzki and Licht 2006; David et al. 2000; Howell 2017), and additional R&D expenses (Aerts and Schmidt 2008; Almus and Czarnitzki 2003; Czarnitzki and Fier 2002; Czarnitzki and Hussinger 2004; Czarnitzki et al. 2007; Hussinger 2008). However, R&D funding can also lead to the crowding out of private investments by public funding (Montmartin and Herrera 2015; Wallsten 2000). R&D funding can affect firms from two perspectives: it will either stimulate firms to increase innovation activity due to the enhanced capability to afford innovation costs or it will give them a possibility to use the R&D subsidy for their own funding because they would have implemented their R&D projects anyway.

In this paper, we want to contribute to the existing debate on the effects of R&D funding. We study the causal impact of R&D funding on the performance and innovation outcomes of German firms from 2000-2020. The main research questions are (1) whether R&D funding improves firm performance and innovative output, (2) whether there are any spillovers for non-funded firms due to the presence of funded firms. The latter question is inspired by the work of earlier scholars, who claim that the presence of FDI firms (Girma et al. 2014) can generate positive externalities via collaboration.

For the first time, we integrate multiple datasets by leveraging the comprehensive German Förderkatalog dataset (includes information on all R&D projects funded in Germany) and linking it with the DAFNE database and the EPO Worldwide Patent Statistical Database (PATSTAT). Our empirical strategy includes several steps. First of all, to address the selection problem implying that the firms with certain characteristics were selected for the funding, we implement matching on several covariates (fixed assets, number of employees, firm age, and industry dummy) on the firm as well as the cluster level. We define as clusters unique industry-region combinations (by using German NUTS1 regions (16 Bundesländer) and NACE1 industry classification (9 broad industries)) resulting in 144 unique clusters. Then, we perform firm- and cluster-level estimations to obtain the direct, indirect, and total effects of funding for different proportions of funded firms in the cluster following Girma et al. (2014).

Although there is a broader literature on R&D policy evaluations, most of the papers analyzing the effect of R&D subsidies either focus on the general effect of funding on the funded firms or collaboration and spillover effects without disentangling the direct from the indirect effects. With this paper, we want to add more insights to the existing literature by making a clear distinction between direct and indirect (spillover) effects of R&D funding on the firms' financial and innovation outcomes, and understanding what drives the overall effect. To the best of our knowledge, only the study by Czarnitzki et al. (2007) looks at both effects by regressing innovation activity on collaboration and separately on R&D funding. Their results suggest that there is no direct effect of R&D funding on private R&D. In contrast, they find that collaboration increases the firms' R&D activity and might generate positive spillovers for other non-funded firms, which drives the overall positive effects of R&D

funding on innovation. However, the latter study does not explore how the effect of R&D varies when we have different shares of treated firms in the cluster. Our key contribution lies in using the share of funded firms as a mechanism to identify spillovers, thereby uncovering insights into the optimal share of funded actors and relevant thresholds to enhance spillover effects.

The remainder of the paper is structured as follows. In the following section, we present the existing literature. Section 3.3 presents the data and the main descriptive statistics. The empirical estimation strategy is explained in Section 3.4. The main findings are discussed in Section 3.5. Finally, Section 3.6 concludes.

3.2 Literature Review

Traditionally, economists ascribe economic growth to technological progress, which can be stimulated via research and development (R&D) funding (Romer 1990). The latter one is very important, because it develops a systematic approach towards innovation and stimulates the diffusion of knowledge and interactive learning among different public and private actors (Clausen 2009).

Another justification for public funding of R&D is the existence of two types of market failures. The first one is related to the classical market failure argument by Arrow (1962) which states that firms have fewer incentives to invest in R&D. The main reason is the existence of knowledge externalities and incomplete appropriability, which makes it difficult to claim the ownership rights and allows free-riding by competitors. The second type of market failure implies that due to very high costs of external capital, firms do not have sufficient incentives to invest in R&D from the societal point of view, thus some innovations will not be developed. This problem is particularly severe for small, young, and cash-limited firms (Hall 2002a; Hall 2002b).

To fill this gap between private and social return rates to R&D investments, governments design different R&D initiatives, which are usually offering either direct funding or a tax incentive (Hall 2002a; Hall 2002b) by stimulating private R&D investments (Jaffe 2002). Existing research analyzes the impact of R&D funding from two different perspectives by concentrating on the direct and spillover effects. The overwhelming majority of studies tackling the direct effects of R&D policies focus on input additionally. Some of them claim that R&D funding increases R&D expenditures of funded firms (Almus and Czarnitzki 2003; Czarnitzki and Fier 2002; Czarnitzki and Licht 2006; Duguet 2004; González et al. 2005; Hud and Hussinger 2015; Hussinger 2008; Li and Bosworth 2018) by supporting the R&D complementarity hypothesis, while the others argue that it leads to the crowding out of R&D spending of funded firms (Montmartin and Herrera 2015; Wallsten 2000). Another strain of research concentrates on output additionality and provides positive evidence of R&D subsidies on innovation (Aschhoff 2009; Bronzini and Piselli 2016; Czarnitzki and Fier 2002; Czarnitzki and Hussinger 2004; Czarnitzki and Hussinger 2017; Czarnitzki and Licht 2006; Czarnitzki and Lopes-Bento 2011; Kleine et al. 2022) and collaboration (Czarnitzki et al. 2007; Kleine et al. 2022). Particularly, the importance of R&D collaboration raised tremendously, which is seen by the firms as a possibility to access new resources, and knowledge (Badillo and Moreno 2015; Kaiser 2002; Li and Bosworth 2018), and share the risks and costs of R&D projects (Antonioli et al. 2016; Belderbos et al. 2004) by leading to the improvement of firm-level outcomes (Falk 2010; Hall et al. 2009; Lerner 1999).

However, R&D collaborations do not necessarily occur between the firms funded within the same scope of certain R&D policies but can be established between different non-funded actors by thus generating positive spillovers and indirect effects of the funding for non-funded firms. This brings us to the second strain of recent literature concentrating on the spillover effects of R&D policies.

In general, inter-firm spillovers are the result of knowledge exchange between scientists (Jaffe 1986). Due to the limited mobility of scientists (Breschi and Lissoni 2009), knowledge exchange is more likely to be stimulated and successfully transferred at a close distance (Audretsch and Feldman 1996; Jong and Freel 2010; Spithoven et al. 2019) by generating geographical spillovers. The latter is the most studied direction in the spillover literature and is found to lead to positive externalities (Acemoglu et al. 2016; Breschi and Lissoni 2005; Petruzzelli 2011). For instance, Petruzzelli (2011) asserts that R&D collaboration at a close geographic distance leads to more innovation.

However, proximity is not only constrained to a spatial context. There is a higher probability of knowledge exchange not only in geographic proximity but also in the same (Bernstein and Nadiri 1989) or more related industries (Jaffe et al. 1993). Bloom et al. (2013) support the idea that technological connectedness leads to more positive spillovers and higher social returns of R&D. They also add that the impact is lower for small firms, which is the reason for their operation in "technological niches" and fewer linkages to other firms in technology space. Further, Li and Bosworth (2018) verifies in the study of the British firms, that own-sector R&D spillovers have the largest and most significant effect on labor productivity.

Building on these concepts, additional studies suggest that agglomeration economies (such as region-industry targeted clusters) can be more advantageous for the growth of innovative firms due to economies of scale and spillover effects (Akcigit et al. 2018; Breschi and Lissoni 2009; Crass et al. 2017; Engel et al. 2017; Nishimura and Okamuro 2010; Uyarra and Ramlogan 2016).

Finally, the study by Czarnitzki et al. (2007) looks both at the direct and indirect (spillover) effects of R&D funding. They argue that R&D collaborations have a more pronounced impact on firms' R&D activity through positive spillovers than direct funding itself.

3.3 Data

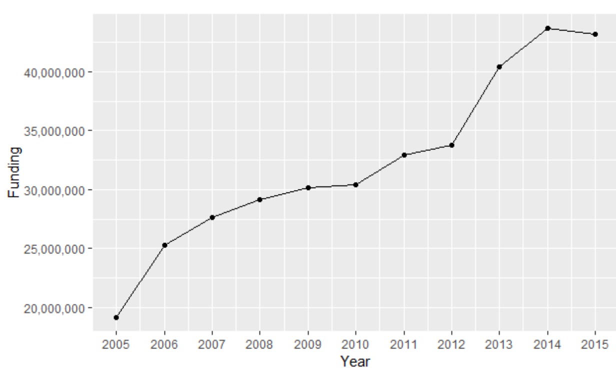
Our final dataset results from linking three types of data sources on the firm level. Firm financial information is taken from DAFNE database (extracted in September 2020). It is geo-referenced data, which provides information on firm location, sector, financial outcomes (e.g. fixed assets, employment costs, sales, revenue, etc.), legal entity details, corporate structures, ownership, etc. In our analysis, we concentrate on the sub-sample of German firms (around 2 million firms) over the period of 2000-2020.

DAFNE data is linked to the EPO Worldwide Patent Statistical Database (PATSTAT), which contains detailed information on all patented innovations including the names and addresses of applicants and inventors, priority year, patent families, the title and abstract of patent applications, citation links, and patents classification by technology class. We restrict PATSTAT data also to the patent applications by German firms.

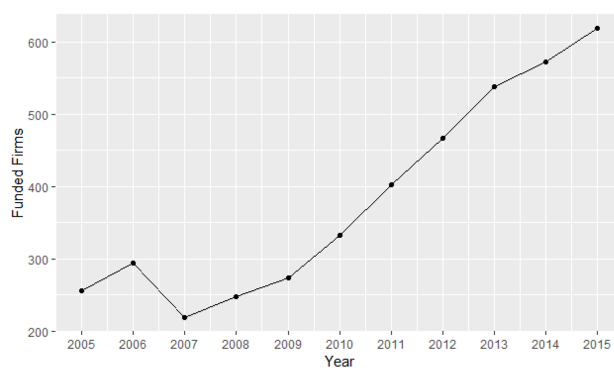
Finally, the information on the direct R&D funding of firms by the German Federal Government is taken from the Förderkatalog (funding catalog) database, which is jointly created by the Federal Ministry of Education and Research (BMBF) and the Federal Ministry of Economics and Technology (BMWI). The Förderkatalog is a comprehensive, unique data that contains detailed information (i.e. funding type, amount, duration, area, recipient groups, responsible institutions/federal states/countries) on 90,000 individual R&D projects that have been funded since 1970. The growing number of funding recipients from 1970 to 2020 led to 28,013 firms with a total budget of 44 bn € (BMBF (2023)). It's noteworthy that this extensive dataset, encompassing all R&D-funded projects in Germany, has not been previously utilized.

Dependent variables: Firm-level outcomes are divided into the firm's fixed assets and employment (number of employees). Assessing innovation outcomes often relies on measuring patents, but this poses challenges as patents vary in their commercial value and technological impact. Research suggests that while most patents hold little value, a select few represent significant technological breakthroughs (Gambardella et al. 2008). When explicit value information is lacking, citation-based measures offer the best approximation of patent quality (Gambardella et al. 2008; Reitzig 2004). Patents that are frequently cited by subsequent patents are generally considered higher quality, provided that differences in citation propensity are controlled for (Harhoff et al. 1999; Jaffe and de Rassenfosse 2016; Trajtenberg 1990).

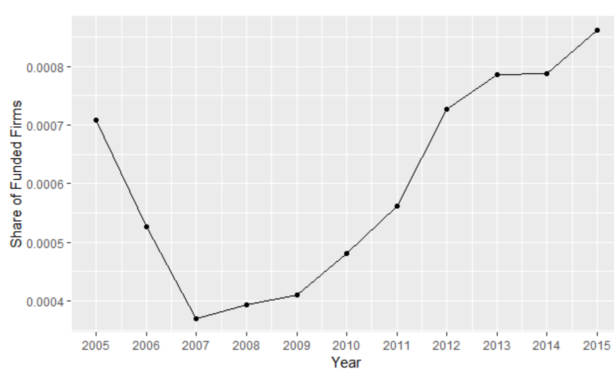
Figure 3.1: Descriptive statistics for funding and patents



(a) Funding increase over time



(b) Funding recipients increase over time



(c) Share of funded firms over time

However, comparing citation counts across countries has limitations. For one, applicants tend to protect more valuable patents abroad (due to the elevated expenses associated with filing and translating patents), making direct comparisons between domestic and foreign applications less informative (Harhoff et al. 2003). Additionally, variations in examination practices across national patent offices result in significant differences in citation counts. Furthermore, patent examiners commonly demonstrate a tendency to favor citing patents from their native country (Michel and Bettels 2001). To address these challenges, we adopt a patent quality measure proposed by Boeing and Mueller (2016) that enables technology-specific cross-country comparisons and incorporates forward citations. This measure accounts for factors such as the number of technology classes, family size (number of countries in which a patent application was filed), and the number of forward citations received by the patent.

Figure 3.1 shows summarising graphs for funding. Both funding (a) as well as funding recipients[1] (b) increased over the years of observation. The funding programs were widened and R&D funding increased in general. Graph (c) shows the decrease in the share of funding recipients/funded firms from 2005 to 2007 due to a slight decrease in funding recipients and an increase in the total number of firms.

Table 3.1: Summary statistics on firm-level characteristics

	Funded (N=207164)	Non-Funded (N=12498684)	Overall (N=12705848)
Firm age	27.2 (31.6)	17.7 (22.3)	17.9 (22.5)
Firm size	2.22 (1.74)	0.805 (0.972)	0.834 (1.01)
Patent (simple)	19.8 (170)	5.42 (85.1)	9.12 (113)
Patent (adjusted)	52.0 (457)	14.1 (214)	23.8 (296)
Fixed assets	107 (1560)	6.12 (303)	8.17 (373)
Total assets	170 (2320)	9.99 (448)	13.2 (554)
Turnover	270 (2420)	20.5 (536)	27.7 (670)
Sales	259 (2430)	23.5 (465)	33.7 (682)
Number of employees	442 (4430)	30.9 (726)	39.1 (953)
Labor costs	46.6 (390)	7.67 (330)	10.9 (336)

Notes: Fixed and total assets, turnover, sales, and labor costs values are in Mln.

Table 3.1 presents summary statistics for variables of interest categorized by R&D funding status. The first column illustrates the average outcomes of funded firms, the second column - the average outcomes of non-funded firms, and the third column - the average overall outcome from 1970 to 2020. Firm-level metrics encompass firm age (computed as the difference between the current year and the year of incorporation), firm size (log of total assets), simple and quality-adjusted number of patents, fixed and total assets, turnover, sales, number of employees, and labor costs. Significantly, substantial discrepancies emerge between R&D funded and non-funded firms, underscoring the necessity of adopting a treatment effects evaluation framework. The raw data indicates that, on average, funded firms are approximately 10 years older and twice as large as non-funded firms. Funded firms also boast around four times more patents and greater firm capital and labor. Regarding industry distribution, funded firms tend to dominate in knowledge-intensive sectors such as Education, Professional, Scientific and Technical Activities, Information and Communication, etc. (refer to Table 3.2).

[1]Funding recipient means the firm that received the funding.

Table 3.2: Summary statistics on industry characteristics

Industry	Funded (N=207164)	Non-Funded (N=12498684)	Overall (N=12705848)
Accommodation and Food Service Activities	897 (0.4%)	138562 (1.1%)	139459 (1.1%)
Administrative and Support Service Activities	914 (0.4%)	136456 (1.1%)	137370 (1.1%)
Arts, Entertainment and Recreation	1042 (0.5%)	46172 (0.4%)	47214 (0.4%)
Construction of Buildings	1932 (0.9%)	417631 (3.3%)	419563 (3.3%)
Education	5019 (2.4%)	97475 (0.8%)	102494 (0.8%)
Financial and Insurance Activities	509 (0.2%)	182500 (1.5%)	183009 (1.4%)
Human Health and Social Work Activities	2729 (1.3%)	84619 (0.7%)	87348 (0.7%)
Information and Communication	3067 (1.5%)	91675 (0.7%)	94742 (0.7%)
Manufacturing	3118 (1.5%)	62820 (0.5%)	65938 (0.5%)
Other Service Activities	2870 (1.4%)	129965 (1.0%)	132835 (1.0%)
Professional, Scientific and Technical Activities	4677 (2.3%)	350926 (2.8%)	355603 (2.8%)
Public Administration and Defense: Compulsory Social Security	1569 (0.8%)	17802 (0.1%)	19371 (0.2%)
Real Estate Activities	3450 (1.7%)	1060267 (8.5%)	1063717 (8.4%)
Transportation and Storage	1668 (0.8%)	84346 (0.7%)	86014 (0.7%)
Water Supply	862 (0.4%)	18567 (0.1%)	19429 (0.2%)
Wholesale and Retail Trade	5749 (2.8%)	819881 (6.6%)	825630 (6.5%)
Activities of households as employers	0 (0%)	290 (0.0%)	290 (0.0%)

As will be discussed in detail in the next section, our identification strategy relies on the partial interference assumption, where firm interaction is freely allowed within well-defined clusters. In this study, we categorize firms into clusters based on 16 geographic areas and 9 broadly defined industries (detailed explanation provided in section 3.4 and the list of regions and industries in Appendix C.1). This classification yields 144 clusters, each comprising an average of 1438 funded firms. Table 3.3 presents summary statistics of cluster-level variables. The average proportion of funded firms within a cluster is approximately 2.5 percent. Moreover, it is evident from the various measures of dispersion presented in Table 3.3 that clusters exhibit considerable heterogeneity in terms of average characteristics. Table 3.4 showcases the number of funded firms and the proportion of funded (treated) firms by cluster indicating substantial heterogeneity across clusters.

Table 3.3: Summary statistics on some cluster-level characteristics

	Mean	SD	Min	Max
Number of funded firms	1438	1905	12	12090
Proportion of funded firms	0.025	0.022	0.002	0.126
Average firm age	17.8	22.4	-249	1014
Average firm size	0.83	1.01	-1.81	11.8
Average patents (simple)	9.11	113.4	1	7116
Patent (adjusted)	23.8	296.2	0	21151.8
Fixed assets	8.16	373.3	-15.57	456537.5
Total assets	13.2	553.5	-6.89	668236.2
Turnover	27.7	670.4	-164.2	221674.2
Sales	33.7	681.8	-4.23	190262.7
Number of employees	39.1	953.4	1	420000
Labor costs	10.9	335.76	-53.46	293906.7
Number of clusters	144			

Notes: Fixed and total assets, turnover, sales, and labor costs values are in Mln.

Table 3.4: The number and the proportion of funded (treated) firms by clusters

C	NF	PF	C	NF	PF	C	NF	PF	C	NF	PF
1	12	0.002	37	2619	0.035	73	1242	0.0098079	109	2968	0.054
2	12090	0.055	38	317	0.016	74	763	0.0241609	110	2110	0.006
3	2270	0.021	39	303	0.009	75	523	0.0537623	111	461	0.09
4	330	0.017	40	3696	0.058	76	976	0.0100608	112	310	0.022
5	126	0.013	41	339	0.005	77	1807	0.0219408	113	2086	0.011
6	652	0.017	42	1189	0.014	78	1022	0.0053821	114	565	0.004
7	1219	0.05	43	1008	0.006	79	269	0.0149769	115	349	0.017
8	349	0.021	44	899	0.029	80	1653	0.0834469	116	322	0.014
9	1108	0.06	45	3453	0.008	81	371	0.0237683	117	449	0.009
10	174	0.014	46	1126	0.023	82	4341	0.0089880	118	101	0.01
11	207	0.012	47	332	0.017	83	51	0.0124299	119	504	0.03
12	1246	0.021	48	378	0.008	84	447	0.0247152	120	3968	0.045
13	453	0.006	49	504	0.025	85	3522	0.0141429	121	503	0.014
14	4904	0.01	50	738	0.024	86	996	0.0299405	122	71	0.006
15	674	0.036	51	1111	0.008	87	1104	0.0287134	123	1032	0.021
16	860	0.024	52	404	0.02	88	615	0.0217437	124	534	0.021
17	2130	0.054	53	4507	0.014	89	361	0.0384124	125	718	0.074
18	446	0.007	54	794	0.032	90	313	0.0110445	126	805	0.008
19	2550	0.034	55	955	0.008	91	1710	0.0103238	127	1210	0.0143984
20	711	0.009	56	298	0.005	92	6288	0.0068924	128	475	0.0063220
21	284	0.027	57	577	0.019	93	634	0.0372350	129	5478	0.1264134
22	211	0.008	58	8083	0.057	94	912	0.0805938	130	2588	0.0110198
23	295	0.01	59	206	0.02	95	596	0.0417425	131	3799	0.0549195
24	11547	0.071	60	1163	0.015	96	1573	0.0105336	132	867	0.0149550
25	239	0.044	61	739	0.004	97	2715	0.0073573	133	5198	0.0076556
26	948	0.02	62	1645	0.094	98	468	0.0207898	134	228	0.0244530
27	3182	0.115	63	205	0.008	99	72	0.0089186	135	2877	0.0212318
28	121	0.009	64	1414	0.053	100	229	0.0126394	136	762	0.0603803
29	759	0.016	65	823	0.018	101	387	0.0148395	137	1509	0.0292187
30	670	0.048	66	321	0.031	102	401	0.0309318	138	160	0.0306279
31	458	0.041	67	356	0.075	103	779	0.0123749	139	4313	0.0091855
32	6516	0.01	68	702	0.006	104	1546	0.0216739	140	246	0.0109362
33	511	0.005	69	107	0.039	105	2207	0.0097025	141	2152	0.0259437
34	4231	0.021	70	1292	0.058	106	1434	0.0197360	142	2660	0.0194057
35	481	0.006	71	1828	0.03	107	1264	0.0061842	143	347	0.0070236
36	646	0.007	72	2345	0.014	108	262	0.0198050	144	200	0.0075259

Note: C = Cluster (region-sector) number; NF = Number of funded firms;
PF = Proportion of funded (treated) firms.

3.4 Empirical Framework

3.4.1 Identification Strategy

The estimation of direct and spillover effects has an inherent identification problem. The Stable Unit Treatment Value Assumption (SUTVA) is violated by construction. For the assumption to be reasonable, the outcome of units must be independent of the treatment status of the other units. If we now consider spillover effects, such that the firms benefit from other firms' increased R&D activity in response to the funding, then SUTVA is violated. Additionally, the funding of firms strengthens their position in a competitive market with other funded and non-funded firms causing changes in the competition of the firms due to the funding. Therefore, for SUTVA to hold we can only compare firms that do not have any innovation spillovers and that do not compete with each other in a common market.

Second, public policy programs usually target certain firms which raises a selection problem in the identification. In this case, there is selection on two levels. The authority deciding who receives the funding might be more interested in specific regions or in specific industries. This might be due to future expectations about the economic outcomes in the industry or region or due to personal preference. Furthermore, the firm selection into the treatment is not random but the result of a decision process. It is a decision both from the firm side to apply for funding and from the side of a policy maker to choose a specific firm. The third identification problem comes from region-industry-specific effects. Industry-specific productivity shocks can affect productivity and the innovation outcomes from innovation inputs such as R&D expenditure. This might lead to treatment as a response to the shock and biased estimates.

Since standard approaches are not able to tackle aforementioned issues, we introduce a non-standard approach by combining the methods suggested by Girma et al. (2014) and Hudgens and Halloran (2008) and Fong et al. (2018). First of all, to consistently estimate both direct and spillover effects, we relax the SUTVA assumption as suggested by the recent econometric literature (Baird et al. 2018; Ferracci et al. 2014; Girma et al. 2014; Hudgens and Halloran 2008; Rosenbaum 2007). As per this literature, treatment externalities, also known as spillovers, occur when an individual's potential outcome is influenced by the treatment status of others within a group. As the effect of the funding is not limited to funded firms, the treatment pollutes the control units. As the funding does not affect all control units, but only groups of them, we refer to partial interference. In contrast to standard methods controlling for the funding of neighboring firms, we can form groups into the spatial dimension as well as the industry dimension to also control for the effect of funding on technologically close non-treated firms. This is important to capture knowledge spillovers and effects on competition in markets as two potential sources of spillover effects. Furthermore, defining clusters and allowing the firms within clusters to interact enables us to perform matching on an additional level, the cluster level. Following the approach outlined by Hudgens and Halloran (2008), we utilize the proportion of treated firms within a group as a metric for interaction among individual firms. Consequently, potential outcomes are modeled as a function of the firm's treatment status and the fraction of R&D funded firms within a specific group or cluster.

Following Girma et al. (2014), we conduct estimations at two distinct levels: the firm level and the cluster level, with clusters defined as unique region-industry combinations. Our main identification assumption is the Partial Interference Assumption, where we posit that the SUTVA holds across clusters but not within clusters. This implies that the potential outcomes of an observed unit are unaffected by the treatment status of other units in other clusters. To satisfy this assumption, we require a cluster definition that provides a sufficient number of clusters with substantial variation to facilitate estimations while ensuring diverse levels of interaction within the clusters. On the other hand, it's essential to minimize interaction in innovative activity and competition across the clusters as much as possible. For this purpose, we utilize region-industry combinations with German NUTS 1 regions (Bundesländer) and the broad NACE 1 industry classification. Our Partial Interference Assumption dictates that interactions occur solely between firms within the same industry and region. This approach aligns with findings in the literature on innovative spillovers, which indicate a sharp spatial decay and limited interactions within broader industries (Acemoglu et al. 2016; Breschi and Lissoni 2005; Glaeser et al. 1992). Moreover, this definition, rather than relying solely on industry classification or regions, offers a sufficiently broad range of clusters. The 9 NACE 1 industries and 16 NUTS 1 regions in Germany give us 144 unique combinations of broad industries and regions. An illustrative example of a cluster could be the manufacturing industry in North Rhine-Westphalia (NRW) or the Information and Communication industry in Bavaria.

3.4.2 Estimation Strategy

We use the empirical methodology suggested by the literature on causal estimands under partial interference that was introduced by Hudgens and Halloran (2008) and shifted to infinite populations by Baird et al. (2018). As suggested by Girma et al. (2014), we perform the first part of the analysis separately for each cluster and then shift the analysis to the cluster level. To give a brief overview, the methodology we use involves the following steps:

Steps on the firm level (separately for each cluster):

1. **Firm-Level Matching**
   We obtain inverse probability weights by regressing the binary treatment variable (1 if a firm was funded, 0 otherwise) on pre-treatment firm characteristics.

2. **Computing cluster-specific treatment effects**
   We regress the firm-level outcome on the binary treatment variable and control variables.

3. **Computing potential outcomes**
   We obtain the potential outcomes based on the estimated coefficients from step 2 and the corresponding observations for those variables. We calculate cluster-average potential outcomes for treated and non-treated units.

Steps on the cluster level:

4. **Aggregating the variables**
   We compute the cluster-means for the firm-specific control variables and the binary treatment variable.

5. **Generalized Propensity Scores**
   We compute Generalized Propensity Scores following Fong et al. (2018) regressing the average binary treatment variable in a cluster on the average firm characteristics in a cluster.

6. **Computing coefficients for the treatment intensity and the GPS**
   We regress the cluster-average potential outcomes across treated firms and across non-treated firms on the Generalized Propensity Scores and the treatment intensity.

7. **Computing Cluster-potential outcomes**
   Based on the coefficients obtained from step 6, we compute the cluster-level potential outcomes for different levels of treatment intensities.

8. **Treatment effect calculation**
   We compute the direct, indirect and overall treatment effects comparing the cluster-level potential outcomes.

We start with the analysis on the firm level. Addressing the effectiveness of R&D funding poses several challenges, primarily stemming from the selective nature of such funding initiatives. Comparing funded and non-funded firms is inherently complex due to existing disparities in outcomes. These disparities give rise to various identification issues. Firstly, the endogeneity of treatment arises as firms may self-select into receiving R&D funding based on unobservable traits that also influence their outcomes. This self-selection introduces bias into estimates of the causal impact of R&D funding on firm performance and innovation outcomes.

Secondly, sample selection bias emerges as the sample of funded firms may not accurately represent the broader population of firms. This discrepancy can lead to biased estimates if treated and untreated groups differ systematically in ways that influence the outcomes of interest.

The latter can be observed in Table 3.5 which summarizes the means of the main variables for the treated firms and the non-treated firms before the first observed treatment. The outcome variables and covariates differ across treated firms and non-treated firms. Treated firms are larger in terms of number of employees and fixed assets. Also, they are slightly older (3 years) and the innovation outcomes (quality-adjusted patents) were already higher before the funding programs were rolled out. We conclude from differences in means between the treatment and control group that firms are non-randomly selected into treatment based on pre-treatment characteristics.

Table 3.5: Pre-treatment differences between treated and non-treated firms

Variables	Treated	Non-treated
Firm age	30	27
Yearly Patents (adjusted)	2.49	0.16
Fixed assets	69036	8012
Number of employees	332	38

Note: Fixed assets are in thousand dollars

Finally, the heterogeneous treatment effects add another layer of complexity as the impact of R&D funding may vary across firms based on diverse characteristics such as size, industry, and age. Failing to consider these differences can result in biased estimates of the average treatment effect.

To overcome these identification issues on the firm level, we perform inverse probability weighting on pre-treatment firm-level characteristics. This approach relies on the assumption of unconfoundedness, which states that conditional on the observed covariates, treatment assignment is independent of potential outcomes. Additionally, matching methods assume common support, which means that there are units in both the treatment and control groups with propensity scores overlapping across the entire range. These assumptions help ensure that the weighted samples are comparable, allowing for a more credible estimation of treatment effects (Hill and Su 2013; Huber 2013).

We define a treatment dummy d_{ict} on the firm level that is 1 if firm i in cluster c received funding in year t, and 0 otherwise. To compute the inverse probability weights, we perform the logistic regression from equation (3.1) for each cluster separately. We regress the treatment dummy d_{ict} on a vector X_{ict} of firm-level covariates before treatment. Those covariates include firm age, number of employees, fixed assets, and a dummy indicating whether the firm is a low-tech, medium-tech, or high-tech firm. The dummy is based on the NACE 2 industry classification of the firm. We exclude the respective outcome variable from the firm covariates in the logistic regressions.

$$P(d_{ict} = 1) = \frac{1}{1 + e^{-b_0 + b_1 X_{ict}}} \tag{3.1}$$

We obtain the inverse probabilities by calculating $\frac{1}{\hat{d_{ict}}}$ for treated firms and $\frac{1}{1-\hat{d}_{ict}}$ for non-treated firms. We use the generated propensity scores as inverse probability weights in the OLS regressions on the firm level.

Table 3.6: Determinants for selection into treatment

	Dependent variable:
	Treatment
Fixed Assets	0.169***
	(0.011)
Number of Employees	0.398***
	(0.016)
Firm Age	0.045
	(0.034)
High-Tech	0.035
	(0.080)
Low-Tech	−1.066***
	(0.057)
Constant	−5.950***
	(0.172)
Observations	39,595
Log Likelihood	−6,913.510
Akaike Inf. Crit.	13,839.020
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 3.6 shows the determinants of the treatment on the firm level. The size of the firm in terms of fixed assets and number of employees have a positive impact on selection into treatment and low-tech firms are less likely to receive the treatment compared to medium- or high-tech firms.

Using the inverse-probability weights from the logistic regressions for each cluster separately, we regress the outcomes on firm-level control variables X_{ict} and the treatment dummy d_{ict},

$$y_{ict} = \alpha + \beta d_{ict} + \delta X_{ict} + \epsilon; \quad i = 1, \ldots, N_{ct} \tag{3.2}$$

With y_{ict} being the firm-level (fixed assets, number of employees) and innovation (quality-adjusted number of patents) outcome of firm i in cluster c, X is the vector of firm-level covariates including firm age, number of employees, and fixed assets, and the error term ϵ. Here, we obtain separate estimates for the effect of receiving funding on the outcome.

In the third step, we calculate the potential outcomes separately for each cluster c and each year t. For that, we use the OLS estimates from equation (3.2) and the firm-level observations for control variables X_{ict}. We set the treatment dummy $d_{ict=1}$ for the potential outcome under treatment $\bar{y}^1_{ct}$ and $d_{ict=0}$ for the potential outcome under no treatment $\bar{y}^0_{ct}$. We aggregate those potential outcomes across the firms in a respective cluster to obtain the average potential outcomes for each cluster and year, for treated and untreated firms.

$$\bar{y}^1_{ct} = \frac{1}{N_{ct}} \sum_{i=1}^{N_{ct}} \hat{\alpha} + \hat{\beta}_c + \hat{\delta}_c X_{ict} \quad \text{and} \quad \bar{y}^0_{ct} = \frac{1}{N_c} \sum_{i=1}^{N} \hat{\alpha} + \hat{\delta}_c X_{ict} \tag{3.3}$$

After obtaining the potential outcomes for each cluster and year separately, we shift our methodology to the cluster level. To solve potential selection issues also on the cluster-level, we again calculate propensity scores. The treatment variable now needs to be adapted, as we now do not observe the treatment of each firm individually, but the overall treatment intensity within each cluster. Therefore we use the proportion of treated firms as suggested by Hudgens and Halloran (2008) in the cluster (p_c). Note that $p_c \in (0, 1)$ and is continuous. The proportion of treated firms in a cluster indicates the intensity of the funding within the cluster and at the same time it is a proxy for the probability of interaction of treated firms in a cluster.

$$p_{ct} = \frac{\sum_{i=1}^{N_{ct}} d_{ict}}{N_{ct}} \tag{3.4}$$

To check the validity of SUTVA assumption on cluster level, we observe the correlation between the share of treated firms in the cluster and other cluster-level variables (see Table 3.7). The clusters show heterogeneity regarding their firm composition and the selection of funding is likely to be non-random but in parts driven by differences in cluster-level characteristics. To overcome a potential selection bias on observed cluster-level variables, we match also on the cluster-level.

Table 3.7: Correlation matrix of cluster-level variables

	% Treated Firms	No. Employees	Fixed Assets	Firm Age	Pot. Outcome
% Treated Firms	1.00***	0.06***	-0.00	-0.22***	0.14***
No. Employees	0.06***	1.00***	0.08***	0.25***	0.51***
Fixed Assets	-0.00	0.08***	1.00***	0.04***	0.18***
Firm Age	-0.22***	0.25***	0.04**	1.00**	0.41***
Pot. Outcome	0.14***	0.51***	0.18***	0.41***	1.00***

Table 3.8: Pre-treatment in clusters

Variables	25% quantile p_c	50% quantile p_c	75% quantile p_c
Firm age	37.0	37.6	37.9
Yearly Patents (adjusted)	25.28	31.13	36.81
Fixed assets	63947	63364	83366
Number of employees	430	464	585

Note: Fixed assets are in thousand dollars

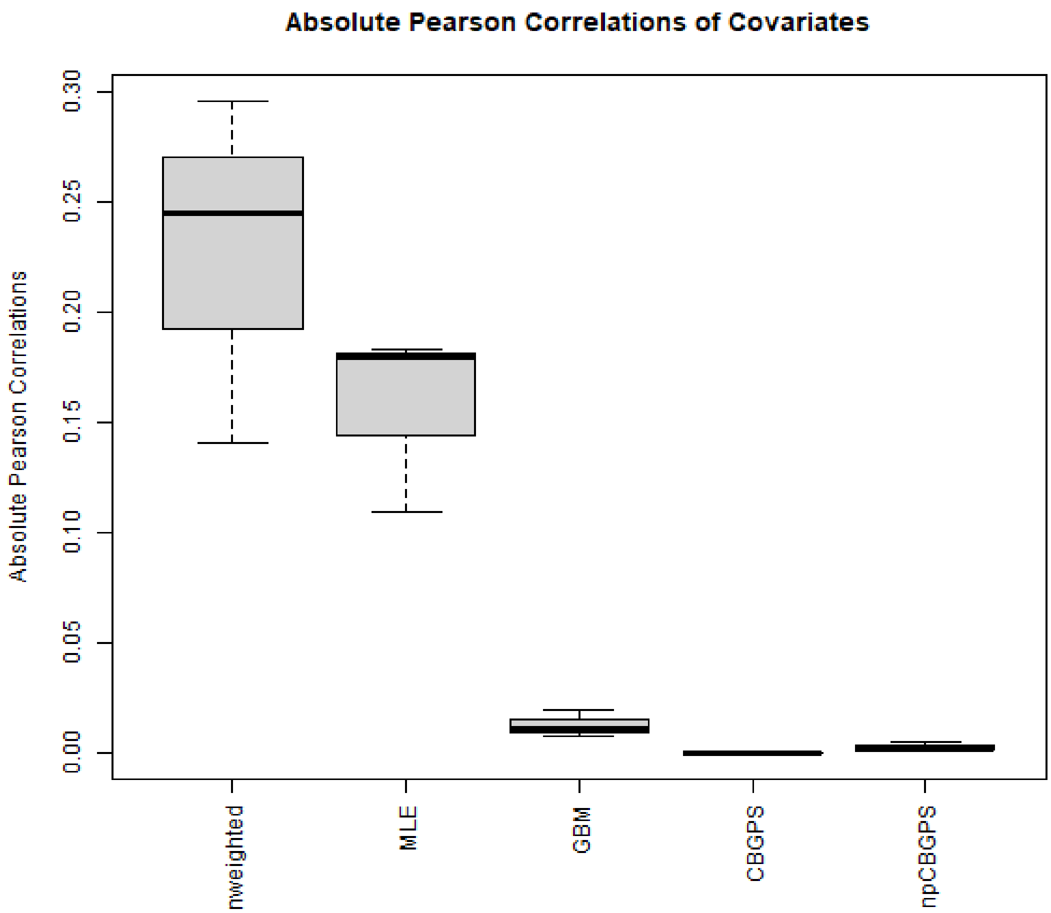
Table 3.8 shows the firm-level characteristics averaged over the cluster. We split by different treatment intensities p_{ct}. However, there are still observable differences in the characteristics between different levels of treatment. The clusters with a higher proportion of treated firms tend to be more capital intensive with a smaller number of employees and higher fixed assets. Additionally, they already had higher patent counts before the observed treatment. Also, they are slightly older. To account for those differences in the characteristics that increase the overall level of treatment in the cluster and at the same time increase average patent outcomes in the clusters, we perform matching also on the cluster level.

Due to the treatment variable now being continuous, we calculate the conditional density of the treatment using the Covariate Balancing Generalized Propensity Score estimation procedure suggested by Fong et al. (2018). We use the same covariates as in the firm-level matching, but now we average them across the firms in the cluster. As we did on the firm-level, we again exclude the outcome variable from the cluster-level covariates. We fit different estimators of the Generalised Propensity Score and pick the one with the best Covariate Balancing. The Covariate Balancing Generalized Propensity Score is the one with the best Covariate Balancing properties. See Appendix C.3 for more details on the generalized propensity score estimation procedure.

Figure C.1 shows the covariate balancing statistics for the generalized propensity score estimations. We use the propensity scores from the Covariate Balancing Generalised Propensity Score (CBPS) estimation in the further procedure, as it minimizes the Correlation of the covariates considerably compared to Generalized Boosted Models and Maximum Likelihood (MLE). The non-parametric version of the CBPS (npCBPS) achieves similar, but slightly worse Covariate Balancing.

We use the estimates of the conditional density of the continuous treatment $G\hat{P}S_c$ and the proportion of treated firms within a cluster $p_c t$ to approximate the expected values of the cluster-level potential outcomes under treatment and no treatment $\bar{y}_c^d t$ we generated in the firm-level estimation in equation 3.3. Following (Hirano and Imbens 2004), we approximate the potential outcomes under treatment $d = 1$ and potential outcomes under no treatment

Figure 3.2: Covariate balancing propensity score estimation



$d = 0$. We use polynomial approximations of the second-order

$$\bar{y}^d_{ct} = \beta_0 + \beta_1 G\hat{P}S_{ct} + \beta_2 p_{ct} + \beta_3 G\hat{P}S^2_{ct} + \beta_4 p^2_{ct} + \beta_5 G\hat{P}S_{ct} p_{ct}. \tag{3.5}$$

The approximation in (3.5) gives us estimates $\hat{\beta}_1$ to $\hat{\beta}_5$ to calculate the average potential cluster outcomes for different levels of proportion of treated firms $\bar{y}^d_p$, where p is the proportion of treated firms.

In the final step, we use a grid of possible values of p_c. We define a grid over the values from the minimum proportion of treated firms in a cluster to the maximum proportion. Therefore, the grid goes from 0 to 0.22 with 0.01 steps. For each of these values of the grid, we calculate the sample average potential outcomes across all clusters under treatment and no treatment. For this, we use the estimated coefficients from equation 3.5 and the conditional treatment density estimates $G\hat{P}S_c$. We follow

$$\bar{y}^d_p = \frac{1}{C} \sum_{r=1}^{C} (\hat{\beta}_0 + \hat{\beta}_1 G\hat{P}S_{ct} + \hat{\beta}_2 p_{ct} + \hat{\beta}_3 G\hat{P}S^2_{ct} + \hat{\beta}_4 p^2_{ct} + \hat{\beta}_5 G\hat{P}S_{ct} p_{ct}) \tag{3.6}$$

We obtain the cluster-level potential outcomes from (3.6) for all different intensities of treatment on the cluster-level. We can now make comparisons of those to calculate treatment effects. We distinguish three types of treatment effects: the direct effect, the spillover effect, and the total effect. For the direct effect γ_{pp}^{10}, we compare cluster potential outcomes with the same treatment intensity p_{ct} within the cluster and under treatment and no treatment. Therefore, we hold the interaction of funded firms in the cluster constant while we assume no interaction across clusters. The potential outcomes only differ due to the treatment status of the respective firm.

$$\bar{\gamma}_{pp}^{10} = \bar{y}_p^1 - \bar{y}_p^0 \tag{3.7}$$

With $\bar{\gamma}_{pp}^{10}$ being the direct treatment effect, $\bar{y}_p^1$ the average potential outcome under treatment for a cluster with the proportion of treated firms p and $\bar{y}_p^0$ the average potential outcome under no treatment for a cluster with the same proportion of treated firms p.

For the spillover effect γ_{p0}^{00}, we compare only non-treated cluster potential outcomes with different levels of the proportion of funded firms in the cluster. The outcomes only differ due to their different levels of treatment in the cluster. Therefore, there is a different level of interaction with funded firms in the compared clusters.

$$\bar{\gamma}_{p0}^{00} = \bar{y}_p^0 - \bar{y}_0^0 \tag{3.8}$$

With $\bar{\gamma}_{p0}^{00}$ being the spillover effect, $\bar{y}_p^0$ the average potential outcome under no treatment for a cluster with proportion of treated firms p and $\bar{y}_0^0$ the average potential outcome under no treatment for a cluster without treated firms, i.e. the proportion of treated firms $p = 0$.

The total effect can be decomposed into direct and indirect effects. It compares treated and untreated potential outcomes at different levels of the proportion of funded firms in the cluster.

$$\bar{\gamma}_{p0}^{10} = \bar{y}_p^1 + \bar{y}_0^0 \tag{3.9}$$

With $\bar{\gamma}_{p0}^{10}$ being the total effect, $\bar{y}_p^1$ the average potential outcome under no treatment for a cluster with proportion of treated firms p and $\bar{y}_0^0$ the average potential outcome under no treatment for a cluster without treated firms, i.e. the proportion of treated firms $p = 0$.

3.5 Results

Our main findings regarding the direct and spillover effects of R&D funding on the firm-level and innovation outcomes are depicted in Figure 3.3. The latter shows that the average treatment effect (i.e. the changes in firm financial and innovation outcomes due to the R&D subsidy) is highly dependent on the proportion of R&D funded firms in the cluster. For example, the direct effect of funding on the quality-adjusted number of patents enhances by around 3.5% when we increase the share of R&D funded firms up to 1.6 %.

When we further increase the share of R&D funded firms to 1.7% and above, its direct effect on patents starts to decline. The picture is different when we observe the spillover effect of R&D funding on innovation, which will be around 10% and 30% when the proportions of R&D funded firms in the cluster are 1% and 2%. This might indicate that the R&D funding indirectly initiated new collaborations between funded and non-funded firms, thus leading to the generation of more patents. Analogous results regarding more pronounced indirect effects were observed by Czarnitzki et al. (2007) who identified no direct impact but rather significant indirect effects on private R&D stemming from positive spillovers facilitated by collaborations.

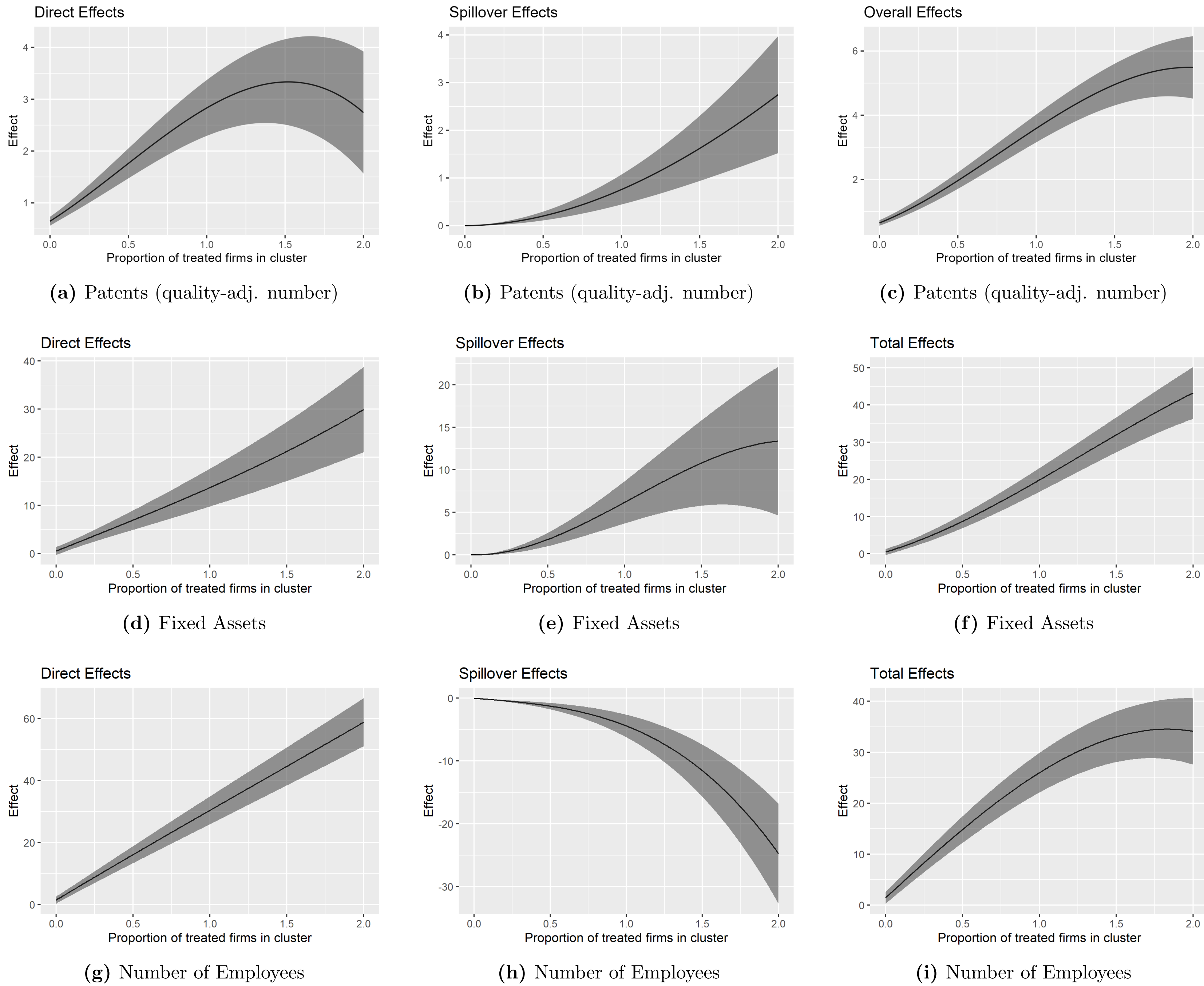
Further, we uncover a significant positive direct impact on the financial outcomes of treated firms. For example, the direct effects of R&D subsidy on fixed assets and the number of employees are 10% and 20% when the proportion of funded firms in the cluster is 1%. When we increase the share of funded firms up to 2%, the direct effect on fixed assets and the number of employees almost triples as well by reaching up to 30% and 20% respectively. The positive trend holds for the spillover effects of R&D subsidy on fixed assets of non-funded firms. However, the indirect effect on the number of employees is negative and reduces by 25% when we increase the share of funded firms up to 2%. This can indicate that the R&D funding of single firms triggered negative competition effects for employment.

Our findings suggest that after the treatment, the funded firms increase their capital and labor. The non-funded firms increase their knowledge output, which can be seen from the increase in the number of quality-adjusted patents for non-funded firms with exposure to funded firms. If we assume high-skill labor for innovative activity to be limited within the cluster, the non-funded firms increase their innovative input by increasing capital but suffer from competing against funded firms in the high-skill labor market. Therefore, they increase fixed assets and the number of employees is decreased, while the funded firms increase their number of employees and fixed assets thus increasing innovative output. Similar results regarding negative competition effects for labor are reported in the previous research on cutting-edge innovation clusters conducted by Lehmann and Menter (2017). While they observed positive impacts on innovation outcomes similar to our results, they also identified 'beggar-thy-neighbor' effects in sunrise regions affecting neighboring areas. They argue that focusing on 'winning' regions increases their capacity and the demand for highly skilled labor, causing neighboring regions to suffer more from the 'power of attraction' of these leading-edge clusters in comparison to non-neighboring regions.

Finally, we obtain the total effect by combining the direct and indirect effects together. We find, that the total effects on both innovation (patents) and firm-level (fixed assets and employment) outcomes are uniformly positive and increasing with the share of funded firms in the cluster. However, for the number of employees and quality-adjusted number of patents, the total effect stabilizes when the proportion of funded firms in the cluster goes beyond 1.6%. Despite having negative spillovers from employment, here the direct effect dominates, and thus the total effect of R&D funding on the number of employees is positive.

These findings align with the broader consensus in the innovation literature, which suggests that R&D funding enhances innovation (Aschhoff 2009; Bronzini and Piselli 2016; Czarnitzki and Fier 2002; Czarnitzki and Hussinger 2017; Czarnitzki and Licht 2006; Czarnitzki and Lopes-Bento 2011; Kleine et al. 2022) and firm-level outcomes (Falk 2010; Hall et al. 2009; Lerner 1999). They also support the ideas of Akcigit et al. (2018), Breschi and Lissoni (2009), Crass et al. (2017), Engel et al. (2017), Nishimura and Okamuro (2010), and Uyarra and Ramlogan (2016), who highlight the importance of agglomeration for firm-level development. Furthermore, the findings reinforce the spillover literature's arguments regarding the positive effects of spatial (Acemoglu et al. 2016; Breschi and Lissoni 2005; Petruzzelli 2011) and industrial (Bloom et al. 2013; Li and Bosworth 2018) proximity demonstrated by the unique region-industry clusters in our analysis. They also verify the findings of Girma et al. (2014) regarding FDI spillovers in innovation setup, as the prevalence of R&D funded firms suggests a greater propensity for positive externalities through increased collaboration.

Figure 3.3: Causal effects of R&D funding on innovation and firm-level performance of funded and non-funded firms with 95% confidence intervals



(a) Patents (quality-adj. number)
(b) Patents (quality-adj. number)
(c) Patents (quality-adj. number)
(d) Fixed Assets
(e) Fixed Assets
(f) Fixed Assets
(g) Number of Employees
(h) Number of Employees
(i) Number of Employees

3.6 Conclusions

This research implements the causal analysis of the direct and spillover effects of R&D funding on firm-level performance and innovation activity of funded and non-funded firms. This study reveals a crucial insight: the effects of research and development (R&D) funding, both direct and spillover impact, exhibit variability contingent on the proportion of funded firms within the cluster. While we see a positive direct impact on firm capital and labor when we increase the proportion of funded firms, the direct effect on patents decreases when the share of funded firms exceeds 1.7%. For spillovers on non-funded firms, we find that they are positive for patents and fixed assets, however, they are negative for the number of employees. What is also interesting, the indirect effects for the quality-adjusted number of patents and fixed assets increase monotonically with the proportion of funded firms. Consequently, policymakers should encourage the concentration of more R&D-funded firms in a cluster, however, not above the threshold of 1.7% to maintain innovation effects.

Another striking finding is that the direct effect of R&D funding on patents is less pronounced compared to the indirect effect: for the clusters with a higher proportion of treated firms, the indirect effect starts to dominate. This signals that funding R&D activities of every single firm in an isolated setup is not a good approach. Rather, the policy aiming at the R&D funding of several firms close to each other (either spatially or industry-wise) will enforce both direct and indirect positive effects. One working example of the aforementioned policy setup can be the cluster type of funding, which according to the existing literature (Acemoglu et al. 2016; Audretsch et al. 2018; Breschi and Lissoni 2005; Crass et al. 2017; Lehmann and Menter 2017) does not only have a positive direct impact on the innovation activity of funded firms but also enforces positive spillovers for non-funded neighbor firms. Finally, the total effect is positive and significant for all outcome variables.

Generally, the results of our study show that the estimation of both direct and indirect treatment effects provides a clearer understanding of different mechanisms through which the share of funded firms impacts the potential outcomes of funded and non-funded firms. In this way, our research significantly contributes to the policy debate regarding the benefits of R&D funding and optimal allocation from a regional development perspective and aligns well with the earlier studies reporting positive direct (Aschhoff 2009; Bronzini and Piselli 2016; Czarnitzki and Hussinger 2017; Falk 2010; Hall et al. 2009) and indirect effects (Czarnitzki et al. 2007).

However, our research also has some potential limitations. The results are obtained under the Partial Inference Assumption, assuming that SUTVA holds across the clusters but not within. There are scenarios where this might not be the case. A potential threat to the SUTVA assumption is the presence of ownership groups. Subsidiaries within these ownership groups are linked across different clusters and might interact as part of the same group, implying inter-firm interactions across clusters, and potentially violating the SUTVA assumption. Additionally, input and output linkages might similarly challenge the validity of SUTVA.

Additionally, the assumption of random treatment assignments conditional on observable characteristics might not hold perfectly because, in an ideal scenario, we would require more detailed information about the firms to ensure genuine randomness. Relying only on observable characteristics means that unobserved variables could still affect treatment assignment, introducing potential biases. More comprehensive data on firm-specific attributes, such as internal capabilities and market conditions, would enhance the reliability and ensure a more accurate analysis of the causal effects. Hence, it might be interesting for future research to observe these linkages and utilize more detailed data to improve empirical validity.

Bibliography

Acemoglu, D., Akcigit, U., & Kerr, W. R. (2016). Innovation network. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 11483–11488.

Aerts, K., & Czarnitzki, D. (2004). *Using innovation survey data to evaluate R&D policy: The case of Belgium* (ZEW Discussion Papers No. 04-55). ZEW - Leibniz Centre for European Economic Research.

Aerts, K., & Schmidt, T. (2008). Two for the price of one? *Research Policy*, *37*(5), 806–822.

Aghion, P., & Howitt, P. (2009). *The economics of growth* (1st ed., Vol. 1). The MIT Press.

Akcigit, U., Caicedo, S., Miguelez, E., Stantcheva, S., & Sterzi, V. (2018). *Dancing with the stars: Innovation through interactions* (tech. rep.). National Bureau of Economic Research.

Akcigit, U., Hanley, D., & Serrano-Velarde, N. (2020). Back to basics: Basic research spillovers, innovation policy, and growth. *The Review of Economic Studies*, *88*(1), 1–43.

Almus, M., & Czarnitzki, D. (2003). The effects of public *R&D* subsidies on firms' innovation activities: The case of Eastern Germany. *Journal of Business Economic Statistics*, *21*(2), 226–236.

Antonioli, D., Marzucchi, A., & Savona, M. (2016). Pain shared, pain halved? cooperation as a coping strategy for innovation barriers. *The Journal of Technology Transfer*, *42*(4), 841–864.

Aranguren, M. J., de la Maza, X., Parrilli, M. D., Vendrell-Herrero, F., & Wilson, J. R. (2013). Nested methodological approaches for cluster policy evaluation: An application to the basque country. *Regional Studies*, *48*(9), 1547–1562.

Arrow, K. Economic Welfare and the Allocation of Resources for Invention. In: In *The Rate and Direction of Inventive Activity: Economic and Social Factors*. NBER Chapters. National Bureau of Economic Research, Inc, 1962, pp. 609–626.

Aschhoff, B. (2009). The effect of subsidies on *R&D* investment and success – do subsidy history and size matter? *SSRN Electronic Journal*.

Audretsch, D. B., & Feldman, M. P. (1996). *R&D* Spillovers and the geography of innovation and production. *The American Economic Review*, *86*(3), 630–640.

Audretsch, D. B., Keilbach, M. C., & Lehmann, E. E. (2006). *Entrepreneurship and economic growth*. Oxford University PressNew York.

Audretsch, D. B., Lehmann, E. E., Menter, M., & Seitz, N. (2018). Public cluster policy and firm performance: Evaluating spillover effects across industries. *Entrepreneurship amp; Regional Development*, *31*(1–2), 150–165.

Autant-Bernard, C., Fadairo, M., & Massard, N. (2013). Knowledge diffusion and innovation policies within the european regions: Challenges based on recent empirical evidence. *Research Policy*, *42*(1), 196–210.

Autio, E., & Rannikko, H. (2016). Retaining winners: Can policy boost high-growth entrepreneurship? *Research Policy*, *45*(1), 42–55.

Badillo, E. R., & Moreno, R. (2015). Are collaborative agreements in innovation activities persistent at the firm level? empirical evidence for the spanish case. *Review of Industrial Organization*, *49*(1), 71–101.

Baird, S., Bohren, J. A., McIntosh, C., & Özler, B. (2018). Optimal design of experiments in the presence of interference. *The Review of Economics and Statistics*, *100*(5), 844–860.
Baptista, R., & Swann, P. (1998). Do firms in clusters innovate more? *Research Policy*, *27*(5), 525–540.
Becker, W., & Dietz, J. (2004). R&D cooperation and innovation activities of firms—evidence for the German manufacturing industry. *Research Policy*, *33*(2), 209–223.
Belderbos, R., Carree, M., & Lokshin, B. (2004). Cooperative *R&D* and firm performance. *Research Policy*, *33*(10), 1477–1492.
Bernstein, J. I., & Nadiri, M. I. (1989). Research and development and intra-industry spillovers: An empirical application of dynamic duality. *The Review of Economic Studies*, *56*(2), 249.
Bérubé, C., & Mohnen, P. (2009). Are firms that receive *R&D* subsidies more innovative? *Canadian Journal of Economics/Revue canadienne d'économique*, *42*(1), 206–225.
Bianchini, S., Llerena, P., & Martino, R. (2019). The impact of *R&D* subsidies under different institutional frameworks. *Structural Change and Economic Dynamics*, *50*, 65–78.
Bioeconomy [Accessed: June 3, 2024]. (2024).
BioM Biotech Cluster Development GmbH. (2024). Leading-edge cluster m4 [Retrieved February 18, 2024].
Blanes, J. V., & Busom, I. (2004). Who participates in *R&D* subsidy programs? *Research Policy*, *33*(10), 1459–1476.
Bloom, N., Schankerman, M., & Reenen, J. V. (2013). Identifying technology spillovers and product market rivalry. *Econometrica*, *81*(4), 1347–1393.
BMBF. (2009). *Leitfaden zur Antragstellung im Spitzencluster-Wettbewerb (2. Wettbewerbsrunde) des Bundesministeriums für Bildung und Forschung* (tech. rep.). Bundesministerium für Bildung und Forschung / Federal Ministry of Education and Research (BMBF), Berlin.
BMBF. (2023). *Förderkatalog*.
Boeing, P., & Mueller, E. (2016). *Measuring patent quality and national technological capacity in cross-country comparison* (ZEW Discussion Papers No. 16-048). ZEW - Leibniz Centre for European Economic Research.
Breschi, S., & Lissoni, F. (2005). "Cross-firm" inventors and social networks: Localized knowledge spillovers revisited. *Annales d'Économie et de Statistique*, (79/80), 189.
Breschi, S., & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows. *Journal of Economic Geography*, *9*(4), 439–468.
Briggs, K., & Wade, M. (2014). More is better: Evidence that joint patenting leads to quality innovation. *Applied Economics*, *46*(35), 4370–4379.
Broekel, T., Fornahl, D., & Morrison, A. (2015). Another cluster premium: Innovation subsidies and ramp;d collaboration networks. *Research Policy*, *44*(8), 1431–1444.
Bronzini, R., & Piselli, P. (2016). The impact of R&D subsidies on firm innovation. *Research Policy*, *45*(2), 442–457.
Brown, R., Mawson, S., & Mason, C. (2017). Myth-busting and entrepreneurship policy: The case of high growth firms. *Entrepreneurship & Regional Development*, *29*(5–6), 414–443.

Burger, M. J., Meijers, E. J., Hoogerbrugge, M. M., & Tresserra, J. M. (2014). Borrowed size, agglomeration shadows and cultural amenities in north-west europe. *European Planning Studies*, *23*(6), 1090–1109.

Busom, I. (2000). An empirical evaluation of the effects of *R&D* subsidies. *Economics of Innovation and New Technology*, *9*(2), 111–148.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230.

Cantner, U. (2013). Innovationsentscheidungen und Innovationspfade: Zwischen Unsicherheit und Risiko. In *Exploring uncertainty* (pp. 295–312). Springer Fachmedien Wiesbaden.

Cantner, U., & Graf, H. (2006). The network of innovators in Jena: An application of social network analysis. *Research Policy*, *35*(4), 463–480.

Cantner, U., Graf, H., & Rothgang, M. (2019). Geographical clustering and the evaluation of cluster policies: Introduction. *The Journal of Technology Transfer*, *44*(6), 1665–1672.

Cantner, U., & Kösters, S. (2011). Picking the winner? Empirical evidence on the targeting of R&D subsidies to start-ups. *Small Business Economics*, *39*(4), 921–936.

Chatterji, A., Glaeser, E., & Kerr, W. (2014). Clusters of entrepreneurship and innovation. *Innovation Policy and the Economy*, *14*, 129–166.

Ci3 - cybersecurity innovation hub [Accessed: June 3, 2024]. (2024).

Clausen, T. H. (2009). Do subsidies have positive impacts on *R&D* and innovation activities at the firm level? *Structural Change and Economic Dynamics*, *20*(4), 239–253.

Crass, D., Rammer, C., & Aschhoff, B. (2017). Geographical clustering and the effectiveness of public innovation programs. *The Journal of Technology Transfer*, *44*(6), 1784–1815.

Czarnitzki, D., & Delanote, J. (2012). Young innovative companies: The new high-growth firms? *Industrial and Corporate Change*, *22*(5), 1315–1340.

Czarnitzki, D., & Delanote, J. (2015). *R&D* Policies for young SMEs: Input and output effects. *Small Business Economics*, *45*(3), 465–485.

Czarnitzki, D., & Fier, A. (2002). Do innovation subsidies crowd out private investment? Evidence from the German service sector. *ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research, ZEW Discussion Papers*, *48*.

Czarnitzki, D., & Hottenrott, H. (2011). *R&D* Investment and financing constraints of small and medium-sized firms. *Small Business Economics*, *36*(1), 65–83.

Czarnitzki, D., & Hussinger, K. (2004). *The link between R&D subsidies, R&D spending and technological performance* (ZEW Discussion Papers No. 04-56). ZEW - Leibniz Centre for European Economic Research.

Czarnitzki, D., & Hussinger, K. (2017). *Input and output additionality of R&D subsidies* (DEM Discussion Paper Series). Department of Economics at the University of Luxembourg.

Czarnitzki, D., & Licht, G. (2006). Additionality of public *R&D* grants in a transition economy. The case of Eastern Germany. *The Economics of Transition*, *14*(1), 101–131.

Czarnitzki, D., Doherr, T., Fier, A., Licht, G., & Rammer, C. (2003). *Öffentliche Förderung der Innovationsaktivitäten von Unternehmen in Deutschland.*

Czarnitzki, D., Ebersberger, B., & Fier, A. (2007). The relationship between R&D collaboration, subsidies and R&D performance: Empirical evidence from Finland and Germany. *Journal of Applied Econometrics*, *22*(7), 1347–1366.

Czarnitzki, D., & Lopes-Bento, C. (2011). *Evaluation of public R&D policies: A cross-country comparison* (ZEW Discussion Papers No. 11-053). ZEW - Leibniz Centre for European Economic Research.
Czarnitzki, D., & Lopes-Bento, C. (2014). Innovation subsidies: Does the funding source matter for innovation intensity and performance? Empirical evidence from Germany. *Industry and Innovation*, *21*(5), 380–409.
Czarnitzki, D., & Toole, A. A. (2007). Business R&D and the interplay of R&D subsidies and product market uncertainty. *Review of Industrial Organization*, *31*(3), 169–181.
David, P. A., Hall, B. H., & Toole, A. A. (2000). Is public *R&D* a complement or substitute for private *R&D*? A review of the econometric evidence. *Research Policy*, *29*(4-5), 497–529.
Delgado, M., Porter, M. E., & Stern, S. (2015). Defining clusters of related industries. *Journal of Economic Geography*, *16*(1), 1–38.
Deschryvere, M. (2014). *R&D*, firm growth and the role of innovation persistence: An analysis of Finnish SMEs and large firms. *Small Business Economics*, *43*(4), 767–785.
Diamond, A. M. (1999). Does federal funding "crowd in" private funding of science? *Contemporary Economic Policy*, *17*(4), 423–431.
Digital hub logistics - effizienzcluster [Accessed: June 3, 2024]. (2024).
Dobkins, H. L., & Ioannides, Y. M. (2001). Spatial interactions among U.S. cities: 1900–1990. *Regional Science and Urban Economics*, *31*(6), 701–731.
Dohse, D., & Staehler, T. (2008). *BioRegio, BioProfile and the rise of the German biotech industry* (tech. rep.).
Doraszelski, U., & Jaumandreu, J. (2013). R&D and productivity: Estimating endogenous productivity. *The Review of Economic Studies*, *80*(4 (285)), 1338–1383.
Duguet, E. (2004). *Are R&D subsidies a substitute or a complement to privately funded R&D ? Evidence from France using propensity score methods for non-experimental data* (Public Economics). University Library of Munich, Germany.
Duranton, G. (2011). California dreamin': The feeble case for cluster policies. *Review of Economic Analysis*, *3*, 3–45.
E-mobil south-west [Accessed: June 3, 2024]. (2024).
Engel, D., Eckl, V., & Rothgang, M. (2017). *R&D* Funding and private *R&D*: Empirical evidence on the impact of the leading-edge cluster competition. *The Journal of Technology Transfer*, *44*(6), 1720–1743.
Engel, D., Mitze, T., Patuelli, R., & Reinkowski, J. (2013). Does cluster policy trigger R&D activity? Evidence from German biotech contests. *European Planning Studies*, *21*(11), 1735–1759.
Eurostat. (2022). R&D expenditure (europa.eu, Ed.) [[Online; posted 23-February-2023]].
Eurostat. (2023). Government budget allocations for R&D (GBARD) (europa.eu, Ed.) [[Online; posted 23-February-2023]].
Falck, O., Heblich, S., & Kipar, S. (2010). Industrial innovation: Direct evidence from a cluster-oriented policy. *Regional Science and Urban Economics*, *40*(6), 574–582.
Falk, M. (2010). Quantile estimates of the impact of *R&D* intensity on firm performance. *Small Business Economics*, *39*(1), 19–37.
Falk, R. (2004). *Behavioural Additionality Effects of R&D Subsidies*. WIFO.

Ferracci, M., Jolivet, G., & van den Berg, G. (2014). Evidence of treatment spillovers within markets. *The Review of Economics and Statistics*, *96*(5), 812–823.

Fong, C. J., Patall, E. A., Vasquez, A. C., & Stautberg, S. (2018). A meta-analysis of negative feedback on intrinsic motivation. *Educational Psychology Review*, *31*(1), 121–162.

Freel, M. S. (2000). Do small innovating firms outperform non-innovators? *Small Business Economics*, *14*(3), 195–210.

Fromhold-Eisebith, M., & Eisebith, G. (2008). Looking behind facades: Evaluating effects of (automotive) cluster promotion. *Regional Studies*, *42*(10), 1343–1356.

Fujita, M., Krugman, P., & Mori, T. (1999). On the evolution of hierarchical urban systems. *European Economic Review*, *43*(2), 209–251.

Gambardella, A., Harhoff, D., & Verspagen, B. (2008). The value of European patents. *European Management Review*, *5*(2), 69–84.

Girma, S., Gong, Y., Görg, H., & Lancheros, S. (2014). *Estimating Direct and Indirect Effects of Foreign Direct Investment on firm Productivity in the Presence of Interactions between Firms* (Kiel Working Papers No. 1961). Kiel Institute for the World Economy.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. *Journal of Political Economy*, *100*(6), 1126–1152.

González, X., & Pazó, C. (2008). Do public subsidies stimulate private R&D spending? *Research Policy*, *37*(3), 371–389.

González, X., Jaumandreu, J., & Pazó, C. (2005). Barriers to innovation and subsidy effectiveness. *RAND Journal of Economics*, *36*(4), 930–949.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277.

Graf, H., & Broekel, T. (2020). A shot in the dark? Policy influence on cluster networks. *Research Policy*, *49*(3), 103920.

Graf, H., & Henning, T. (2008). Public research in regional networks of innovators: A comparative study of four East German regions. *Regional Studies*, *43*(10), 1349–1368.

Griliches, Z. (1979). Issues in Assessing the Contribution of Research and Development to Productivity Growth. *Bell Journal of Economics*, *10*(1), 92–116.

Grossman, G. M., & Helpman, E. (1994). Endogenous innovation in the theory of growth. *The Journal of Economic Perspectives*, *8*(1), 23–44.

Hagedoorn, J., Kranenburg, H. v., & Osborn, R. N. (2003). Joint patenting amongst companies – exploring the effects of inter-firm R&D partnering and experience. *Managerial and Decision Economics*, *24*(2–3), 71–84.

Hall, B., Mairesse, J., & Mohnen, P. (2009). *Measuring the returns to R&D* (NBER Working Papers No. 15622). National Bureau of Economic Research, Inc.

Hall, B. H. (2002a). The Assessment: Technology Policy. *Oxford Review of Economic Policy*, *18*(1), 1–9.

Hall, B. H. (2002b). The financing of research and development. *Oxford Review of Economic Policy*, *18*(1), 35–51.

Harhoff, D., Narin, F., Scherer, F. M., & Vopel, K. (1999). Citation frequency and the value of patented inventions. *Review of Economics and Statistics*, *81*(3), 511–515.

Harhoff, D., Scherer, F. M., & Vopel, K. (2003). Citations, family size, opposition and the value of patent rights. *Research Policy*, *32*(8), 1343–1363.

Hassine, H. B., & Mathieu, C. (2020). R&D crowding out or R&D leverage effects: An evaluation of the French cluster-oriented technology policy. *Technological Forecasting and Social Change*, *155*, 120025.

Hill, J., & Su, Y.-S. (2013). Assessing lack of common support in causal inference using bayesian nonparametrics: Implications for evaluating the effect of breastfeeding on children's cognitive outcomes. *The Annals of Applied Statistics*, *7*(3).

Hinzmann, S., Cantner, U., & Graf, H. (2017). The role of geographical proximity for project performance: Evidence from the German Leading-Edge Cluster Competition. *The Journal of Technology Transfer*, *44*(6), 1744–1783.

Hirano, K., & Imbens, G. W. The propensity score with continuous treatments. In: Wiley, 2004, pp. 73–84. ISBN: 9780470090459.

Hottenrott, H., & Lopes-Bento, C. (2016). *R&D* Partnerships and innovation performance: Can there be too much of a good thing? *Journal of Product Innovation Management*, *33*(6), 773–794.

Howell, S. T. (2017). Financing innovation: Evidence from *R&D* grants. *The American Economic Review*, *107*(4), 1136–1164.

Huber, M. (2013). Identifying causal mechanisms (primarily) based on inverse probability weighting: Identifying causal mechanisms. *Journal of Applied Econometrics*, *29*(6), 920–943.

Hud, M., & Hussinger, K. (2015). The impact of *R&D* subsidies during the crisis. *Research Policy*, *44*(10), 1844–1855.

Hudgens, M. G., & Halloran, M. E. (2008). Toward causal inference with interference. *Journal of the American Statistical Association*, *103*(482), 832–842.

Hughes, A. E., & Mina, A. The impact of the patent system on smes. In: 2010.

Hussinger, K. (2008). R&D and subsidies at the firm level: An application of parametric and semiparametric two-step selection models. *Journal of Applied Econometrics*, *23*(6), 729–747.

Hyytinen, A., & Toivanen, O. (2005). Do financial constraints hold back innovation and growth? *Research Policy*, *34*(9), 1385–1403.

Iacus, S. M., King, G., & Porro, G. (2011). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association*, *106*(493), 345–361.

Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, *20*(1), 1–24.

Intelligent Technical Systems OstWestfalenLippe (it's OWL). (2024). Strategy | it's owl [Retrieved February 18, 2024].

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, *108*(3), 577–598.

Jaffe, A. B. (1986). Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. *The American Economic Review*, *76*(5), 984–1001.

Jaffe, A. B. (2002). Building programme evaluation into the design of public research-support programmes. *Oxford Review of Economic Policy*, *18*(1), 22–34.

Jaffe, A. B., & de Rassenfosse, G. (2016). *Patent Citation Data in Social Science Research: Overview and Best Practices* (NBER Working Papers No. 21868). National Bureau of Economic Research, Inc.

Jong, J. P. D., & Freel, M. (2010). Absorptive capacity and the reach of collaboration in high technology small firms. *Research Policy*, *39*(1), 47–54.

Kaiser, U. (2002). An empirical test of models explaining research expenditures and research cooperation: Evidence for the German service sector. *International Journal of Industrial Organization*, *20*(6), 747–774.

Kiese, M., & Hundt, C. (2014). Cluster policies, organising capacity and regional resilience: Evidence from German case studies. *Raumforschung und Raumordnung | Spatial Research and Planning*, *72*(2).

Kleine, M., Heite, J., & Huber, L. R. (2022). Subsidized *R&D* collaboration: The causal effect of innovation vouchers on innovation outcomes. *Research Policy*, *51*(6), 104515.

Koehler, M. (2018). Estimating the benefits of R&D subsidies for Germany. *SSRN Electronic Journal*.

Kuhlmann, S., & Edler, J. (2003). Scenarios of technology and innovation policies in europe: Investigating future governance. *Technological Forecasting and Social Change*, *70*(7), 619–637.

Lach, S. (2002). Do *R&D* subsidies stimulate or displace private R&D? Evidence from Israel. *Journal of Industrial Economics*, *50*(4), 369–390.

Lehmann, E. E., & Menter, M. (2017). Public cluster policy and neighboring regions: Beggar-thy-neighbor? *Economics of Innovation and New Technology*, *27*(5–6), 420–437.

Lerner, J. (1999). The government as venture capitalist: The long-run impact of the SBIR program. *The Journal of Business*, *72*(3), 285–318.

Li, Y., & Bosworth, D. (2018). R&D spillovers in a supply chain and productivity performance in British firms. *The Journal of Technology Transfer*, *45*(1), 177–204.

Lööf, H., & Heshmati, A. (2004). *The Impact of Public Funding on Private R&D investment: New Evidence from a Firm Level Innovation Study* (Working Paper Series in Economics and Institutions of Innovation No. 6). Royal Institute of Technology, CESIS - Centre of Excellence for Science and Innovation Studies.

MAI Carbon - Composites United cluster [Accessed: June 3, 2024]. (2020).

Manski, C., & Pepper, J. (2018). How do right-to-carry laws affect crime rates? coping with ambiguity using bounded-variation assumptions. *The Review of Economics and Statistics*, *100*(2), 232–244.

Martin, P., Mayer, T., & Mayneris, F. (2011). Public support to clusters. *Regional Science and Urban Economics*, *41*(2), 108–123.

Medical valley EMN [Accessed: June 3, 2024]. (2020).

Meuleman, M., & Maeseneire, W. D. (2012). Do *R&D* subsidies affect SMEs' access to external financing? *Research Policy*, *41*(3), 580–591.

Michel, J., & Bettels, B. (2001). Patent citation analysis. A closer look at the basic input data from patent search reports. *Scientometrics*, *51*(1), 185–201.

MicroTEC Südwest. (2024). Microtec Südwest [Retrieved February 18, 2024].

Montmartin, B., & Herrera, M. (2015). Internal and external effects of *R&D* subsidies and fiscal incentives: Empirical evidence using spatial dynamic panel models. *Research Policy*, *44*, 1065–1079.

Müller, S., & Korsgaard, S. (2017). Resources and bridging: The role of spatial context in rural entrepreneurship. *Entrepreneurship & Regional Development*, *30*(1–2), 224–255.

Nishimura, J., & Okamuro, H. (2010). R&D productivity and the organization of cluster policy: An empirical evaluation of the Industrial Cluster Project in Japan. *The Journal of Technology Transfer*, *36*(2), 117–144.

Nishimura, J., & Okamuro, H. (2011). Subsidy and networking: The effects of direct and indirect support programs of the cluster policy. *Research Policy*, *40*(5), 714–727.

Partridge, M. D., Rickman, D. S., Ali, K., & Olfert, M. R. (2009). Do new economic geography agglomeration shadows underlie current population dynamics across the urban hierarchy? *Papers in Regional Science*, *88*(2), 445–467.

Petruzzelli, A. M. (2011). The impact of technological relatedness, prior ties, and geographical distance on university–industry collaborations: A joint-patent analysis. *Technovation*, *31*(7), 309–319.

Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, *90*(5), 2555–2591.

Reitzig, M. (2004). Improving patent valuations for management purposes—validating new indicators by analyzing application rationales. *Research Policy*, *33*(6–7), 939–957.

Rogers, M. (2004). Networks, firm size and innovation. *Small Business Economics*, *22*(2), 141–153.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, *98*(5), S71–S102.

Rosenbaum, P. R. (2007). Interference between units in randomized experiments. *Journal of the American Statistical Association*, *102*, 191–200.

Rosenberg, J. B. (1976). Research and market share: A reappraisal of the Schumpeter hypothesis. *Journal of Industrial Economics*, *25*(2), 101–12.

Rothgang, M., Cantner, U., Dehio, J., Engel, D., Fertig, M., Graf, H., Hinzmann, S., Linshalm, E., Ploder, M., Scholz, A. M., & Töpfer, S. (2017a). Cluster policy: Insights from the German leading-edge cluster competition. *Journal of Open Innovation: Technology, Market, and Complexity*, *3*(1).

Rothgang, M., Cantner, U., Dehio, J., Engel, D., Fertig, M., Graf, H., Hinzmann, S., Linshalm, E., Ploder, M., Scholz, A.-M., Töpfer, S., Impressum, Breuer, M., Schulte, R., Vorsitzende, S., Hans, G., Fabritius, P., Justus, H., Jürgen, H., & Riphahn, T. (2014). *Begleitende Evaluierung des Förderinstruments "Spitzencluster-Wettbewerb" des BMBF* (tech. rep.).

Rothgang, M., Dehio, J., & Lageman, B. (2017b). Analysing the effects of cluster policy: What can we learn from the German leading-edge cluster competition? *The Journal of Technology Transfer*, *44*(6), 1673–1697.

Rothgang, M., Lageman, B., & Scholz, A.-M. (2021). Why are there so few hard facts about the impact of cluster policies in Germany? A critical review of evaluation studies. *Review of Evolutionary Political Economy*, *2*, 105–139.

Scott, J. Firm versus industry variability in R&D intensity. In: In *R&D, patents, and productivity*. National Bureau of Economic Research, Inc, 1984, pp. 233–248.

Smits, R., & Kuhlmann, S. (2004). The rise of systemic instruments in innovation policy. *International Journal of Foresight and Innovation Policy*, *1*(1/2), 4.

Software-Cluster [Accessed: June 3, 2024]. (2024).

Solow, R. M. (1957). Technical change and the aggregate production function. *The Review of Economics and Statistics*, *39*(3), 312.

Spithoven, A., Vlegels, J., & Ysebaert, W. (2019). Commercializing academic research: A social network approach exploring the role of regions and distance. *The Journal of Technology Transfer*, *46*(4), 1196–1231.

Streicher, G., Schibany, A., & Gretzmacher, N. (2004). *Input Additionality Effects of R&D Subsidies in Austria. Empirical Evidence from Firm-level Panel Data*. WIFO.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199.

Svensson, P. (1998). Strategic trade policy and endogenous *R&D* subsidies: An empirical study. *Kyklos*, *51*(2), 259–275.

Sörensen, A., Kongsted, H. C., & Marcusson, M. (2003). R&D, public innovation policy, and productivity: The case of Danish manufacturing. *Economics of Innovation and New Technology*, *12*(2), 163–178.

Trajtenberg, M. (1990). A penny for your quotes: Patent citations and the value of innovations. *The RAND Journal of Economics*, *21*(1), 172.

Töpfer, S., Cantner, U., & Graf, H. (2017). Structural dynamics of innovation networks in German Leading-Edge Clusters. *The Journal of Technology Transfer*, *44*(6), 1816–1839.

UNESCO. (2022). World development indicators - world bank (2022.05.26) (ourworldindata.org, Ed.) [[Online; accessed 4-July-2023]].

Uyarra, E., & Ramlogan, R. The impact of cluster policy on innovation (J. Edler, P. Cunningham, A. Gök, & P. Shapira, Eds.). In: *Handbook of Innovation Policy Impact* (J. Edler, P. Cunningham, A. Gök, & P. Shapira, Eds.). Ed. by Edler, J., Cunningham, P., Gök, A., & Shapira, P. Chapters. Edward Elgar Publishing, 2016. Chap. 7, pp. 196–238.

Wallsten, S. J. (2000). The effects of government-industry R&D programs on private R&D: The case of the small business innovation research program. *The RAND Journal of Economics*, *31*(1), 82.

Wolf, T., Cantner, U., Graf, H., & Rothgang, M. (2017). Cluster ambidexterity towards exploration and exploitation: Strategies and cluster management. *The Journal of Technology Transfer*, *44*(6), 1840–1866.

Zúñiga-Vicente, J. Á., Alonso-Borrego, C., Forcadell, F. J., & Galán, J. I. (2012). Assesing the effect of public subsidies on firm *R&D* investment: A survey. *Journal of Economic Surveys*, *28*(1), 36–67.

APPENDIX A

Chapter 1

A.1 Leading-Edge Cluster Competition

A.1.1 The main goals of the "Leading-edge cluster" policy

The "Leading-edge Cluster Competition" (LECC) was launched in 2007 as part of Germany's "High-Tech Strategy 2020" [1] initiative by the Federal Ministry of Education and Research (BMBF) (Lehmann and Menter 2017). With a total budget of 600 million euros, LECC stands as the largest cluster program established by the government to date (Rothgang et al. 2017a), resulting in over 40 business startups, 300 patents, and 900 innovations (Lehmann and Menter 2017). The primary objective of the program was to foster innovative capabilities and promote the development of efficient, highly productive clusters in Germany by offering financial support to collaborative R&D projects in selected cluster regions (Lehmann and Menter 2017). Various strategies were employed to create an innovative ecosystem and enhance regional innovation performance, including strengthening R&D collaboration among diverse stakeholders such as firms, universities, and research institutes (Töpfer et al. 2017).

In greater detail, the LECC aimed to generate positive impacts across six target levels: (1) projects, (2) actors, (3) cluster organizations, (4) clusters, (5) cluster regions, and (6) the national economy. As such, the program established a wide range of objectives regarding the future development of the Leading-Edge Clusters (Rothgang et al. 2017a).

- Enhancing the technological and economic competencies of Leading-Edge Clusters based on their strategic plans.

[1] The "High-Tech Strategy 2020" was initiated by the German Federal Ministry of Education and Research (BMBF) with the objective of allocating public funding to foster innovation, entrepreneurship, and enhance productivity and prosperity across German regions, leveraging their unique factors and resources (Lehmann and Menter 2017)

- Leveraging regional innovative potential to foster sustained value creation, with a focus on promoting novel forms of R&D collaboration, particularly in innovative domains.
- Elevating the international visibility and competitiveness of the clusters as premier technology hubs, while attracting innovative foreign enterprises to establish operations in cluster regions.
- Prioritizing the training of highly skilled professionals and workers.

In total, the cluster competition consisted of three rounds (in 2008, 2010, and 2012), during which the Federal Ministry selected 15 clusters based on the recommendations of an expert jury. These clusters, designated as Leading-Edge Clusters, received funding of up to 40 million euros per cluster over a five-year period (Hinzmann et al. 2017). The funding for the third wave of the LEC concluded in 2017 (Rothgang et al. 2017b). The awarded funds were utilized by the winning clusters to execute research and development projects in collaboration with partners sharing the same thematic focus (Hinzmann et al. 2017).

Unlike other political initiatives that concentrated on the identification and financing of specific regions through a top-down approach, the "Leading-edge cluster competition" encouraged a self-selection process among German regions. In this scenario, regions had to independently apply for financial and governmental support, thereby fostering a political bottom-up approach. This approach aided regions in identifying and leveraging their unique regional factors and resources, ultimately contributing to successful outcomes and enhancing the overall efficiency of the cluster (Lehmann and Menter 2017).

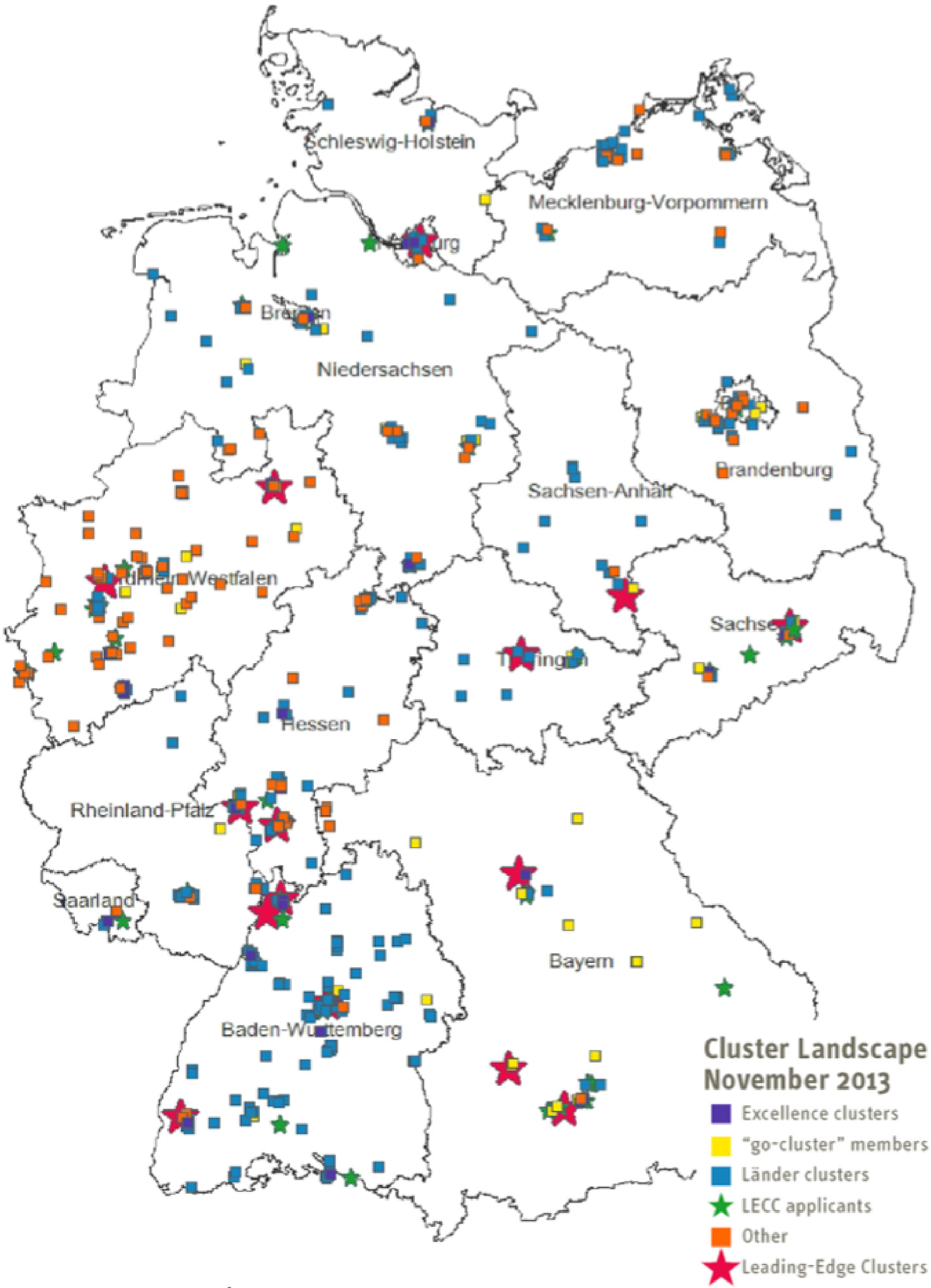
In summary, the overarching objective of the LECC was to bolster the German economy through multifaceted regional development. Consequently, policymakers focused on fostering collaboration between the public and private sectors and elevating the involvement of universities as active partners within funded clusters. These clusters, in turn, were tasked with cultivating an enabling environment, especially for small- and medium-sized enterprises, to facilitate groundbreaking innovation and regional expansion (Lehmann and Menter 2017).

A.1.2 Policy targets

As outlined earlier, the LECC designated funding for 15 cluster initiatives. Figure A.1 depicts the geographical dispersion of these Leading-Edge Clusters alongside approximately 640 cluster programs established at the national or Länder level in 2013. This visualization underscores the pivotal role of cluster support in Germany's innovation policy. While all German Länder endorses clusters, there are notable variations in the number of clusters funded and the extent of support provided. Moreover, it's important to note that Leading-Edge Clusters are integrated into the cluster structures and networks sanctioned by the Länder. Typically, the Länder offer financial backing to clusters for a limited duration, occasionally leveraging funding avenues from the European Fund for Regional Development (EFRE) (Rothgang et al. 2017a).

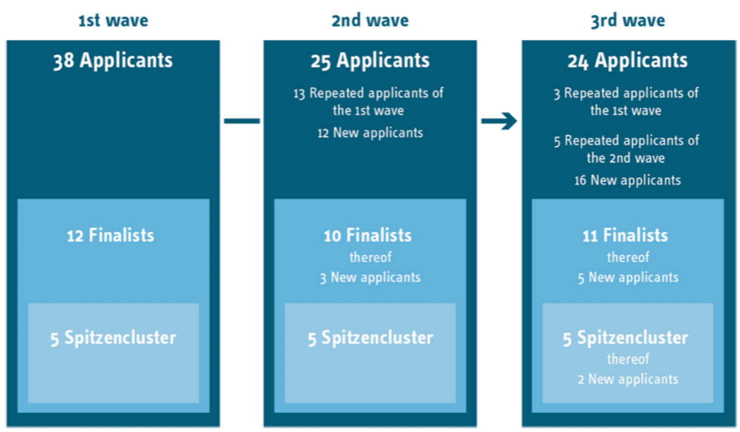
Before delving into the supported programs, let's briefly examine the selection process within the LECC. This process involved a two-step selection mechanism overseen by a jury comprising representatives from both industry and academia. Notably, the competition welcomed proposals from all technology domains. The funding was allocated across two rounds, with the evaluation outcomes of the initial stage influencing the distribution of funds for the subsequent phase (Rothgang et al. 2017a).

Figure A.1: The German networks and clusters at the federal and Länder levels.
Source: Adapted from Rothgang et al. (2017a), p.4



The LECC's selection process was transparent and well-structured, with clear proposal deadlines and thorough selection criteria. Figure A.2 illustrates how LECC participants were distributed across three waves. The figure demonstrates that the LECC effectively encouraged new cluster initiatives in both the second and third rounds. Furthermore, the program played a role in enhancing the organizational and strategic processes of initially unsuccessful clusters, some of which went on to succeed in subsequent rounds. For example, in the second wave, 7 out of 10 finalists had previously been unsuccessful applicants in the first wave. Additionally, three applicants from the first wave and five from the second wave participated in the third wave of the cluster competition (Rothgang et al. 2017a).

Figure A.2: Applicants in the three selection waves of the LECC
Source: Adapted from Rothgang et al. (2017a), p.2



The BMBF outlined specific requirements to uphold the innovative success of clusters, with two key stipulations being:

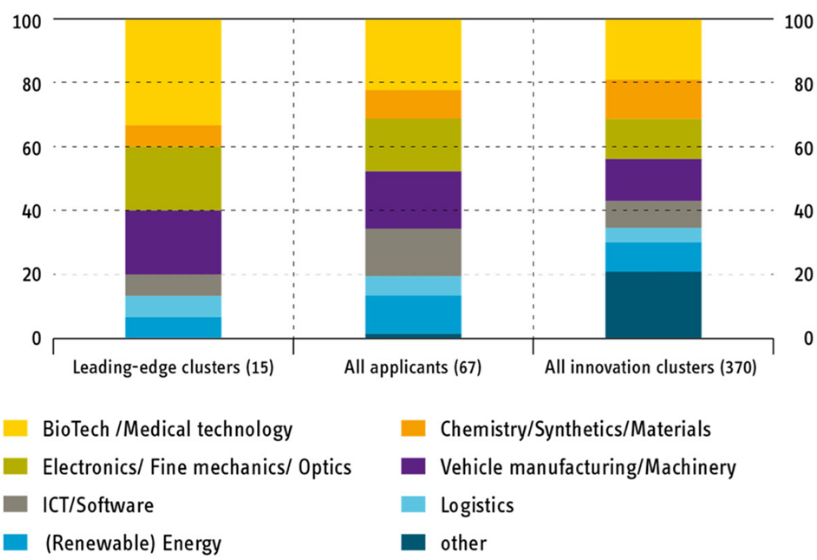
- The presence of effective cluster management capable of coordinating activities within the cluster and facilitating connections between cluster participants by promoting collaboration between academia and industry (Töpfer et al. 2017).

- The cluster management, in collaboration with the cluster members' committee, was tasked with pre-selecting projects and devising a collective cluster strategy that balanced exploration (new research) and exploitation (previous research) (Wolf et al. 2017). To achieve this strategy, a strong emphasis was placed on fostering cooperation among all stakeholders. According to Töpfer et al. (2017), social cohesion notably improved during the LECC project phase, attributed to various initiatives organized by cluster management, such as regular formal and informal gatherings, as well as open seminars.

As previously mentioned, the LECC welcomed participation from all sectors and technological domains, which could explain its engagement across various industries. Additionally, the regional composition of industries and diversity play significant roles in determining the geographical distribution of cluster initiatives. Regions with higher levels of industrial development tend to host more cluster initiatives and networks. In essence, the prevalence of cluster initiatives correlates positively with the degree of regional industrialization. Consequently, nine out of fifteen Leading-Edge Clusters are situated in the more industrialized southern regions of Germany (Rothgang et al. 2017a).

Figure A.3 illustrates the sectoral breakdown of the LECC compared to all applicants of the LECC and other innovative clusters in Germany. It is noticeable from this figure that within the LECC, the biotechnology and medical technology sectors are prominent, with an average number of cluster applications. This prominence may be attributed to the existence of previous funding programs for networks and clusters, such as BioRegio, which facilitated the development of structures conducive to LECC candidacy. Additionally, other notable industries in the LECC include Electronics/Fine Mechanics/Optics and ICT/Software, which also exhibit a higher representation compared to other comparison groups (Rothgang et al. 2017a).

Figure A.3: The industry distribution of the clusters in the LECC
Source: Adapted from Rothgang et al. (2017a), p.5



Shifting the focus to the regions receiving support, the following is a brief overview of 15 unique clusters spanning different industries funded within the framework of the LECC:

1. **The "BioRN"** cluster, situated in the Rhein-Neckar region, specializes in red biotechnology. SMEs primarily lead this cluster, with significant involvement from large firms as major consumers. The cluster boasts a substantial presence of actors from the pharmaceutical industry, supplemented by knowledge-intensive services in life sciences and software development (Töpfer et al. 2017).

2. **"Cool Silicon"** operates across Dresden and Chemnitz, prioritizing enhanced energy efficiency within information and communication technology, particularly in microelectronics. As a result, cluster participants maintain close ties with electronic semiconductors and devices, medical technology, and industrial process and control technology sectors. Notably, unlike other major IT clusters globally (e.g., Silicon Valley in the USA), research and development activities are integrated with productive operations within this cluster (Töpfer et al. 2017).

3. The cluster **"Forum Organic Electronics" (FOE)**, situated in the Rhein-Neckar region, focuses on organic electronics, a cross-sectional technology offering a wide range of potential applications for the processing industry. The participants specialize in advanced and high-tech applications, mainly comprising large firms, renowned research institutes, and universities due to the innovative technology and significant emphasis on basic research. In contrast, the representation of SMEs is relatively lower compared to other selected clusters. Research objectives within FOE include organic LED and photovoltaic technologies, organic sensors, memories, and circuits (Töpfer et al. 2017).

4. Another thriving cluster is **"Hamburg Aviation"**, located in the Hamburg region. This cluster encompasses two facets of the aviation industry. Primarily, it involves activities related to aerospace technology and engineering, such as the design and production of cabin systems, as well as innovative applications for fuel cells. Additionally, it encompasses a wide array of services associated with aeronautical techniques and air traffic management. "Hamburg Aviation" comprises a significant number of SMEs, alongside a few large firms, research organizations, and universities (Töpfer et al. 2017).

5. **"Solar Valley"**, situated in Middle Germany encompassing Saxony, Saxony-Anhalt, and Thuringia, specializes in silicon-based photovoltaic and thin-film technology. The cluster participants are intricately involved in the semiconductor industry, with a particular emphasis on the production of electronic semiconductors, electronic and fine mechanical optical devices, and optical instruments, among others (Töpfer et al. 2017).

6. **The Effizienz Cluster Logistic Ruhr**, located in Ruhr, serves as a global hub for pioneering the development of efficient and effective logistic services. Currently, the cluster comprises 160 companies and 12 scientific and educational institutions actively engaged in various projects. In addition to the cluster partners, the Effizienz Cluster Logistic Ruhr receives support from numerous other firms and organizations beyond the cluster's immediate scope (*Digital Hub Logistics - Effizienzcluster* 2024).

7. Since 2010, **Medical Valley EMN** has been integrated into the LECC, covering the Nürnberg region. It focuses on developing efficient products and services in the realms of prevention, diagnosis, therapy, and rehabilitation, aiming to enhance overall healthcare structures. Within the LECC framework, Medical Valley partners have executed 45 R&D projects between 2010 and 2015, concentrating on five main areas: ophthalmology, diagnostics, imaging, horizontal innovation, and intelligent sensors and therapy systems. These endeavors have notably contributed to improving the efficiency of the healthcare system (*Medical Valley EMN* 2020).

8. **The Munich Biotech Cluster**, situated in the Munich region, received LECC funding from 2010 to 2015. Specializing in personalized drugs and targeted therapies, the cluster engages over 100 partners from both industry and academia (BioM Biotech Cluster Development GmbH 2024).

9. One of the pioneering internationally renowned Leading-Edge Clusters (LECs) is **MicroTEC Southwest**, located in the Baden-Württemberg region. The primary objectives of this cluster are to offer intelligent microsystem solutions in the domains of

production and life sciences, with the aim of boosting the competitiveness of manufacturing companies while simultaneously enhancing safety, security, resource efficiency, and overall quality of life (MicroTEC Südwest 2024).

10. **The Software Cluster** stands as Europe's most robust network of companies and research institutions in the realm of software development. Encompassing a wide area in the southwest of Germany, including the cities of Darmstadt, Kaiserslautern, Karlsruhe, Saarbrücken, and Walldorf, along with their surroundings, the cluster places particular emphasis on enterprise software. It aims to bolster its technological leadership in cloud-based platforms and Industry 4.0 solutions. To realize these objectives, since 2017, the cluster has forged strong collaborations with partners from the world's most innovative regions, such as Silicon Valley (USA), Singapore, and Bahia (Brazil). These partnerships offer exceptional living laboratories for data-driven projects, capitalizing on vast amounts of diverse data (*Software-Cluster* 2024).

11. Since its designation as a Leading-Edge Cluster in 2012, **the BioEconomy Cluster** has received funding from the German Federal Ministry of Education and Research until 2017. Operating across the regions of Sachsen-Anhalt and Sachsen, it brings together industry and scientific stakeholders to advance the material foundation and energetic utilization of non-food biomass. The BioEconomy cluster serves as a regional hub of expertise in bioeconomics, engaging a diverse range of industries including timber and forestry, chemicals, plastics, and plant engineering. Through the incorporation of upscaling methods, the cluster facilitates the swift transition of processes from laboratory-scale to industrial-scale development (*Bioeconomy* 2024).

12. **The Cluster for Individual Immune Intervention (Ci3**) is situated in the Rhine-Main region, spanning across Rheinland-Pfalz, Hessen, and Baden-Württemberg. Similar to other Leading-Edge Clusters of the third wave, Ci3 received government funding from 2012 to 2017. The cluster is dedicated to consolidating exceptional and innovative expertise in individualized immune intervention, covering a spectrum of knowledge in drugs, diagnostics, and therapeutic approaches, with the overarching aim of positioning the region as a global leader in this field. Ci3's projects predominantly address challenges associated with oncology, infections, and autoimmune diseases (*CI3 - Cybersecurity Innovation Hub* 2024).

13. **The Electric Mobility South-West Cluster** extends across the regions of Karlsruhe, Mannheim, Stuttgart, and Ulm, engaging over 120 stakeholders from both industry and science. The primary objective of the cluster is to drive technological advancements in the automotive sector, with a focus on reducing pollution, enhancing efficiency, and promoting the acceptance of electric mobility in the market. To achieve these goals, the cluster management has formulated a comprehensive strategy that includes targets such as "Markets and Costs" (aiming to create competitive life-cycle expenditures), "Handling and Comfort" (meeting customer demands for electric vehicles), and "Connected Mobility" (improving the availability of transportation modes) (*e-mobil South-West* 2024).

14. **It's OWL (Intelligent Technical Systems Ostwestfalen-Lippe)** is among the five

Leading-Edge Clusters of the third wave, situated in the Nordrhein-Westfalen region. The primary objective of this cluster is to create solutions for Intelligent Technical Systems, encompassing intelligent products and production methods. These systems emerge from the convergence of engineering sciences and information technology, characterized by their adaptability to the environment, robustness, user-friendliness, and high predictive capabilities. To achieve its overarching goal, the cluster fosters the development of new technological fields within its network, including machine intelligence, the design of socio-technical systems, digital infrastructure, safety and security, value-added networks, and advanced Systems Engineering (Intelligent Technical Systems OstWestfalenLippe (it's OWL) 2024).

15. **MAI Carbon** stands as the final Leading-Edge Cluster endorsed by an independent jury during the third round (2012) of the LECC. Positioned strategically within the triangle of Munich, Augsburg, and Ingolstadt, the cluster benefits from a highly advantageous location. Leveraging quantum leap innovation, MAI Carbon endeavors to harness the lightweight construction potential of CFRP technology (Carbon fiber reinforced plastics), aiming to establish it as a standard not only in the automotive and aerospace sectors but also in the realm of mechanical engineering (*MAI Carbon - Composites United Cluster* 2020).

A.2 Data

The table A.1 outlines a detailed process for generating and refining our dataset, encompassing ten sequential steps. Beginning with matching LECC data to Orbis, the process involves integrating matched data, removing outliers, and limiting the dataset to the years 2000-2016. Firms with less than five observations are filtered out, and duplicates are eliminated. Subsequently, the dataset is restricted to the year 2007 for Coarsened Exact Matching (CEM), since the matching can be implemented in the pre-treatment year (reminder, the LECC funding started in 2008).

Table A.1: Data generation process

			Number of firms			
	Description	Year	All	LECC	Non-LECC	Notes
1	Match LECC data to Orbis	1009-2018	2,771,687	1,322	2,770,365	Matched by firms name and address.
2	Join matched LECC data to Orbis.	1009-2018	2,771,666	1,301	2,770,365	21 LECC firms not matched, absent in Orbis.
3	Outliers in 99% & 1% quantiles are removed	1009-2018	2,771,666	1,301	2,770,365	No firms removed, only 21 observations
4	Data is limited to 2000-2016	2000-2016	2,271,842	1,290	2,270,552	11 LECC firms drop, no data is available for 2000-2016.
5	Firms with less than 5 observations are removed	2000-2016	762,300	925	761,375	365 LECC & 1 509 542 non-LECC firms drop
6	Duplicates are removed	2000-2016	762,300	925	761,375	No duplicates are found
7	Data is limited to year 2007 for CEM[a]	2007	585,047	751	584,296	Matching is allowed on covariates before the treatment.
8	CEM on firm age[b]	2007	1502	751	751	Matching rate- 100%
9	CEM on firm age + firm size[c]	2007	1500	750	750	Matching rate- 99.87%
10	CEM on firm age + firm size + industry[d]	2007	1444	722	722	Matching rate- 96.19%

[a] Coarsened Exact Matching
[b] Firm age = current year (2007) - year incorporation
[c] Firm size = log(total assets)

[d] Industry is distinguished based on the 2-digit NACE code

Dependent Variables-Quality-Adjusted Number of Patents: When assessing innovation outcomes, patent applications are usually used. However, measuring patents poses challenges as they vary in commercial value and technological impact making the use of simple patent count irrelevant. Research indicates that while most patents have limited value, a select few represent significant technological breakthroughs (Gambardella et al. 2008). In the absence of explicit value data, citation-based measures provide the best approximation of patent quality (Gambardella et al. 2008; Reitzig 2004). Patents frequently cited by subsequent patents are generally considered higher quality, provided that differences in citation propensity are controlled for (Harhoff et al. 1999; Jaffe and de Rassenfosse 2016; Trajtenberg 1990).

However, comparing citation counts across countries has limitations. For one, applicants tend to protect more valuable patents abroad (due to the elevated expenses associated with filing and translating patents), making direct comparisons between domestic and foreign applications less informative (Harhoff et al. 2003). Additionally, variations in examination practices across national patent offices result in significant differences in citation counts. Furthermore, patent examiners commonly demonstrate a tendency to favor citing patents from their native country (Michel and Bettels 2001).

To address these challenges, we adopt a patent quality measure proposed by Boeing and Mueller (2016) that enables technology-specific cross-country comparisons and incorporates forward citations. This measure accounts for factors such as the number of technology classes, family size (number of countries in which a patent application was filed), and the number of forward citations received by the patent.

A.2.1 Direct effects: LECC vs non-LECC firms

The table A.2 provides descriptive statistics for various variables in the year 2007, comparing LECC-funded firms to non-LECC firms. The dataset comprises a total of 585,047 observations, with 751 belonging to LECC-funded firms and 584,296 to non-LECC firms.

In terms of firm age, LECC-funded firms have an average age of 18.9 years, whereas non-LECC firms' average is 16.0. Regarding firm size, measured by the log of total assets, LECC-funded firms demonstrate a higher average size of 15.5 compared to 12.9 for non-LECC firms.

Table A.2: Descriptive statistics on firm characteristics in 2007

	LECC (n= 751)	Non-LECC (n=584296)	Overall (n=585047)
Firm age	18.9 (27.0)	16.0 (21.0)	16.0 (21.0)
Firm size	15.5 (2.84)	12.9 (2.19)	12.9 (2.19)
Patent (simple)	79.6 (389)	5.41 (67.2)	7.48 (93.5)
Patent (adjusted)	226 (1120)	16.0 (203)	21.9 (276)
Joint patents	20.6 (143)	0.501 (13.9)	0.773 (21.7)
Co-partners	32.4 (242)	0.877 (23.5)	1.30 (36.7)
Fixed assets	709 (6810)	5.65 (184)	6.41 (290)
Total assets	1060 (8620)	9.48 (284)	10.6 (402)
Turnover	1670 (8610)	33.4 (494)	37.3 (652)
Sales	1680 (8680)	43.8 (502)	49.8 (731)
Employee costs	282 (1470)	8.72 (119)	10.2 (161)

Notes: Fixed & total assets, turnover, sales, and employee costs values are in Mln.

Patent activity is notably higher among LECC-funded firms, both in simple counts (79.6 vs. 5.41) and adjusted counts (226 vs. 16.0). Moreover, LECC-funded firms tend to engage more in joint patents and co-partnerships compared to their non-LECC counterparts. Examining financial metrics, LECC-funded firms generally exhibit higher values for fixed assets, total assets, turnover, sales, and labor costs compared to non-LECC firms.

Across different industries, the distribution of LECC and non-LECC firms varies significantly (see Table A.3). LECC-funded firms are found across a diverse range of industries, with a notable presence in sectors such as Information and Communication, Wholesale and Retail Trade, and Professional, Scientific, and Technical Activities. These industries account for a significant portion of LECC-funded firms, indicating a preference for investment or funding in sectors with potentially high growth or innovation prospects.

Additionally, LECC-funded firms are also observed in industries like Manufacturing, Administrative and Support Service Activities, and Construction of Buildings, albeit in relatively smaller numbers compared to other sectors. This suggests a diversified portfolio of investments by cluster policy funding, spanning both traditional and knowledge-intensive industries (see Table A.3).

Further, the cross-sectional data from 2007 is used to conduct CEM. CEM is applied sequentially, initially focusing on firm age, then incorporating firm size, and finally extending to include industry classification. Each step aims to enhance dataset quality and prepare it for subsequent analyses, ensuring robustness and accuracy in research findings (see more details in Chapter 1.5).

Table A.3: Descriptive statistics on industry distribution in 2007

Industry	LECC (n= 751)	Non-LECC (n=584296)	Overall (n=585047)
Administrative and Support Service Activities	7 (0.9%)	8275 (1.0%)	8282 (1.0%)
Construction of Buildings	2 (0.3%)	27686 (3.4%)	27688 (3.4%)
Education	4 (0.5%)	5518 (0.7%)	5522 (0.7%)
Financial and Insurance Activities	2 (0.3%)	8129 (1.0%)	8131 (1.0%)
Human Health and Social Work Activities	4 (0.5%)	4373 (0.5%)	4377 (0.5%)
Information and Communication	19 (2.4%)	5833 (0.7%)	5852 (0.7%)
Manufacturing	13 (1.7%)	4301 (0.5%)	4314 (0.5%)
Other Service Activities	6 (0.8%)	6386 (0.8%)	6392 (0.8%)
Professional, Scientific and Technical Activities	22 (2.8%)	25187 (3.1%)	25209 (3.1%)
Real Estate Activities	5 (0.6%)	70568 (8.6%)	70573 (8.6%)
Transportation and Storage	4 (0.5%)	5895 (0.7%)	5899 (0.7%)
Wholesale and Retail Trade	21 (2.7%)	54526 (6.7%)	54547 (6.7%)
Accommodation and Food Service Activities	0 (0%)	8363 (1.0%)	8363 (1.0%)
Activities of households as employers	0 (0%)	19 (0.0%)	19 (0.0%)
Arts, Entertainment and Recreation	0 (0%)	2404 (0.3%)	2404 (0.3%)
Public Administration and Defense: Compulsory Social Security	0 (0%)	1354 (0.2%)	1354 (0.2%)
Water Supply	0 (0%)	1256 (0.2%)	1256 (0.2%)

A.2.2 Spillover effects: Non-funded neighbors vs other firms

In the second part of our analysis, our data comprises the non-funded cluster neighbors (up to a 10 km radius) and all other non-funded firms above a 10 km radius. Table A.4 presents descriptive statistics comparing non-funded neighbor firms to other non-funded firms in 2007. On average, neighbor firms are slightly younger (average firm age is 17.0 years) compared to other firms (18.4 years). In terms of firm size, neighbor firms have a slightly higher average (0.854) compared to other firms (0.774).

Interestingly, neighbor firms have a lower average number of adjusted patents (0.212) and joint patents (0.510) compared to other firms (0.239 and 1.00 respectively). Regarding financial metrics, neighbor firms tend to have higher average fixed assets (9.57 Mln) and total assets (14.8 Mln) compared to other firms (6.37 Mln and 11.0 Mln, respectively). This trend is also reflected in turnover, sales, and employee costs, where neighbor firms generally have higher averages compared to other firms.

The examination of Table A.5 highlights distinct variations in industry representation between neighbor firms and other firms. Notably, neighbor firms show elevated proportions in sectors like Construction of Buildings, Information and Communication, Financial and Insurance Activities, Real Estate Activities, and Professional, Scientific, and Technical Activities, where LECC firms also hold significant influence. In contrast to other firms, neighbor firms exhibit lower representation in industries related to Arts, Entertainment and Recreation, Accommodation and Food Service Activities, and Household activities as employers.

Table A.4: Descriptive statistics on non-funded neighbors vs other firms in 2007

	Neighbor firms (n=157,804)	Other firms (n=427,243)	Overall (n=585,047)
Firm age	17.0 (20.0)	18.4 (22.5)	18.0 (21.9)
Firm size	0.854 (1.12)	0.774 (0.975)	0.796 (1.02)
Patent (adjusted)	0.212 (7.59)	0.239 (34.1)	0.232 (29.4)
Joint patents	0.510 (5.29)	1.00 (28.8)	0.817 (23.0)
Fixed assets	9.57 (243)	6.37 (352)	7.24 (326)
Total assets	14.8 (358)	11.0 (480)	12.0 (450)
Turnover	49.5 (701)	46.3 (782)	47.0 (763)
Sales	54.6 (475)	58.8 (901)	57.8 (815)
Employee costs	11.0 (57.0)	10.2 (154)	10.4 (135)

Notes: Fixed & total assets, turnover, sales, and employee costs values are in Mln.

Table A.5: Descriptive statistics on industry distribution of non-funded neighbors vs other firms in 2007

Industry	Neighbor firms (n=157,804)	Other firms (n=427,243)	Overall (n=585,047)
Administrative and Support Service Activities	1252 (0.8%)	4346 (1.0%)	5598 (1.0%)
Arts, Entertainment and Recreation	24 (0.0%)	1427 (0.3%)	1451 (0.2%)
Construction of Buildings	3370 (2.1%)	17946 (4.2%)	21316 (3.6%)
Education	733 (0.5%)	3158 (0.7%)	3891 (0.7%)
Financial and Insurance Activities	2903 (1.8%)	2528 (0.6%)	5431 (0.9%)
Human Health and Social Work Activities	650 (0.4%)	2847 (0.7%)	3497 (0.6%)
Information and Communication	1397 (0.9%)	2685 (0.6%)	4082 (0.7%)
Manufacturing	927 (0.6%)	2416 (0.6%)	3343 (0.6%)
Other Service Activities	1207 (0.8%)	2546 (0.6%)	3753 (0.6%)
Professional, Scientific and Technical Activities	7961 (5.0%)	9897 (2.3%)	17858 (3.1%)
Public Administration and Defense: Compulsory Social Security	161 (0.1%)	816 (0.2%)	977 (0.2%)
Real Estate Activities	17169 (10.9%)	32330 (7.6%)	49499 (8.5%)
Transportation and Storage	188 (0.1%)	4038 (0.9%)	4226 (0.7%)
Water Supply	158 (0.1%)	840 (0.2%)	998 (0.2%)
Wholesale and Retail Trade	5932 (3.8%)	32767 (7.7%)	38699 (6.6%)
Accommodation and Food Service Activities	0 (0%)	5581 (1.3%)	5581 (1.0%)
Activities of households as employers	0 (0%)	6 (0.0%)	6 (0.0%)

A.3 Coarsened Exact Matching (CEM)

A.3.1 Direct effects: Main sample of cluster and non-cluster firms

Table A.6 presents descriptive statistics comparing matched and unmatched firms after applying the Coarsened Exact Matching (CEM) method on different firm-level covariates (firm age, firm size, industry) on the pre-treatment period (2007). Both matched and unmatched firms are divided into two groups: LECC and non-LECC firms. The matching process resulted in 722 LECC firms being successfully matched to an equal number of non-LECC firms, thereby creating a total sample size of 1444 firms. In contrast, a significant portion of the initial sample, comprising 583,609 firms, remained unmatched. Among these unmatched firms, 29 belong to the LECC category, while 583,574 are non-LECC firms.

Table A.6: Descriptive statistics: Matched vs unmatched-firms after CEM

	Matched firms			Unmatched firms		
	LECC (n=722)	Non-LECC (n=722)	Overall (n=1444)	LECC (n=29)	Non-LECC (n=583574)	Overall (n=583603)
Firm age	16.7 (20.3)	16.7 (20.1)	16.7 (20.2)	92.2 (61.5)	18.0 (21.9)	18.0 (21.9)
Firm size	15.2 (2.72)	15.2 (2.72)	15.2 (2.72)	22.7 (1.80)	12.4 (3.64)	12.4 (3.64)
Patent (simple)	10.8 (29.9)	7.12 (11.6)	9.56 (25.4)	448 (910)	0.0536 (7.05)	0.0759 (9.97)
Patent (adjusted)	7.93 (53.4)	2.38 (14.2)	5.15 (39.2)	1230 (2650)	0.158 (21.3)	0.220 (29.4)
Joint patents	3.02 (25.2)	1.31 (4.21)	2.44 (20.7)	240 (473)	0.0219 (3.05)	0.0338 (4.79)
Co-partners	3.90 (18.4)	2.88 (9.21)	3.55 (15.9)	386 (819)	0.0384 (5.17)	0.0576 (8.14)
Fixed assets	75.4 (717)	72.7 (588)	74.0 (655)	17200 (31400)	5.82 (194)	6.67 (315)
Total assets	137 (994)	154 (1090)	145 (1040)	25000 (37800)	9.79 (299)	11.0 (435)
Turnover	271 (840)	268 (885)	270 (860)	13700 (23700)	5.67 (211)	6.35 (284)
Sales	257 (789)	267 (863)	262 (824)	12900 (23200)	4.70 (171)	5.34 (252)
Employee costs	43.5 (155)	31.9 (76.3)	38.2 (125)	2740 (4250)	0.792 (20.8)	0.928 (40.9)

Notes: Fixed & total assets, turnover, sales, and employee costs values are in Mln.

Table A.6 shows that the average age of matched firms is similar across LECC and Non-LECC categories, with an average age of 16.7 years. In contrast, unmatched LECC firms are substantially older, with an average age of 92.2 years. Matched firms, regardless of LECC status, display comparable sizes, with an average size of around 15.2. Unmatched firms, however, present a dichotomy in terms of size. LECC firms among the unmatched group have an average size of around 22.7, indicating they are relatively large. On the other hand, non-LECC firms among the unmatched group have a smaller average size of 12.4, suggesting they are relatively small.

Despite matching has not been implemented on innovation (e.g. patents (simple and quality-adjusted)) and collaboration variables (joint patents, co-partners), we observe that matched firms exhibit lower patent counts compared to unmatched firms. This trend is consistent across all financial indicators (i.e. fixed assets, total assets, turnover, sales, and labor costs), indicating that matched firms tend to have lower levels compared to their unmatched counterparts.

Further, table A.7 compares the distribution of firms across different industries between matched and unmatched groups after employing the CEM. Notably, the matching process dropped mostly LECC firms that were not represented in certain industries by dropping comparable non-LECC firms as well.

Table A.7: Descriptive statistics: Matched vs unmatched-firms after CEM

Industry	Matched firms LECC (n=722)	Matched firms Non-LECC (n=722)	Matched firms Overall (n=1444)	Unmatched firms LECC (n=29)	Unmatched firms Non-LECC (n=583574)	Unmatched firms Overall (n=583603)
Administrative and Support Service Activities	5 (0.7%)	4 (0.6%)	9 (0.6%)	0 (0%)	5504 (0.1%)	5504 (0.1%)
Construction of Buildings	3 (0.4%)	1 (0.1%)	4 (0.3%)	0 (0%)	21348 (0.3%)	21348 (0.3%)
Education	3 (0.4%)	3 (0.4%)	6 (0.4%)	0 (0%)	3885 (0.0%)	3885 (0.0%)
Financial and Insurance Activities	2 (0.3%)	0 (0%)	2 (0.1%)	1 (3.4%)	5403 (0.1%)	5404 (0.1%)
Human Health and Social Work Activities	4 (0.6%)	3 (0.4%)	7 (0.5%)	0 (0%)	3564 (0.0%)	3564 (0.0%)
Information and Communication	25 (3.5%)	19 (2.6%)	44 (3.0%)	0 (0%)	4093 (0.1%)	4093 (0.1%)
Manufacturing	10 (1.4%)	8 (1.1%)	18 (1.2%)	1 (3.4%)	3270 (0.0%)	3271 (0.0%)
Other Service Activities	4 (0.6%)	3 (0.4%)	7 (0.5%)	0 (0%)	3656 (0.0%)	3656 (0.0%)
Professional, Scientific and Technical Activities	25 (3.5%)	28 (3.9%)	53 (3.7%)	0 (0%)	17978 (0.2%)	17978 (0.2%)
Real Estate Activities	5 (0.7%)	5 (0.7%)	10 (0.7%)	0 (0%)	49489 (0.6%)	49489 (0.6%)
Transportation and Storage	4 (0.6%)	2 (0.3%)	6 (0.4%)	0 (0%)	4197 (0.1%)	4197 (0.1%)
Water Supply	3 (0.4%)	3 (0.4%)	6 (0.4%)	0 (0%)	1012 (0.0%)	1012 (0.0%)
Wholesale and Retail Trade	14 (1.9%)	16 (2.2%)	30 (2.1%)	0 (0%)	38698 (0.5%)	38698 (0.5%)

Overall, the outcomes of both tables suggest that the matching on selected variables worked successfully. Unmatched firms appear to be older, larger, and have higher patent counts and financial indicators compared to matched firms.

A.3.2 SMEs and large firms

The same procedure of Coarsened Exact Matching is implemented for the sub-samples of SMEs and large firms by choosing the same set of covariates (i.e. firm age, size, industry) (please, check Chapter 1.5 for more details).

Table A.8: Descriptive statistics: SMEs sample

	Before Matching LECC (n=496)	Before Matching Non-LECC (n=654,879)	Before Matching Overall (n=655,375)	After Matching LECC (n=496)	After Matching Non-LECC (n=496)	After Matching Overall (n=992)
Firm age	11.1 (10.2)	15.8 (19.2)	15.8 (19.2)	11.0 (10.1)	11.0 (10.1)	11.0 (10.1)
Firm size	13.9 (1.48)	12.6 (1.87)	12.6 (1.87)	13.7 (2.19)	13.7 (2.19)	13.7 (2.19)
Patent (simple)	2.59 (2.77)	2.18 (13.0)	2.19 (12.9)	3.97 (11.9)	2.16 (1.34)	3.54 (10.4)
Patent (adjusted)	7.86 (9.60)	5.77 (40.0)	5.80 (39.7)	1.66 (16.6)	0.233 (1.54)	0.944 (11.8)
Joint patents	0.492 (1.63)	0.134 (2.61)	0.138 (2.60)	0.600 (2.14)	0.288 (1.08)	0.521 (1.93)
Co-partners	1.06 (4.13)	0.232 (4.48)	0.240 (4.48)	1.27 (4.79)	0.788 (2.67)	1.15 (4.36)
Fixed assets	0.516 (0.884)	0.385 (1.05)	0.385 (1.05)	0.490 (0.834)	0.584 (1.09)	0.537 (0.971)
Total assets	2.30 (2.48)	1.09 (1.82)	1.09 (1.82)	2.19 (2.31)	2.18 (2.29)	2.18 (2.30)
Turnover	6.91 (8.09)	4.46 (18.7)	4.46 (18.7)	56.6 (294)	7.05 (9.32)	34.3 (219)
Sales	6.67 (8.03)	4.55 (19.0)	4.55 (19.0)	6.71 (8.09)	7.47 (9.56)	7.05 (8.76)
Employee costs	5.10 (4.54)	2.12 (3.23)	2.13 (3.24)	5.13 (4.64)	3.39 (2.83)	4.49 (4.14)

Notes: Fixed & total assets, turnover, sales, and employee costs values are in Mln.

Table A.9: Descriptive statistics: SMEs sample

Industry	Before Matching LECC (n=496)	Before Matching Non-LECC (n=654879)	Before Matching Overall (n=655375)	After Matching LECC (n=496)	After Matching Non-LECC (n=496)	After Matching Overall (n=992)
Administrative and Support Service Activities	5 (1%)	6176 (0.9%)	6181 (0.9%)	5 (1.0%)	7 (1.4%)	12 (1.2%)
Education	3 (0.6%)	4570 (0.7%)	4573 (0.7%)	3 (0.6%)	3 (0.6%)	6 (0.6%)
Financial and Insurance Activities	2 (0.4%)	5292 (0.8%)	5294 (0.8%)	1 (0.2%)	0 (0%)	1 (0.1%)
Human Health and Social Work Activities	3 (0.6%)	3435 (0.5%)	3438 (0.5%)	3 (0.6%)	2 (0.4%)	5 (0.5%)
Information and Communication	21 (4.2%)	4960 (0.8%)	4981 (0.8%)	15 (3.0%)	17 (3.4%)	32 (3.2%)
Manufacturing	9 (1.8%)	3414 (0.5%)	3423 (0.5%)	7 (1.4%)	3 (0.6%)	10 (1.0%)
Other Service Activities	5 (1.0%)	4493 (0.7%)	4498 (0.7%)	5 (1.0%)	3 (0.6%)	8 (0.8%)
Professional, Scientific and Technical Activities	26 (5.2%)	22005 (3.4%)	22031 (3.4%)	28 (5.6%)	30 (6.0%)	58 (5.8%)
Real Estate Activities	1 (0.2%)	52370 (8.0%)	52371 (8.0%)	1 (0.2%)	1 (0.2%)	2 (0.2%)
Wholesale and Retail Trade	15 (3.0%)	41880 (6.4%)	41895 (6.4%)	15 (3.0%)	9 (1.8%)	24 (2.4%)
Accommodation and Food Service Activites	0 (0%)	6766 (1.0%)	6766 (1.0%)	0	0	0
Activities of households as employers	0 (0%)	12 (0.0%)	12 (0.0%)	0	0	0
Arts, Entertainment and Recreation	0 (0%)	1747 (0.3%)	1747 (0.3%)	0	0	0
Construction of Buildings	0 (0%)	24131 (3.7%)	24131 (3.7%)	0	0	0
Public Administration and Defense: Compulsory Social Security	0 (0%)	1018 (0.2%)	1018 (0.2%)	0	0	0
Transportation and Storage	0 (0%)	4537 (0.7%)	4537 (0.7%)	0	0	0
Water Supply	0 (0%)	1032 (0.2%)	1032 (0.2%)	0	0	0

Table A.8, A.9, A.10 and A.11 present summary statistics for the LECC-funded and non-funded SMEs and large firms before and after the matching. Despite due to the sample split into SMEs and large firms, treatment and control groups for both cases become automatically compatible with respect to firm age and size, there is still less compatibility between industries. After matching (columns four and five), the gaps become less and we observe very small differences between LECC-funded and non-funded firms, indicating that matching worked well. The LECC-funded and non-funded SMEs and large firms are now identical with respect to coarsened covariates explicitly matched on. However, they become also more identical in terms of covariates that were not used in the matching procedure.

Table A.10: Descriptive statistics: Large firm sample

	Before Matching LECC (n=269)	Before Matching Non-LECC (n=46,362)	Before Matching Overall (n=46,631)	After Matching LECC (n=241)	After Matching Non-LECC (n=241)	After Matching Overall (n=482)
Firm age	34.7 (39.4)	24.2 (31.1)	24.3 (31.2)	25.3 (25.6)	24.5 (24.8)	24.9 (25.2)
Firm size	18.6 (2.12)	17.4 (1.11)	17.4 (1.12)	18.0 (1.50)	18.0 (1.50)	18.0 (1.50)
Patent (simple)	119 (476)	9.92 (104)	14.8 (145)	15.2 (40.0)	23.4 (82.7)	18.0 (57.9)
Patent (adjusted)	336 (1370)	30.3 (314)	44.0 (427)	19.8 (80.3)	13.9 (99.4)	16.9 (90.3)
Joint patents	42.0 (204)	1.59 (27.7)	2.66 (43.4)	3.47 (9.72)	5.13 (24.7)	4.11 (17.1)
Co-partners	65.5 (345)	2.79 (46.8)	4.45 (73.3)	6.43 (20.0)	9.46 (37.6)	7.60 (28.1)
Fixed assets	2050 (11500)	80.6 (718)	92.0 (1140)	203 (1210)	157 (660)	180 (975)
Total assets	3050 (14500)	129 (1100)	146 (1570)	372 (1660)	380 (1620)	376 (1640)
Turnover	2270 (10000)	154 (994)	175 (1430)	317 (864)	314 (914)	315 (887)
Sales	2270 (10100)	153 (970)	175 (1410)	309 (814)	311 (916)	310 (863)
Employee costs	380 (1700)	18.9 (190)	22.4 (256)	51.0 (161)	33.9 (70.5)	42.8 (127)

Notes: Fixed & total assets, turnover, sales, and employee costs values are in Mln.

Table A.11: Descriptive statistics: Large firm sample

Industry	Before Matching			After Matching		
	LECC (n=269)	Non-LECC (n=46,362)	Overall (n=46,631)	LECC (n=241)	Non-LECC (n=241)	Overall (n=482)
Financial and Insurance Activities	1 (0.4%)	833 (1.7%)	834 (1.6%)	1 (0.4%)	2 (0.8%)	3 (0.6%)
Human Health and Social Work Activities	2 (0.7%)	582 (1.2%)	584 (1.2%)	2 (0.8%)	2 (0.8%)	4 (0.8%)
Information and Communication	6 (2.5%)	205 (0.4%)	211 (0.4%)	6 (2.5%)	5 (2.1%)	11 (2.3%)
Manufacturing	6 (2.5%)	429 (0.9%)	435 (0.9%)	6 (2.5%)	1 (0.4%)	7 (1.5%)
Other Service Activities	3 (1.1%)	199 (0.4%)	202 (0.4%)	1 (0.4%)	2 (0.8%)	3 (0.6%)
Professional, Scientific and Technical Activities	8 (2.9%)	1094 (2.2%)	1102 (2.2%)	8 (3.3%)	7 (2.9%)	15 (3.1%)
Real Estate Activities	4 (1.4%)	10224 (20.3%)	10228 (20.2%)	4 (1.7%)	4 (1.7%)	8 (1.7%)
Transportation and Storage	2 (0.7%)	504 (1.0%)	506 (1.0%)	1 (0.4%)	3 (1.2%)	4 (0.8%)
Water Supply	0 (0%)	139 (0.3%)	139 (0.3%)	0	0	0
Wholesale and Retail Trade	8 (2.9%)	2170 (4.3%)	2178 (4.3%)	5 (2.1%)	7 (2.9%)	12 (2.5%)
Accommodation and Food Service Activities	0 (0%)	123 (0.2%)	123 (0.2%)	0	0	0
Administrative and Support Service Activities	4 (1.7%)	347 (0.7%)	351 (0.7%)	4 (1.7%)	1 (0.4%)	5 (1.0%)
Arts, Entertainment and Recreation	0 (0%)	96 (0.2%)	96 (0.2%)	0	0	0
Construction of Buildings	0 (0%)	695 (1.4%)	695 (1.4%)	0	0	0
Education	0 (0%)	171 (0.3%)	171 (0.3%)	0	0	0
Public Administration and Defense: Compulsory Social Security	0 (0%)	157 (0.3%)	157 (0.3%)	0	0	0

A.3.3 Coarsened Exact Matching: Funded and non-funded cluster members

The same procedure of Coarsened Exact Matching is implemented for the sub-samples of funded and non-funded cluster members by choosing the same set of covariates (i.e. firm age, size, industry) (please, check Chapter 1.5 for more details).

Table A.12: Descriptive statistics: Funded cluster members

	Before Matching			After Matching		
	LECC (n=280)	Non-LECC (n=597248)	Overall (n=597528)	LECC (n=211)	Non-LECC (n=211)	Overall (n=422)
Firm age	23.1 (30.7)	17.6 (21.8)	17.6 (21.8)	17.3 (20.7)	16.9 (19.6)	17.1 (20.1)
Firm size	16.3 (2.93)	13.1 (2.09)	13.1 (2.09)	15.9 (2.55)	15.9 (2.55)	15.9 (2.55)
Patents (simple)	114 (502)	5.51 (70.3)	7.21 (94.7)	36.7 (141)	37.5 (152)	36.9 (144)
Patents (adjusted)	316 (1440)	16.3 (212)	21.0 (279)	95.0 (333)	85.3 (312)	91.9 (325)
Joint patents	30.8 (176)	0.510 (14.5)	0.715 (20.6)	8.28 (47.0)	4.88 (25.3)	7.09 (40.7)
Co-partners	44.8 (270)	0.900 (24.6)	1.20 (33.3)	11.2 (53.8)	7.79 (39.5)	10.0 (49.2)
Fixed assets	833 (4710)	6.22 (199)	6.63 (225)	228 (1640)	266 (2800)	247 (2290)
Total assets	1370 (7610)	10.5 (307)	11.1 (352)	447 (3120)	567 (5150)	507 (4250)
Turnover	2020 (8460)	41.4 (565)	44.4 (656)	1600 (9190)	1480 (9640)	1540 (9400)
Sales	2060 (8540)	50.2 (553)	54.3 (679)	1630 (9210)	1430 (9190)	1530 (9170)
Employee costs	344 (1470)	8.76 (68.2)	9.66 (104)	154 (839)	157 (1050)	155 (948)

Notes: Fixed & total assets, turnover, sales, and employee costs values are in Mln.

Table A.13: Descriptive statistics: Funded cluster members

	Before Matching			After Matching		
Industry	LECC (n=224)	Non-LECC (n=295,573)	Overall (n=295,797)	LECC (n=211)	Non-LECC (n=211)	Overall (n=422)
Administrative and Support Service Activites	32 (0.6%)	148810 (1.1%)	148842 (1.1%)	5 (0.6%)	4 (0.5%)	9 (0.6%)
Construction of Buildings	7 (0.1%)	452799 (3.3%)	452806 (3.3%)	1 (0.1%)	1 (0.1%)	2 (0.1%)
Education	24 (0.5%)	109145 (0.8%)	109169 (0.8%)	0	0	0
Financial and Insurance Activities	15 (0.3%)	199520 (1.5%)	199535 (1.5%)	0	0	0
Human Health and Social Work Activities	28 (0.5%)	90791 (0.7%)	90819 (0.7%)	4 (0.5%)	6 (0.8%)	10 (0.6%)
Information and Communication	141 (2.7%)	100648 (0.7%)	100789 (0.7%)	15 (1.9%)	24 (3.1%)	39 (2.5%)
Manufacturing	92 (1.7%)	69135 (0.5%)	69227 (0.5%)	19 (2.4%)	14 (1.8%)	33 (2.1%)
Other Service Activities	41 (0.8%)	145770 (1.1%)	145811 (1.1%)	7 (0.9%)	6 (0.8%)	13 (0.8%)
Professional, Scientific and Technical Activities	169 (3.2%)	368966 (2.7%)	369135 (2.7%)	26 (3.3%)	26 (3.3%)	52 (3.3%)
Real Estate Activities	25 (0.5%)	1117296 (8.1%)	1117321 (8.1%)	3 (0.4%)	3 (0.4%)	6 (0.4%)
Transportation and Storage	60 (1.1%)	93034 (0.7%)	93094 (0.7%)	2 (0.3%)	3 (0.4%)	5 (0.3%)
Water Supply	6 (0.1%)	20414 (0.1%)	20420 (0.1%)	1 (0.1%)	0 (0%)	1 (0.1%)
Wholesale and Retail Trade	100 (1.9%)	903209 (6.6%)	903309 (6.6%)	11 (1.4%)	9 (1.2%)	20 (1.3%)
Accommodation and Food Service Activites	0 (0%)	155762 (1.1%)	155762 (1.1%)	0	0	0
Activities of households as employers	0 (0%)	341 (0.0%)	341 (0.0%)	0	0	0
Arts, Entertainment and Recreation	0 (0%)	50604 (0.4%)	50604 (0.4%)	0	0	0
Public Administration and Defense: Compulsory Social Security	0 (0%)	20121 (0.1%)	20121 (0.1%)	0	0	0

Table A.12, A.13, A.14 and A.15 present summary statistics for the LECC-funded and non-funded cluster members before and after the matching. Before the matching treatment and control groups for both cases have different industry structures, however are more compatible with respect to firm age and size. After matching (columns four and five), the gaps become less and we observe very small differences between funded and non-funded cluster members, indicating that matching worked well. The LECC-funded and non-funded cluster members are now identical with respect to coarsened covariates explicitly matched on. However, they became also more identical in terms of covariates that were not used in the matching procedure.

Table A.14: Descriptive statistics: Non-funded cluster members

	Before Matching			After Matching		
	LECC (n=765)	Non-LECC (n=597248)	Overall (n=598013)	LECC (n=394)	Non-LECC (n=394)	Overall (n=788)
Firm age	19.3 (27.2)	17.6 (21.8)	17.6 (21.8)	14.0 (16.8)	14.0 (17.0)	14.0 (16.9)
Firm size	15.6 (2.85)	13.1 (2.09)	13.1 (2.10)	14.8 (2.40)	14.8 (2.40)	14.8 (2.39)
Patents (simple)	81.3 (394)	5.51 (70.3)	7.78 (98.0)	7.67 (14.0)	6.14 (10.1)	7.16 (12.8)
Patents(adjusted)	16.3 (212)	22.7 (289)	24.9 (46.5)	20.5 (40.3)	23.4 (44.4)	21,95 (42.5)
Joint patents	21.0 (145)	0.510 (14.5)	0.811 (22.8)	1.36 (4.05)	1.02 (3.28)	1.23 (3.77)
Co-partners	32.9 (244)	0.900 (24.6)	1.37 (38.5)	2.48 (7.26)	2.33 (8.07)	2.42 (7.57)
Fixed assets	727 (6890)	6.22 (199)	7.19 (323)	34.5 (312)	67.6 (681)	51.1 (530)
Total assets	1080 (8720)	10.5 (307)	11.9 (445)	82.7 (589)	89.6 (725)	86.1 (660)
Turnover	1710 (8720)	41.4 (565)	46.9 (760)	180 (721)	227 (970)	200 (834)
Sales	1700 (8750)	50.2 (553)	57.6 (812)	179 (699)	234 (995)	202 (835)
Employee costs	8.76 (68.2)	10.4 (134)	36.5 (169)	25.4 (59.4)	31.6 (133)	27.5 (72)

Notes: Fixed & total assets, turnover, sales, and employee costs values are in Mln.

Table A.15: Descriptive statistics: Non-funded cluster members

	Before Matching			After Matching		
Industry	LECC (n=401)	Non-LECC (n=295,573)	Overall (n=295,974)	LECC (n=394)	Non-LECC (n=394)	Overall (n=788)
Administrative and Support Service Activities	76 (0.9%)	148810 (1.1%)	148886 (1.1%)	12 (0.9%)	11 (0.9%)	23 (0.9%)
Construction of Buildings	31 (0.4%)	452799 (3.3%)	452830 (3.3%)	2 (0.2%)	2 (0.2%)	4 (0.2%)
Education	67 (0.8%)	109145 (0.8%)	109212 (0.8%)	6 (0.5%)	4 (0.3%)	10 (0.4%)
Financial and Insurance Activities	40 (0.5%)	199520 (1.5%)	199560 (1.5%)	5 (0.4%)	8 (0.7%)	13 (0.5%)
Human Health and Social Work Activities	48 (0.6%)	90791 (0.7%)	90839 (0.7%)	9 (0.7%)	7 (0.6%)	16 (0.6%)
Information and Communication	213 (2.5%)	100648 (0.7%)	100861 (0.7%)	41 (3.1%)	35 (2.8%)	76 (3.0%)
Manufacturing	90 (1.1%)	69135 (0.5%)	69225 (0.5%)	17 (1.3%)	13 (1.1%)	30 (1.2%)
Other Service Activities	59 (0.7%)	145770 (1.1%)	145829 (1.1%)	8 (0.6%)	7 (0.6%)	15 (0.6%)
Professional, Scientific and Technical Activities	345 (4.1%)	368966 (2.7%)	369311 (2.7%)	64 (4.9%)	61 (5.0%)	125 (4.9%)
Public Administration and Defense: Compulsory Social Security	10 (0.1%)	20121 (0.1%)	20131 (0.1%)	0	0	0
Real Estate Activities	67 (0.8%)	1117296 (8.1%)	1117363 (8.1%)	6 (0.5%)	2 (0.2%)	8 (0.3%)
Transportation and Storage	25 (0.3%)	93034 (0.7%)	93059 (0.7%)	0	0	0
Water Supply	8 (0.1%)	20414 (0.1%)	20422 (0.1%)	1 (0.1%)	1 (0.1%)	2 (0.1%)
Wholesale and Retail Trade	217 (2.6%)	903209 (6.6%)	903426 (6.6%)	31 (2.4%)	26 (2.1%)	57 (2.3%)
Accommodation and Food Service Activities	0 (0%)	155762 (1.1%)	155762 (1.1%)	0	0	0
Activities of households as employers	0 (0%)	341 (0.0%)	341 (0.0%)	0	0	0
Arts, Entertainment and Recreation	0 (0%)	50604 (0.4%)	50604 (0.4%)	0	0	0

A.4 Main Results: Direct Effects of the LECC on 3 Funded Cohorts

In this section, we present the estimates from a Sun & Abraham Two-Way Fixed Effects (TWFE) model for 3 cohorts funded in 2008, 2010, and 2012. For Cohort 1 (see A.16 and Figure A.16), we bin the last relative periods as 4 to 8. For Cohort 2 (see A.17 and Figure A.17), we bin the pre-periods from -4 to -5 and the post-periods from 4 to 6. Finally, for Cohort 3 (see A.18 and Figure A.18), we only bin from -4 to -7 relative period. Further, for all cohorts we drop the 1st relative period before the treatment.

The results suggest that LECC had a positive significant impact on the number of quality-adjusted patents, fixed assets, and employment costs, indicating that being in the treated group is associated with higher values for these variables compared to the control group. However, we notice that the impacts are more significant for the initial cohort, diminishing gradually with subsequent cohorts.

Table A.16: Sun&Abraham TWFE estimate for the Cohort 1

Dependent Variables:	Patent (adjusted)	Joint patents	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)
Variables				
Year = -3	0.0233	0.0391	-0.1845	-0.0563
	(0.1084)	(0.0587)	(0.3807)	(0.0983)
Year = -2	0.1319	0.0342	-0.0816	0.0338
	(0.0885)	(0.0507)	(0.3187)	(0.0817)
Year = 0	0.1547*	0.0684	0.3759	0.2283***
	(0.0831)	(0.0546)	(0.2486)	(0.0763)
Year = 1	0.1569*	0.1215**	0.7392***	0.2762***
	(0.0854)	(0.0548)	(0.2526)	(0.0924)
Year = 2	0.1660**	0.0695	0.9871***	0.3899***
	(0.0846)	(0.0587)	(0.2344)	(0.0811)
Year = 3	0.2336**	0.0895	0.9480***	0.4344***
	(0.0923)	(0.0564)	(0.2399)	(0.0841)
Year = 4	0.1557**	-0.1004**	0.8949***	0.4175***
	(0.0667)	(0.0393)	(0.2039)	(0.0757)
ATT	0.1658**	-0.0125	0.8320***	0.3780***
	(0.0769)	(0.0425)	(0.2752)	(0.0853)
Fixed-effects				
Firm	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	9,099	9,099	9,081	3,805
R^2	0.81651	0.63788	0.81665	0.92384
Within R^2	0.00330	0.01266	0.00885	0.01994

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Figure A.4: Event study: Staggered treatment for funded firms (Cohort 1)

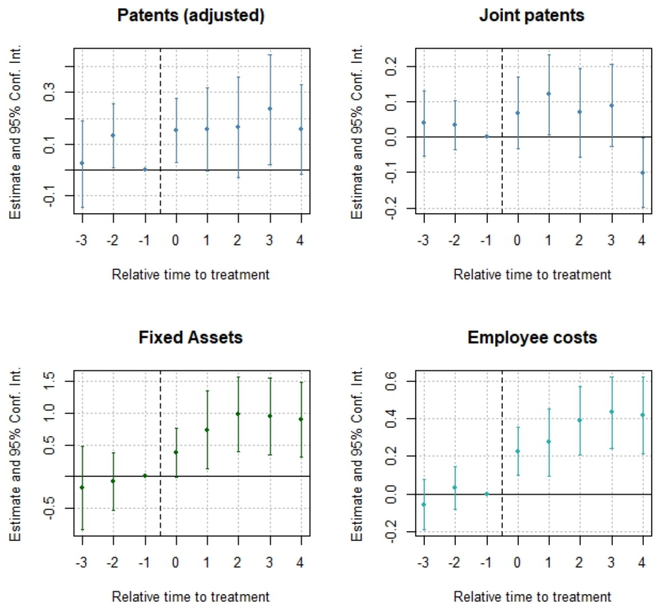


Table A.17: Sun&Abraham TWFE estimate for the Cohort 2

Dependent Variables:	Patent (adjusted)	Joint patents	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)
Variables				
Year = -4	0.0026	-0.0205	-0.2042*	-0.0189
	(0.0335)	(0.0224)	(0.1184)	(0.0544)
Year = -3	0.0255	-0.0003	-0.1559	0.0641
	(0.0367)	(0.0253)	(0.1213)	(0.0564)
Year = -2	0.0475	-0.0102	-0.0946	0.0180
	(0.0383)	(0.0250)	(0.1106)	(0.0631)
Year = 0	0.0460	0.0336	0.1735	0.1900***
	(0.0351)	(0.0252)	(0.1092)	(0.0533)
Year = 1	0.0488	0.0278	0.0506	0.2434***
	(0.0350)	(0.0276)	(0.1108)	(0.0559)
Year = 2	0.1043***	-0.0072	0.1077	0.2917***
	(0.0371)	(0.0276)	(0.1132)	(0.0616)
Year = 3	0.1146***	-0.0179	-0.0104	0.2688***
	(0.0389)	(0.0268)	(0.1178)	(0.0564)
Year = 4	0.0569*	-0.0523**	0.0053	0.2742***
	(0.0325)	(0.0209)	(0.1030)	(0.0563)
ATT	0.0687**	-0.0141	0.0524	0.2580***
	(0.0297)	(0.0224)	(0.1039)	(0.0635)
Fixed-effects				
Firm	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	12,714	12,714	12,671	4,873
R^2	0.80496	0.54090	0.83633	0.93360
Within R^2	0.00174	0.00232	0.00131	0.01624

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Figure A.5: Event study: Staggered treatment for funded firms (Cohort 2)

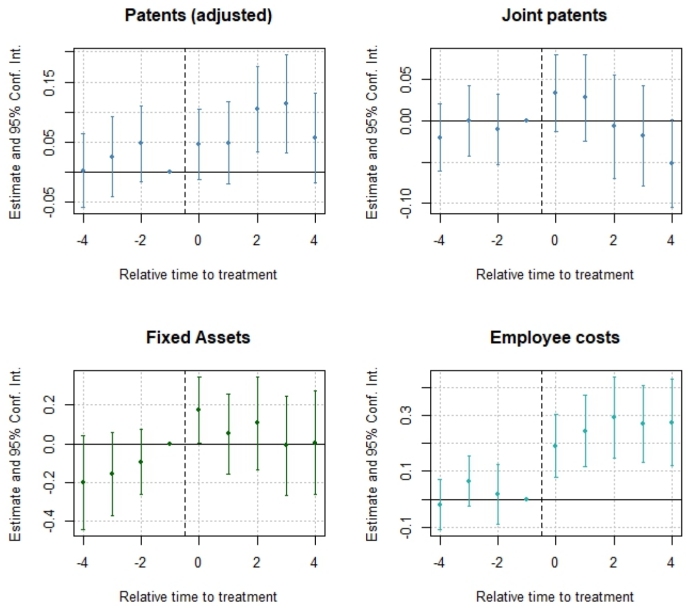
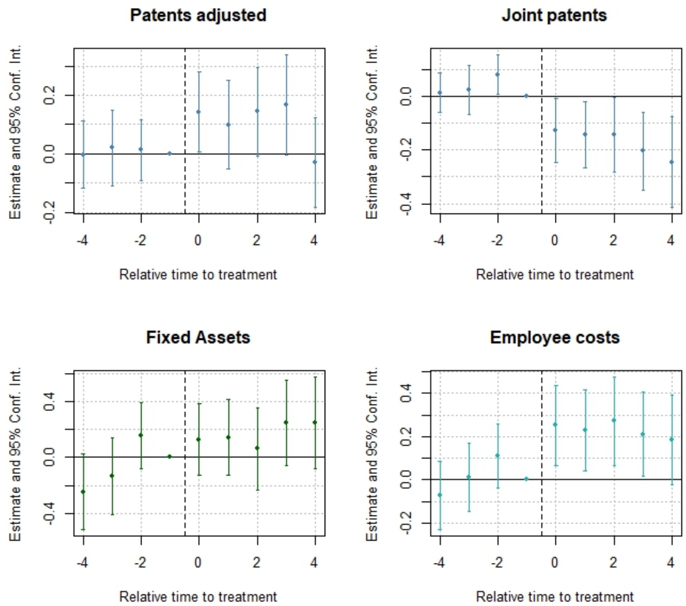


Table A.18: Sun&Abraham TWFE estimate for the Cohort 3

Dependent Variables:	Patent (adjusted)	Joint patents	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)
Variables				
Year = -4	-0.0028	0.0124	-0.2470*	-0.0739
	(0.0590)	(0.0398)	(0.1280)	(0.0823)
Year = -3	0.0201	0.0240	-0.1329	0.0106
	(0.0746)	(0.0554)	(0.1428)	(0.0893)
Year = -2	0.0129	0.0806	0.1583	0.1109
	(0.0709)	(0.0538)	(0.1397)	(0.0893)
Year = 0	0.1431*	-0.1291**	0.1292	0.2513***
	(0.0778)	(0.0516)	(0.1442)	(0.0948)
Year = 1	0.0995	-0.1430***	0.1436	0.2294**
	(0.0805)	(0.0508)	(0.1438)	(0.0918)
Year = 2	0.1442*	-0.1430**	0.0610	0.2718***
	(0.0772)	(0.0575)	(0.1528)	(0.1010)
Year = 3	0.1684**	-0.2039***	0.2490	0.2103**
	(0.0818)	(0.0572)	(0.1543)	(0.0967)
Year = 4	-0.0300	-0.2465***	0.2487	0.1859*
	(0.0802)	(0.0714)	(0.1667)	(0.1049)
ATT	0.1065*	-0.1721***	0.1657	0.2306**
	(0.0624)	(0.0648)	(0.1364)	(0.0925)
Fixed-effects				
Firm	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	9,175	9,175	9,157	3,882
R^2	0.85438	0.64151	0.84616	0.92761
Within R^2	0.00398	0.01777	0.00203	0.01203

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Figure A.6: Event study: Staggered treatment for funded firms (Cohort 3)



A.5 Heterogeneity Analysis

A.5.1 Funded vs non-funded cluster members: SMEs and large firms

In addition to heterogeneity analysis targeting the effects of the LECC, we observe firm size characteristics specifically for funded and non-funded cluster members. We subsample funded and non-funded firms into SMEs and large firms based on the median of total assets by categorizing firms up to the threshold into SMEs, and above into the large firms.

Table A.19 presents descriptive statistics comparing funded and non-funded cluster members, categorized by SMEs and large firms. Table A.19 provides information on the number of firms in each category, with 200 SMEs and 222 large firms among the funded cluster members, and 405 SMEs and 381 large firms among the non-funded cluster members. In general, we notice a significant difference in capital size between funded and non-funded cluster members, indicating that funded cluster members tend to have at least twice as large total and fixed assets compared to non-funded cluster members.

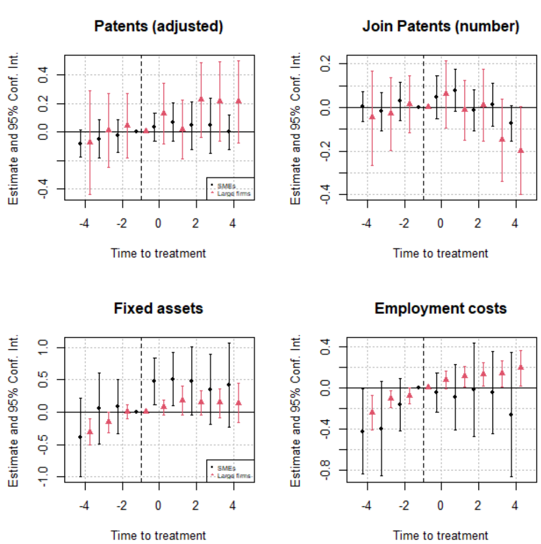
Table A.19: Descriptive statistics on funded and non-funded cluster members

	Funded Cluster Members		Non-Funded Cluster Members	
	SMEs	Large	SMEs	Large
Total Assets	2,230,729	1,409,956,530	926,967.4	160,828,790.2
Fixed Assets	535651.4	761,854,701.6	203,200.9	95,429,122.0
Firm Number	200	222	405	381
Observations	2499	2499	4324	4325

1. SMEs and large funded cluster members

Figure A.7 and Table A.20 present the results of a Sun & Abraham Two-Way Fixed Effects (TWFE) model for funded cluster members, categorized by SMEs and large firms, across different cohorts.

Figure A.7: Direct impact of the LECC policy on funded cluster member SMEs and large firms



Interestingly, detailed analysis in Table A.20 shows, that the notably large effect on fixed assets is primarily driven by SMEs rather than large firms. This observation suggests that the injection of funding has a meaningful impact on the capital investments of SMEs, indicating their capacity for expansion and growth in response to financial stimuli.

Table A.20: Sun&Abraham TWFE estimate for all cohorts of funded cluster members

	SMEs				Large firms			
Dependent Variables:	Patent (adjusted)	Joint patents	Fixed assets	Employment costs	Patent (adjusted)	Joint patents	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Year = -4	-0.0842	0.0037	-0.3923*	-0.4236**	-0.0760	-0.0481	-0.3107***	-0.2395***
	(0.0541)	(0.0345)	(0.2337)	(0.1916)	(0.1682)	(0.0899)	(0.0845)	(0.0739)
Year = -3	-0.0488	-0.0181	0.0593	-0.3953*	0.0108	-0.0310	-0.1563*	-0.1049*
	(0.0657)	(0.0477)	(0.2528)	(0.2160)	(0.1423)	(0.0937)	(0.0894)	(0.0634)
Year = -2	-0.0285	0.0282	0.0845	-0.1651	0.0433	0.0143	0.0026	-0.0757
	(0.0679)	(0.0466)	(0.2477)	(0.1564)	(0.1388)	(0.0821)	(0.0894)	(0.0657)
Year = 0	0.0349	0.0466	0.4782**	-0.0482	0.1273	0.0606	0.0694	0.0748
	(0.0618)	(0.0457)	(0.2335)	(0.1547)	(0.1317)	(0.0906)	(0.0846)	(0.0586)
Year = 1	0.0692	0.0786*	0.5099***	-0.0893	0.0177	-0.0129	0.1784*	0.1099**
	(0.0662)	(0.0460)	(0.1925)	(0.1486)	(0.1180)	(0.0790)	(0.1018)	(0.0532)
Year = 2	0.0435	-0.0135	0.4807**	-0.0171	0.2260*	0.0097	0.1483*	0.1319**
	(0.0770)	(0.0421)	(0.2356)	(0.1993)	(0.1286)	(0.0883)	(0.0852)	(0.0604)
Year = 3	0.0453	0.0138	0.3528	-0.0428	0.2137*	-0.1500*	0.1421	0.1355**
	(0.0860)	(0.0457)	(0.2155)	(0.1670)	(0.1279)	(0.0857)	(0.0898)	(0.0616)
Year = 4	-0.0014	-0.0735*	0.4244*	-0.2625	0.2132*	-0.1990**	0.1397	0.1902***
	(0.0626)	(0.0399)	(0.2417)	(0.2207)	(0.1210)	(0.0827)	(0.1017)	(0.0621)
ATT	0.0275	-0.0115	0.4432*	-0.1326	0.1728	-0.0952	0.1370	0.1452**
	(0.0569)	(0.0381)	(0.2523)	(0.2018)	(0.1148)	(0.0746)	(0.1080)	(0.0604)
CATT (cohort = 2008)	-0.0469	-0.0824	0.7070	-0.0823	0.1103	0.0391	0.3827*	0.1529
	(0.1358)	(0.1068)	(0.5154)	(0.3417)	(0.1990)	(0.1117)	(0.2053)	(0.1050)
CATT (cohort = 2010)	0.0832	0.0340	0.3678	-0.2550	0.1586	-0.1662	-0.0041	0.1456**
	(0.0593)	(0.0296)	(0.3067)	(0.2155)	(0.1091)	(0.1081)	(0.1132)	(0.0597)
CATT (cohort = 2012)	-0.0471	-0.0449	0.1039	0.0856	0.3349	-0.2324	-0.0904	0.1296
	(0.1265)	(0.0659)	(0.2661)	(0.2249)	(0.3128)	(0.1675)	(0.1249)	(0.1153)
Fixed-effects								
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Observations	2,499	2,499	2,499	433	2,499	2,499	2,499	2,201
R^2	0.66123	0.45012	0.70747	0.93811	0.82303	0.78500	0.94077	0.93181
Within R^2	0.01079	0.02387	0.01254	0.13780	0.02118	0.02571	0.03980	0.04841

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

2. SMEs and large non-funded cluster members

In parallel, when examining non-funded SMEs and large firms (see Figure A.8 and Table A.21), it becomes clear that SMEs are driving a substantial effect observed on employment costs. This finding suggests that SMEs, despite not receiving funding, may prioritize investments in human resources and labor-related expenses, potentially reflecting their focus on workforce development and operational sustainability. This distinction underscores the importance of considering firm size when analyzing the effects of funding initiatives, as the dynamics of capital and labor cost allocation and utilization may differ significantly between SMEs and large corporations.

Figure A.8: Direct impact of the LECC policy on non-funded cluster member SMEs and large firms

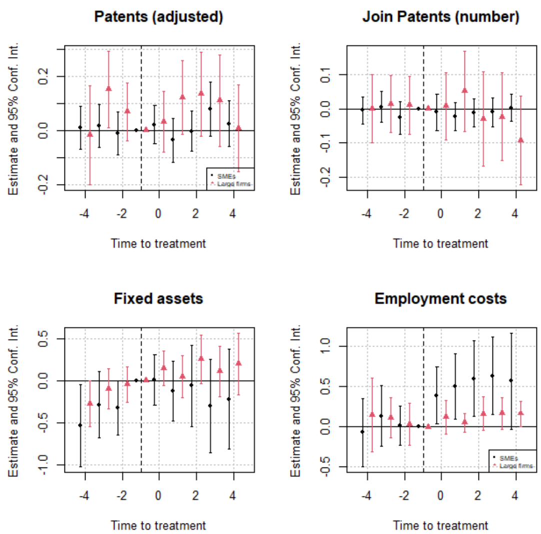


Table A.21: Sun&Abraham TWFE estimate for all cohorts of non-funded cluster members

	SMEs				Large firms			
Dependent Variables:	Patent (adjusted)	Joint patents	Fixed assets	Employment costs	Patent (adjusted)	Joint patents	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Year = -4	0.0096	-0.0052	-0.5325***	-0.0731	-0.0171	0.0009	-0.2680**	0.1451
	(0.0395)	(0.0196)	(0.1995)	(0.1954)	(0.0849)	(0.0510)	(0.1110)	(0.2300)
Year = -3	0.0182	0.0047	-0.2840	0.1331	0.1527*	0.0144	-0.0939	0.1126
	(0.0383)	(0.0249)	(0.2084)	(0.1804)	(0.0798)	(0.0540)	(0.1101)	(0.1317)
Year = -2	-0.0103	-0.0260	-0.3199	0.0146	0.0697	0.0108	-0.0437	0.0303
	(0.0415)	(0.0236)	(0.2091)	(0.1569)	(0.0811)	(0.0553)	(0.1238)	(0.1371)
Year = 0	0.0226	-0.0101	0.0122	0.3880**	0.0338	0.0073	0.1485	0.1190
	(0.0373)	(0.0233)	(0.1999)	(0.1909)	(0.0756)	(0.0538)	(0.1138)	(0.1109)
Year = 1	-0.0355	-0.0235	-0.1257	0.5006***	0.1229*	0.0510	0.0508	0.0499
	(0.0373)	(0.0203)	(0.1872)	(0.1920)	(0.0690)	(0.0566)	(0.1204)	(0.0650)
Year = 2	-0.0015	-0.0123	-0.0592	0.5973***	0.1354*	-0.0296	0.2553*	0.1620
	(0.0370)	(0.0228)	(0.2127)	(0.2253)	(0.0757)	(0.0567)	(0.1336)	(0.1077)
Year = 3	0.0795*	-0.0102	-0.2956	0.6330***	0.1112	-0.0237	0.1099	0.1648*
	(0.0455)	(0.0217)	(0.2267)	(0.2167)	(0.0753)	(0.0540)	(0.1253)	(0.0973)
Year = 4	0.0261	0.0019	-0.2178	0.5652**	0.0098	-0.0935*	0.1975	0.1626**
	(0.0368)	(0.0198)	(0.2032)	(0.2291)	(0.0665)	(0.0501)	(0.1403)	(0.0673)
ATT	0.0178	-0.0085	-0.1472	0.5294**	0.0628	-0.0386	0.1648	0.1390*
	(0.0368)	(0.0192)	(0.2107)	(0.2078)	(0.0657)	(0.0540)	(0.1391)	(0.0747)
CATT (cohort = 2008)	0.0333	0.1676**	-0.2790	0.6779**	0.0442	-0.0400	0.8145*	0.1360
	(0.1724)	(0.0814)	(0.3022)	(0.3034)	(0.0936)	(0.0944)	(0.4359)	(0.1120)
CATT (cohort = 2010)	0.0178	-0.0369*	-0.1475	0.5328**	0.0353	-0.0322	-0.0426	0.0239
	(0.0363)	(0.0206)	(0.2368)	(0.2639)	(0.0929)	(0.0707)	(0.1219)	(0.0589)
CATT (cohort = 2012)	-0.0056	0.0063	0.0594	0.0234	0.2032	-0.0624	-0.0185	0.5272
	(0.0854)	(0.0238)	(0.2571)	(0.2114)	(0.1397)	(0.0867)	(0.1746)	(0.4203)
Fixed-effects								
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Observations	4,324	4,324	4,324	445	4,325	4,325	4,325	2,715
R^2	0.66281	0.45399	0.73184	0.93796	0.79779	0.58026	0.85557	0.88359
Within R^2	0.00961	0.01188	0.00533	0.15100	0.01277	0.00666	0.00856	0.01428

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

A.5.2 Spillover effects: SMEs and large firms

We analyze the heterogeneous spillover effect of the LECC on neighbor firms within 0-5 km (see Figure A.9 and Table A.22) and 5-10 km (see Figure A.10 and Table A.23) distances by categorizing them into SMEs and large firms based on median total assets definition.

Figure A.9: Spillover impact of the LECC policy on SMEs and large firms within 0-5 km

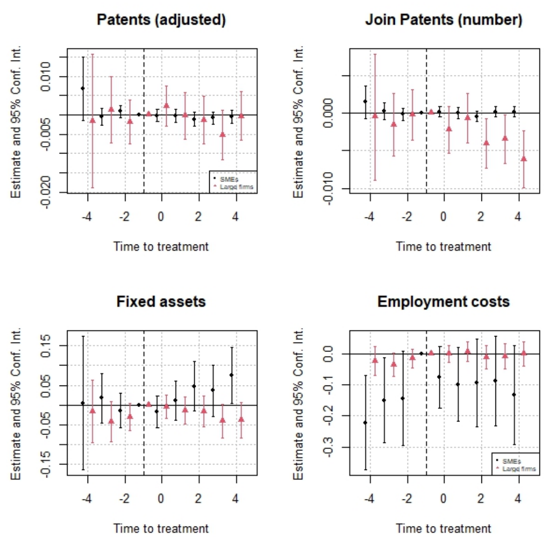


Table A.22: Dynamic spillover effects of the LECC for non-funded neighbor firms within 0-5 km

	SMEs				Large firms			
Dependent Variables:	Patent (adjusted)	Joint patents	Fixed assets	Employment costs	Patent (adjusted)	Joint patents	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Year = -4	0.0067**	0.0014	0.0054	-0.2221***	-0.0016	-0.0005	-0.0159	-0.0232
	(0.0033)	(0.0010)	(0.0709)	(0.0790)	(0.0076)	(0.0037)	(0.0329)	(0.0222)
Year = -3	-0.0006	0.0002	0.0175	-0.1502**	0.0013	-0.0016	-0.0418	-0.0346
	(0.0011)	(0.0005)	(0.0359)	(0.0762)	(0.0046)	(0.0023)	(0.0289)	(0.0224)
Year = -2	0.0009	-0.0002	-0.0135	-0.1440	-0.0018	-0.0003	-0.0305	-0.0156
	(0.0009)	(0.0004)	(0.0303)	(0.0876)	(0.0035)	(0.0019)	(0.0235)	(0.0200)
Year = 0	-0.0002	0.0002	-0.0177	-0.0752	0.0023	-0.0022	-0.0035	0.0002
	(0.0008)	(0.0004)	(0.0276)	(0.0627)	(0.0032)	(0.0018)	(0.0206)	(0.0169)
Year = 1	-0.0002	5.52×10^{-6}	0.0115	-0.0983	-0.0002	-0.0007	-0.0131	0.0069
	(0.0008)	(0.0004)	(0.0274)	(0.0679)	(0.0032)	(0.0018)	(0.0195)	(0.0170)
Year = 2	-0.0011	-0.0005	0.0476*	-0.0928	-0.0014	-0.0041**	-0.0165	-0.0108
	(0.0008)	(0.0003)	(0.0278)	(0.0777)	(0.0032)	(0.0018)	(0.0193)	(0.0166)
Year = 3	-0.0008	0.0002	0.0367	-0.0869	-0.0052*	-0.0035**	-0.0401**	-0.0078
	(0.0008)	(0.0003)	(0.0283)	(0.0786)	(0.0031)	(0.0017)	(0.0192)	(0.0172)
Year = 4	-0.0004	0.0002	0.0751***	-0.1335	-0.0003	-0.0062***	-0.0374**	-0.0002
	(0.0008)	(0.0003)	(0.0257)	(0.0814)	(0.0027)	(0.0015)	(0.0170)	(0.0149)
ATT	-0.0005	5.3×10^{-5}	0.0426	-0.1069*	-0.0007	-0.0046***	-0.0291	-0.0016
	(0.0007)	(0.0003)	(0.0266)	(0.0636)	(0.0028)	(0.0015)	(0.0186)	(0.0174)
Fixed-effects								
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Observations	658,017	658,017	658,017	23,615	658,017	658,017	658,017	134,254
R^2	0.50448	0.44822	0.85688	0.92637	0.71698	0.57260	0.86361	0.94360
Within R^2	3.13×10^{-5}	1.82×10^{-5}	4.59×10^{-5}	0.00108	1.19×10^{-5}	8.44×10^{-5}	2.36×10^{-5}	8.12×10^{-5}

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The table A.22 examines the LECC impact on non-funded neighbor firms within 0-5 kilometers of funded clusters, focusing on SMEs and large firms. While SMEs experience no significant impact in patenting activity, they show an increased investment in fixed assets by up to 7.5% following the LECC. On the contrary, the effects completely disappear for distant neighbors (within a 5-10 km radius), affecting both non-funded neighbor SMEs and large firms, as observed earlier in the entire sample (see Figure A.10 and Table A.23).

Figure A.10: Spillover impact of the LECC policy on SMEs and large firms within 5-10 km

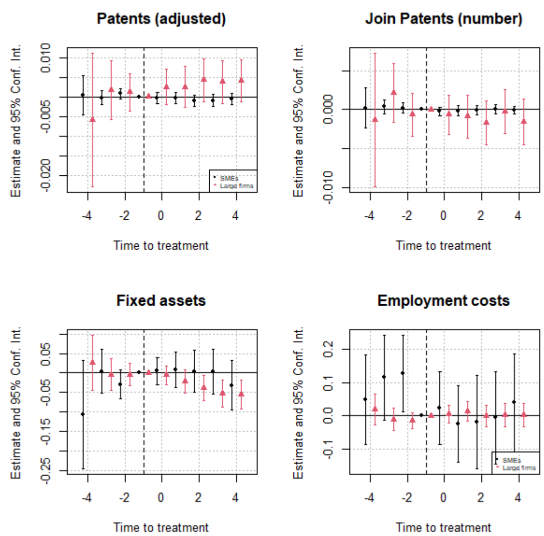


Table A.23: Dynamic spillover effects of the LECC for non-funded neighbor firms within 5-10 km

	SMEs				Large firms			
Dependent Variables:	Patent (adjusted)	Joint patents	Fixed assets	Employment costs	Patent (adjusted)	Joint patents	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Year = -4	0.0005	0.0002	-0.1077*	0.0481	-0.0058	-0.0013	0.0263	0.0191
	(0.0026)	(0.0013)	(0.0599)	(0.0684)	(0.0068)	(0.0035)	(0.0282)	(0.0202)
Year = -3	-0.0001	0.0004	0.0046	0.1151*	0.0019	0.0021	-0.0044	-0.0109
	(0.0009)	(0.0004)	(0.0319)	(0.0693)	(0.0040)	(0.0021)	(0.0229)	(0.0191)
Year = -2	0.0009	0.0002	-0.0290	0.1277**	0.0013	-0.0006	-0.0046	-0.0140
	(0.0008)	(0.0003)	(0.0272)	(0.0644)	(0.0030)	(0.0016)	(0.0194)	(0.0169)
Year = 0	-0.0002	-0.0002	0.0055	0.0235	0.0026	-0.0007	-0.0062	0.0052
	(0.0008)	(0.0003)	(0.0245)	(0.0621)	(0.0028)	(0.0015)	(0.0170)	(0.0155)
Year = 1	-0.0002	-0.0001	0.0083	-0.0234	0.0025	-0.0009	-0.0215	0.0140
	(0.0007)	(0.0003)	(0.0243)	(0.0652)	(0.0027)	(0.0015)	(0.0164)	(0.0159)
Year = 2	-0.0009	-3.5×10^{-5}	0.0046	-0.0194	0.0043	-0.0017	-0.0389**	-0.0006
	(0.0007)	(0.0003)	(0.0246)	(0.0759)	(0.0027)	(0.0015)	(0.0162)	(0.0159)
Year = 3	-0.0008	4.68×10^{-5}	0.0037	-0.0059	0.0039	-0.0003	-0.0527***	0.0020
	(0.0007)	(0.0003)	(0.0250)	(0.0714)	(0.0027)	(0.0015)	(0.0161)	(0.0160)
Year = 4	-0.0005	-6.16×10^{-5}	-0.0315	0.0389	0.0042*	-0.0015	-0.0549***	0.0024
	(0.0007)	(0.0003)	(0.0227)	(0.0716)	(0.0023)	(0.0012)	(0.0141)	(0.0143)
ATT	-0.0005	-7.25×10^{-5}	-0.0101	0.0119	0.0038	-0.0012	-0.0441***	0.0036
	(0.0006)	(0.0002)	(0.0238)	(0.0604)	(0.0024)	(0.0012)	(0.0152)	(0.0154)
Fixed-effects								
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Observations	720,476	720,476	720,476	25,961	720,629	720,629	720,629	145,851
R^2	0.49605	0.45238	0.85568	0.92295	0.71204	0.56164	0.86320	0.93899
Within R^2	1.19×10^{-5}	7.32×10^{-6}	1.99×10^{-5}	0.00093	1.63×10^{-5}	1.33×10^{-5}	7.47×10^{-5}	6.06×10^{-5}

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

A.6 Robustness Checks

In this section, we conduct robustness checks by incorporating additional fixed effects into our analysis. These fixed effects, namely Industry-Year and State-Year, allow us to account for potential unobserved heterogeneity at the industry and state levels. By including these fixed effects, we aim to strengthen the validity of our findings by controlling for any confounding factors that may vary across different industries or states over time. This helps ensure that our estimated treatment effects are robust and not driven by factors unrelated to the treatment variable. Robustness checks with additional fixed effects are applied to both the analysis of direct effects and spillover effects, as well as to various heterogeneity analyses.

A.6.1 Direct effects

Table A.24 presents the direct effects of the LECC for funded cluster firms, with three different models (I, II, III) and various dependent variables including adjusted patents, joint patents, fixed assets, and employment costs. When comparing the models, it's evident that the inclusion of additional fixed effects (such as Industry-Year and State-Year) improves the fit statistics, increasing the R-squared values and indicating better model performance. These fixed effects help to account for unobserved heterogeneity and potential biases, enhancing the accuracy of the estimates. Overall, these results underscore the significant impact of the LECC on the innovation and economic performance of funded cluster firms, with additional fixed effects enhancing the accuracy of the estimates.

Table A.24: Direct effects of the LECC for funded cluster firms

	I				II				III			
Dependent Variables:	Patent(adjusted)	Joint patents	Fixed assets	Employment costs	Patent(adjusted)	Joint patents	Fixed assets	Employment costs	Patent(adjusted)	Joint patents	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variables												
Year = -4	-0.0141	-0.0270	-0.1782**	-0.1124**	-0.0330	-0.0426*	-0.1899*	-0.1382***	-0.0252	-0.0485*	-0.1211	-0.1240**
	(0.0299)	(0.0202)	(0.0891)	(0.0472)	(0.0353)	(0.0238)	(0.1044)	(0.0530)	(0.0397)	(0.0261)	(0.1101)	(0.0544)
Year = -3	0.0125	-0.0101	-0.1447	-0.0502	0.0309	-0.0146	-0.1473	-0.0552	0.0400	-0.0211	-0.1897	-0.0578
	(0.0312)	(0.0215)	(0.0903)	(0.0360)	(0.0370)	(0.0249)	(0.1098)	(0.0450)	(0.0432)	(0.0283)	(0.1169)	(0.0505)
Year = -2	0.0303	0.0010	-0.0495	-0.0270	0.0323	-0.0114	-0.0293	-0.0333	0.0269	-0.0099	-0.0283	-0.0398
	(0.0320)	(0.0216)	(0.0921)	(0.0415)	(0.0382)	(0.0248)	(0.1113)	(0.0468)	(0.0432)	(0.0280)	(0.1176)	(0.0497)
Year = 0	0.0578*	-0.0053	0.1973**	0.1351***	0.0604*	-0.0072	0.2646**	0.1106**	0.0649	0.0002	0.2597**	0.1138**
	(0.0305)	(0.0217)	(0.0871)	(0.0369)	(0.0356)	(0.0259)	(0.1072)	(0.0498)	(0.0400)	(0.0297)	(0.1121)	(0.0562)
Year = 1	0.0434	-0.0045	0.1663**	0.1503***	0.0575*	0.0143	0.1543	0.1396***	0.0610	0.0213	0.2179**	0.1400***
	(0.0298)	(0.0220)	(0.0789)	(0.0315)	(0.0346)	(0.0255)	(0.0979)	(0.0401)	(0.0402)	(0.0302)	(0.1057)	(0.0521)
Year = 2	0.0919***	-0.0384	0.2698***	0.2250***	0.1262***	-0.0329	0.2268**	0.1950***	0.1542***	-0.0144	0.2981***	0.2125***
	(0.0320)	(0.0237)	(0.0875)	(0.0401)	(0.0381)	(0.0264)	(0.1076)	(0.0518)	(0.0431)	(0.0316)	(0.1144)	(0.0583)
Year = 3	0.1133***	-0.0488**	0.2047**	0.2088***	0.1163***	-0.0572**	0.1621	0.2045***	0.1286***	-0.0437	0.2115*	0.2126***
	(0.0337)	(0.0231)	(0.0881)	(0.0366)	(0.0392)	(0.0258)	(0.1070)	(0.0456)	(0.0446)	(0.0304)	(0.1123)	(0.0517)
Year = 4	0.0564*	-0.0940***	0.2730***	0.2372***	0.0492	-0.0904***	0.2017*	0.2374***	0.0752*	-0.0840***	0.2635**	0.2061***
	(0.0308)	(0.0200)	(0.0956)	(0.0419)	(0.0363)	(0.0225)	(0.1135)	(0.0530)	(0.0400)	(0.0260)	(0.1237)	(0.0604)
ATT	0.0678**	-0.0515**	0.2342**	0.2017***	0.0729**	-0.0478**	0.2022*	0.1910***	0.0729**	-0.0478**	0.2022*	0.1910***
	(0.0284)	(0.0211)	(0.0952)	(0.0465)	(0.0347)	(0.0229)	(0.1145)	(0.0509)	(0.0347)	(0.0229)	(0.1145)	(0.0509)
CATT (cohort = 2008)	0.0730	-0.0273	0.7193***	0.2531***	0.0610	0.0014	0.5818**	0.2701***	0.1250	0.0006	0.7693***	0.2011*
	(0.0745)	(0.0450)	(0.2485)	(0.0689)	(0.0840)	(0.0475)	(0.2553)	(0.0890)	(0.0870)	(0.0547)	(0.2948)	(0.1101)
CATT (cohort = 2010)	0.0573*	-0.0346	0.0598	0.1661***	0.0709*	-0.0387	0.0327	0.1645**	0.0722	-0.0194	0.0745	0.1943***
	(0.0308)	(0.0237)	(0.1000)	(0.0525)	(0.0382)	(0.0266)	(0.1314)	(0.0646)	(0.0466)	(0.0319)	(0.1253)	(0.0631)
CATT (cohort = 2012)	0.1085*	-0.1753***	0.1978	0.2234**	0.1040	-0.1793***	0.3309**	0.1374	0.1152	-0.1987***	0.1832	0.1198
	(0.0627)	(0.0649)	(0.1387)	(0.0932)	(0.0694)	(0.0682)	(0.1386)	(0.0876)	(0.0750)	(0.0729)	(0.1557)	(0.0921)
Fixed-effects												
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year					Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Year									Yes	Yes	Yes	Yes
Fit statistics												
Observations	15,654	15,654	15,611	6,592	15,654	15,654	15,611	6,592	15,654	15,654	15,611	6,592
R^2	0.81800	0.61432	0.84057	0.93237	0.85212	0.69508	0.86531	0.96071	0.85415	0.70113	0.86884	0.96277
Within R^2	0.00392	0.01254	0.00729	0.02403	0.00449	0.01023	0.00560	0.02683	0.00513	0.01116	0.00813	0.02189

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

A.6.2 Spillover effects

The table A.25 presents the estimated effects of the LECC policy on non-funded neighbor firms within a 0-5 km radius across different outcomes such as patents, joint patents, fixed assets, and employment costs. Overall, the results indicate that the LECC had a significant impact on the fixed assets of non-funded neighbors within 5 km. When including extra fixed effects such as Industry-Year and State-Year, the coefficients remain largely similar, indicating that the LECC policy's impact on these outcomes is consistent even after controlling for additional factors. However, the inclusion of these fixed effects helps to account for industry-specific and state-specific variations, providing a more robust analysis of the LECC policy's effects.

Further, despite the inclusion of these additional fixed effects, the R-squared values remain relatively high, suggesting that the model with extra fixed effects captures a substantial portion of the variation in the outcomes. This indicates that the model's explanatory power is robust even after accounting for industry-specific and state-specific variations, further reinforcing the credibility of the estimated effects.

Table A.25: Dynamic spillover effects of the LECC for non-funded neighbor firms within 0-5 km

	I				II				III			
Dependent Variables:	Patent(adjusted)	Joint patents	Fixed assets	Employment costs	Patent(adjusted)	Joint patents	Fixed assets	Employment costs	Patent(adjusted)	Joint patents	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variables												
Year = -4	0.0026	0.0009	0.0466	-0.0142	-0.0007	-9.94×10^{-5}	0.0348	-0.0283	0.0003	0.0008	0.0309	-0.0253
	(0.0049)	(0.0023)	(0.0350)	(0.0220)	(0.0047)	(0.0023)	(0.0354)	(0.0217)	(0.0050)	(0.0024)	(0.0360)	(0.0225)
Year = -3	-1.58×10^{-5}	-0.0002	0.0103	-0.0368*	0.0006	-0.0003	-0.0023	-0.0324	0.0003	-0.0002	0.0020	-0.0371*
	(0.0019)	(0.0009)	(0.0256)	(0.0222)	(0.0018)	(0.0009)	(0.0257)	(0.0220)	(0.0019)	(0.0010)	(0.0264)	(0.0223)
Year = -2	8.74×10^{-5}	-0.0001	-0.0188	-0.0157	0.0004	-0.0002	-0.0233	-0.0121	0.0002	-0.0003	-0.0307	-0.0166
	(0.0015)	(0.0008)	(0.0221)	(0.0199)	(0.0015)	(0.0008)	(0.0221)	(0.0197)	(0.0016)	(0.0008)	(0.0226)	(0.0204)
Year = 0	0.0011	-0.0005	-0.0060	-0.0019	0.0014	-0.0007	0.0005	0.0013	0.0012	-0.0008	-0.0012	0.0074
	(0.0014)	(0.0008)	(0.0195)	(0.0174)	(0.0014)	(0.0008)	(0.0196)	(0.0171)	(0.0015)	(0.0008)	(0.0201)	(0.0172)
Year = 1	0.0008	7.58×10^{-5}	0.0139	0.0065	0.0008	3.99×10^{-5}	0.0246	-0.0042	0.0006	4.36×10^{-6}	0.0256	0.0024
	(0.0014)	(0.0008)	(0.0188)	(0.0178)	(0.0014)	(0.0008)	(0.0189)	(0.0175)	(0.0015)	(0.0008)	(0.0193)	(0.0177)
Year = 2	-0.0004	-0.0014*	0.0326*	-0.0104	-0.0001	-0.0015*	0.0497***	-0.0148	-0.0004	-0.0016*	0.0556***	-0.0127
	(0.0014)	(0.0008)	(0.0185)	(0.0174)	(0.0014)	(0.0008)	(0.0186)	(0.0171)	(0.0015)	(0.0008)	(0.0191)	(0.0174)
Year = 3	-0.0013	-0.0008	0.0105	-0.0170	-0.0013	-0.0010	0.0307	-0.0162	-0.0016	-0.0010	0.0407**	-0.0118
	(0.0014)	(0.0007)	(0.0186)	(0.0178)	(0.0013)	(0.0007)	(0.0187)	(0.0174)	(0.0014)	(0.0008)	(0.0192)	(0.0177)
Year = 4	0.0019	-0.0014**	0.0425***	-0.0007	0.0011	-0.0026***	0.0677***	-0.0010	0.0007	-0.0028***	0.0682***	-0.0036
	(0.0013)	(0.0007)	(0.0164)	(0.0160)	(0.0012)	(0.0007)	(0.0165)	(0.0155)	(0.0013)	(0.0007)	(0.0169)	(0.0157)
ATT	0.0009	-0.0010	0.0273	-0.0033	0.0006	-0.0017**	0.0466***	-0.0048	0.0003	-0.0018**	0.0489***	-0.0039
	(0.0013)	(0.0007)	(0.0179)	(0.0180)	(0.0012)	(0.0007)	(0.0179)	(0.0174)	(0.0013)	(0.0007)	(0.0184)	(0.0177)
Fixed-effects												
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year					Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Year									Yes	Yes	Yes	Yes
Fit statistics												
Observations	1,352,023	1,352,023	1,351,840	161,222	1,352,023	1,352,023	1,351,840	161,222	1,352,023	1,352,023	1,351,840	161,222
R^2	0.69075	0.55170	0.87388	0.94579	0.69717	0.55934	0.87537	0.94770	0.69741	0.55971	0.87542	0.94788
Within R^2	8.6×10^{-6}	1.22×10^{-5}	2.3×10^{-5}	7.19×10^{-5}	4.93×10^{-6}	3.42×10^{-5}	4.94×10^{-5}	7.08×10^{-5}	3.88×10^{-6}	3.87×10^{-5}	5.13×10^{-5}	7.41×10^{-5}

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.26 highlights, that despite the incorporation of extra fixed effects, such as Industry-Year and State-Year, the coefficients and R-squared values for the spillover analysis on distant neighbors (5-10 km) remain largely unchanged. This consistency suggests that the observed effects in this distance range are not significantly impacted by the LECC policy. Consequently, it appears that the LECC policy does not have a discernible influence on distant neighbors within the 5-10 km range, as indicated by the stable coefficients and R-squared values across different model specifications.

Table A.26: Dynamic spillover effects of the LECC for non-funded neighbor firms within 5-10 km

	I				II				III			
Dependent Variables:	Patent(adjusted)	Joint patents	Fixed assets	Employment costs	Patent(adjusted)	Joint patents	Fixed assets	Employment costs	Patent(adjusted)	Joint patents	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variables												
Year = -4	-0.0018	0.0005	0.0117	0.0244	-0.0035	0.0003	-0.0022	0.0175	-0.0034	-0.0003	-0.0160	0.0106
	(0.0044)	(0.0022)	(0.0292)	(0.0201)	(0.0043)	(0.0021)	(0.0296)	(0.0198)	(0.0044)	(0.0022)	(0.0300)	(0.0202)
Year = -3	2.12×10^{-5}	0.0012	-0.0346	0.0045	0.0007	0.0011	-0.0378*	-4.5×10^{-5}	0.0010	0.0012	-0.0321	-0.0045
	(0.0017)	(0.0008)	(0.0225)	(0.0191)	(0.0016)	(0.0008)	(0.0225)	(0.0186)	(0.0017)	(0.0008)	(0.0228)	(0.0186)
Year = -2	0.0002	1.41×10^{-5}	-0.0335*	0.0028	0.0006	-8.26×10^{-5}	-0.0343*	-0.0058	0.0011	-0.0001	-0.0311	-0.0057
	(0.0013)	(0.0007)	(0.0191)	(0.0175)	(0.0013)	(0.0007)	(0.0191)	(0.0170)	(0.0013)	(0.0007)	(0.0195)	(0.0171)
Year = 0	0.0003	-0.0005	-0.0174	0.0072	0.0007	-0.0006	-0.0153	0.0021	0.0010	-0.0005	-0.0141	0.0037
	(0.0013)	(0.0006)	(0.0167)	(0.0160)	(0.0012)	(0.0006)	(0.0168)	(0.0155)	(0.0013)	(0.0007)	(0.0170)	(0.0155)
Year = 1	0.0006	-0.0003	-0.0213	0.0204	0.0009	-0.0004	-0.0183	0.0055	0.0015	-0.0004	-0.0214	0.0064
	(0.0012)	(0.0006)	(0.0163)	(0.0166)	(0.0012)	(0.0006)	(0.0163)	(0.0160)	(0.0012)	(0.0007)	(0.0166)	(0.0160)
Year = 2	0.0008	-0.0005	-0.0156	-0.0048	0.0013	-0.0006	-0.0139	-0.0148	0.0019	-0.0006	-0.0201	-0.0165
	(0.0012)	(0.0006)	(0.0161)	(0.0166)	(0.0012)	(0.0006)	(0.0161)	(0.0161)	(0.0012)	(0.0007)	(0.0164)	(0.0163)
Year = 3	0.0008	6.03×10^{-5}	-0.0115	0.0002	0.0010	-0.0001	-0.0105	-0.0048	0.0016	-4.1×10^{-5}	-0.0177	-0.0069
	(0.0012)	(0.0006)	(0.0161)	(0.0165)	(0.0012)	(0.0006)	(0.0161)	(0.0160)	(0.0012)	(0.0006)	(0.0164)	(0.0162)
Year = 4	0.0019*	-4.49×10^{-5}	-0.0276*	0.0080	0.0017	-0.0008	-0.0187	0.0013	0.0022*	-0.0008	-0.0254*	-0.0029
	(0.0011)	(0.0005)	(0.0141)	(0.0151)	(0.0011)	(0.0005)	(0.0141)	(0.0144)	(0.0011)	(0.0006)	(0.0144)	(0.0146)
ATT	0.0013	-0.0002	-0.0219	0.0068	0.0013	-0.0006	-0.0165	-0.0009	0.0018	-0.0006	-0.0218	-0.0033
	(0.0011)	(0.0005)	(0.0153)	(0.0163)	(0.0011)	(0.0005)	(0.0153)	(0.0156)	(0.0011)	(0.0006)	(0.0156)	(0.0156)
Fixed-effects												
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year					Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Year									Yes	Yes	Yes	Yes
Fit statistics												
Observations	1,480,529	1,480,529	1,480,373	175,493	1,480,529	1,480,529	1,480,373	175,493	1,480,529	1,480,529	1,480,373	175,493
R^2	0.68696	0.54223	0.87297	0.94174	0.69242	0.54980	0.87437	0.94361	0.69267	0.55021	0.87444	0.94382
Within R^2	5.79×10^{-6}	5.28×10^{-6}	7.05×10^{-6}	5.01×10^{-5}	5.52×10^{-6}	8.13×10^{-6}	5.38×10^{-6}	4.08×10^{-5}	6.97×10^{-6}	8.16×10^{-6}	4.1×10^{-6}	3.3×10^{-5}

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

A.6.3 Heterogeneity analysis

A.6.3.1 SMEs vs large firms

Upon comparing Table A.27 and Table A.28, it seems that the inclusion of extra fixed effects does not substantially alter the coefficients or their statistical significance for the direct effects of the LECC policy on either funded SMEs or large firms. In both cases, the coefficients across different models (I, II, III) and years exhibit similar patterns and significance levels regardless of whether additional fixed effects are included.

The R-squared values for the models both for SMEs and large firms indicate a good fit overall, with values ranging from 0.69 to 0.98 across different specifications. Interestingly, the inclusion of extra fixed effects does not lead to substantial changes in the R-squared values, suggesting that the additional controls do not significantly improve the explanatory power of the models. This implies that the main determinants of the outcomes considered in these analyses are captured well by the existing set of fixed effects and independent variables, and the inclusion of further controls does not substantially enhance the models' ability to explain the variation in the dependent variables.

Table A.27: Direct effects of the LECC policy on funded SMEs

	I				II				III			
Dependent Variables:	Patent(adjusted)	Joint patents	Fixed assets	Employment costs	Patent(adjusted)	Joint patents	Fixed assets	Employment costs	Patent(adjusted)	Joint patents	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variables												
Year = -4	-0.0278	-0.0036	-0.3727**	-0.1213	-0.0268	0.0042	-0.2863*	-0.1563	0.0115	0.0281	-0.2260	-0.2354
	(0.0318)	(0.0167)	(0.1488)	(0.0890)	(0.0333)	(0.0198)	(0.1694)	(0.1095)	(0.0342)	(0.0197)	(0.1844)	(0.1626)
Year = -3	0.0214	0.0004	-0.0677	-0.0556	0.0191	-0.0017	-0.0632	-0.1269	0.0636	0.0131	-0.0756	-0.1533
	(0.0344)	(0.0209)	(0.1519)	(0.0849)	(0.0376)	(0.0231)	(0.1674)	(0.1319)	(0.0415)	(0.0243)	(0.1844)	(0.2045)
Year = -2	-0.0142	-0.0068	-0.0736	-0.0142	-0.0096	-0.0086	-0.0579	-0.0749	0.0103	0.0068	-0.0227	-0.0223
	(0.0366)	(0.0214)	(0.1592)	(0.0768)	(0.0406)	(0.0232)	(0.1733)	(0.1001)	(0.0425)	(0.0249)	(0.1901)	(0.1436)
Year = 0	0.0177	0.0073	0.3198**	0.1065	0.0242	0.0005	0.3464**	0.0824	0.0310	0.0206	0.3674*	0.1195
	(0.0344)	(0.0212)	(0.1560)	(0.0728)	(0.0366)	(0.0233)	(0.1711)	(0.0903)	(0.0402)	(0.0240)	(0.1892)	(0.1111)
Year = 1	0.0140	0.0182	0.3512***	0.1382**	0.0233	0.0247	0.3280**	0.1652*	0.0277	0.0270	0.4042**	0.2630**
	(0.0338)	(0.0210)	(0.1314)	(0.0663)	(0.0359)	(0.0224)	(0.1480)	(0.0905)	(0.0420)	(0.0254)	(0.1711)	(0.1277)
Year = 2	0.0436	-0.0229	0.5188***	0.1737**	0.0552	-0.0324	0.4112**	0.2029*	0.0581	-0.0003	0.5020***	0.3426**
	(0.0381)	(0.0218)	(0.1499)	(0.0785)	(0.0422)	(0.0237)	(0.1619)	(0.1104)	(0.0472)	(0.0265)	(0.1843)	(0.1549)
Year = 3	0.0672	-0.0213	0.4052***	0.1704**	0.0634	-0.0190	0.3167*	0.2408*	0.0793	0.0188	0.3726**	0.3533**
	(0.0411)	(0.0206)	(0.1526)	(0.0809)	(0.0445)	(0.0228)	(0.1635)	(0.1246)	(0.0498)	(0.0252)	(0.1790)	(0.1663)
Year = 4	0.0431	-0.0622***	0.5894***	0.1730*	0.0377	-0.0592***	0.4791***	0.2352	0.0716*	-0.0330	0.5178***	0.1869
	(0.0359)	(0.0198)	(0.1586)	(0.0905)	(0.0400)	(0.0209)	(0.1731)	(0.1496)	(0.0432)	(0.0220)	(0.1935)	(0.2115)
ATT	0.0379	-0.0270	0.4728***	0.1571*	0.0394	-0.0270	0.4016**	0.1967	0.0573	-0.0030	0.4532**	0.2393*
	(0.0328)	(0.0177)	(0.1647)	(0.0881)	(0.0379)	(0.0180)	(0.1785)	(0.1363)	(0.0385)	(0.0209)	(0.1951)	(0.1423)
CATT (cohort = 2008)	0.0876	-0.0133	1.253***	0.2749**	0.0739	0.0045	0.9962**	0.3717*	0.1619	0.0229	1.307***	0.1597
	(0.1069)	(0.0604)	(0.4527)	(0.1288)	(0.1138)	(0.0573)	(0.4353)	(0.1946)	(0.1290)	(0.0650)	(0.5054)	(0.2465)
CATT (cohort = 2010)	0.0229	-0.0295*	0.2651	0.1203	0.0376	-0.0317	0.2549	0.1148	0.0359	-0.0064	0.2392	0.3067*
	(0.0311)	(0.0168)	(0.1698)	(0.1100)	(0.0357)	(0.0196)	(0.1959)	(0.1619)	(0.0377)	(0.0199)	(0.2075)	(0.1620)
CATT (cohort = 2012)	0.0325	-0.0404	0.1869	-0.0488	-0.0274	-0.0647	0.0980	0.0427	-0.0280	-0.0379	0.0426	0.1412
	(0.0744)	(0.0344)	(0.1646)	(0.1282)	(0.0696)	(0.0441)	(0.1796)	(0.1459)	(0.0769)	(0.0469)	(0.2417)	(0.1535)
Fixed-effects												
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year					Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Year									Yes	Yes	Yes	Yes
Fit statistics												
Observations	7,437	7,437	7,430	1,386	7,437	7,437	7,430	1,386	7,437	7,437	7,430	1,386
R^2	0.63813	0.34488	0.71222	0.94625	0.68106	0.41879	0.75077	0.97158	0.69335	0.44614	0.76129	0.98039
Within R^2	0.00417	0.01133	0.01554	0.05226	0.00473	0.01037	0.00982	0.08285	0.00594	0.01021	0.01296	0.11852

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.28: Direct effects of the LECC policy on funded large firms

	I				II				III			
Dependent Variables:	Patent(adjusted)	Joint patents	Fixed assets	Employment costs	Patent(adjusted)	Joint patents	Fixed assets	Employment costs	Patent(adjusted)	Joint patents	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variables												
Year = -4	-0.0655	-0.0367	-0.4253***	0.0174	-0.2212*	-0.2081	-0.5924***	0.0115	-0.2263	-0.2354	-0.5860***	-0.1043
	(0.1017)	(0.1005)	(0.1158)	(0.2265)	(0.1285)	(0.1458)	(0.1701)	(0.2288)	(0.1572)	(0.1843)	(0.2266)	(0.1502)
Year = -3	0.0648	-0.0127	-0.3951***	0.0281	-0.0054	0.0202	-0.3959**	0.0504	-0.0469	-0.0251	-0.5290**	-0.0300
	(0.0941)	(0.0975)	(0.1511)	(0.1378)	(0.1202)	(0.1278)	(0.1700)	(0.1414)	(0.1499)	(0.1576)	(0.2071)	(0.0996)
Year = -2	0.0191	0.0040	-0.2407**	0.0270	0.0602	-0.0399	-0.2249	0.0630	0.0039	-0.0827	-0.2637	-0.0176
	(0.0986)	(0.1029)	(0.1215)	(0.1353)	(0.1278)	(0.1438)	(0.1384)	(0.1398)	(0.1597)	(0.1906)	(0.1805)	(0.1135)
Year = 0	0.0402	0.0647	0.0130	0.2220*	0.0255	0.0612	0.0659	0.2752**	0.0290	0.0055	0.0973	0.2292*
	(0.0954)	(0.1081)	(0.0892)	(0.1220)	(0.1242)	(0.1560)	(0.1241)	(0.1383)	(0.1507)	(0.2039)	(0.1687)	(0.1380)
Year = 1	0.0598	0.0426	0.0392	0.1278**	0.0501	0.0132	0.0190	0.1380**	-0.0294	-0.1583	0.0825	0.1063
	(0.0842)	(0.0918)	(0.1064)	(0.0539)	(0.1154)	(0.1271)	(0.1307)	(0.0616)	(0.1470)	(0.1561)	(0.1575)	(0.0898)
Year = 2	0.1338	0.0017	0.1904*	0.3328***	0.2030	0.0269	0.1646	0.3903***	0.2482	0.0575	0.2240	0.3283**
	(0.0946)	(0.1110)	(0.1033)	(0.1221)	(0.1247)	(0.1455)	(0.1334)	(0.1331)	(0.1597)	(0.1945)	(0.1898)	(0.1290)
Year = 3	0.1976**	-0.0435	0.1152	0.2677***	0.2387**	-0.0710	0.1524	0.3363***	0.2427	-0.1800	0.2483	0.2697***
	(0.0920)	(0.0992)	(0.0973)	(0.0858)	(0.1210)	(0.1355)	(0.1389)	(0.0965)	(0.1530)	(0.1871)	(0.1915)	(0.0954)
Year = 4	0.1450	-0.1265	0.1188	0.3171***	0.1717	-0.1175	0.1814	0.4167***	0.1330	-0.1317	0.1587	0.3611***
	(0.0882)	(0.0968)	(0.1208)	(0.0728)	(0.1147)	(0.1307)	(0.1707)	(0.0900)	(0.1408)	(0.1820)	(0.2180)	(0.1019)
ATT	0.1221	-0.0403	0.1005	0.2688***	0.1453	-0.0421	0.1325	0.3369***	0.1252	-0.0942	0.1603	0.2838***
	(0.0853)	(0.1091)	(0.1127)	(0.0885)	(0.1155)	(0.1315)	(0.1589)	(0.1071)	(0.1511)	(0.1718)	(0.1661)	(0.1073)
CATT (cohort = 2008)	0.0706	0.0610	0.2855	0.2339**	0.0550	0.0315	0.3515	0.1921	0.0286	-0.1370	0.3848	0.2448
	(0.1444)	(0.1359)	(0.1879)	(0.1118)	(0.1851)	(0.1720)	(0.2562)	(0.1383)	(0.2090)	(0.2591)	(0.3237)	(0.1678)
CATT (cohort = 2010)	0.1610	-0.1552	0.0207	0.1828**	0.2531	-0.1150	-0.0066	0.2941**	0.2244	-0.0911	-0.0262	0.2453*
	(0.1230)	(0.1672)	(0.1206)	(0.0875)	(0.1579)	(0.2024)	(0.1704)	(0.1174)	(0.2200)	(0.2569)	(0.1846)	(0.1471)
CATT (cohort = 2012)	0.1121	0.0247	-0.0038	0.5231	0.0347	-0.0146	0.1134	0.6474	0.0448	-0.0370	0.2437	0.4315*
	(0.1616)	(0.1769)	(0.1833)	(0.3854)	(0.2161)	(0.2644)	(0.2177)	(0.4173)	(0.2700)	(0.3303)	(0.2260)	(0.2485)
Fixed-effects												
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year					Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Year									Yes	Yes	Yes	Yes
Fit statistics												
Observations	3,276	2,162	3,276	2,852	3,276	2,162	3,276	2,852	3,276	2,162	3,276	2,852
R^2	0.78255	0.56969	0.79500	0.82030	0.85450	0.74817	0.85901	0.90297	0.87374	0.81080	0.88242	0.92350
Within R^2	0.01136	0.00749	0.02411	0.02415	0.01987	0.01218	0.02786	0.03106	0.02421	0.01573	0.03813	0.02854

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

A.6.3.2 Funded vs non-funded cluster members

When comparing the two tables (see Table A.29 and Table A.28), it's evident that the addition of extra fixed effects does not alter the coefficients significantly. The coefficients for the LECC policy effects remain largely consistent across both tables for funded and non-funded cluster members. This suggests that the inclusion of additional fixed effects does not fundamentally change the estimated effects of the LECC policy on the outcomes of interest.

In terms of R-squared values, the inclusion of extra fixed effects does lead to slight improvements in some cases. For example, in Models II and III for both funded and non-funded cluster members, there are small increases in the R-squared values compared to Model I. This indicates that the additional fixed effects may capture some additional variation in the outcomes, leading to a slightly better model fit.

Table A.29: Direct effects of the LECC on funded cluster members

	I				II				III			
Dependent Variables:	Patent(adjusted)	Joint patents	Fixed assets	Employment costs	Patent(adjusted)	Joint patents	Fixed assets	Employment costs	Patent(adjusted)	Joint patents	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variables												
Year = -4	0.0219	-0.0002	-0.2936**	0.1123	-0.0224	-0.0089	-0.3408**	0.0469	0.0302	0.0099	-0.2228	0.0052
	(0.0459)	(0.0255)	(0.1181)	(0.2185)	(0.0442)	(0.0270)	(0.1368)	(0.1614)	(0.0517)	(0.0312)	(0.1462)	(0.1006)
Year = -3	0.0939**	0.0075	-0.1752	0.1129	0.0849*	0.0119	-0.1677	0.1311	0.1297**	0.0290	-0.2374	0.0432
	(0.0429)	(0.0281)	(0.1300)	(0.1264)	(0.0483)	(0.0302)	(0.1386)	(0.1036)	(0.0624)	(0.0372)	(0.1532)	(0.0688)
Year = -2	0.0408	-0.0107	-0.1365	0.0214	0.0429	-0.0136	-0.1434	0.0637	0.0715	0.0082	-0.1063	-0.0382
	(0.0443)	(0.0285)	(0.1354)	(0.1320)	(0.0493)	(0.0305)	(0.1487)	(0.1008)	(0.0619)	(0.0356)	(0.1550)	(0.0742)
Year = 0	0.0443	-0.0021	0.0820	0.1283	0.0470	-0.0025	0.0780	0.1305	0.0386	0.0146	0.0567	0.0741
	(0.0423)	(0.0283)	(0.1262)	(0.1027)	(0.0460)	(0.0311)	(0.1384)	(0.0850)	(0.0559)	(0.0379)	(0.1496)	(0.0687)
Year = 1	0.0581	0.0164	0.0564	0.0930	0.0590	0.0239	0.0091	0.0984*	0.0548	0.0419	0.0304	0.0718
	(0.0402)	(0.0296)	(0.1144)	(0.0575)	(0.0452)	(0.0326)	(0.1243)	(0.0591)	(0.0566)	(0.0425)	(0.1406)	(0.0854)
Year = 2	0.0976**	-0.0187	0.2862**	0.1943**	0.1082**	-0.0177	0.1946	0.2326***	0.1186**	0.0237	0.2100	0.2137***
	(0.0437)	(0.0307)	(0.1336)	(0.0962)	(0.0472)	(0.0321)	(0.1385)	(0.0847)	(0.0582)	(0.0424)	(0.1496)	(0.0782)
Year = 3	0.1310***	-0.0191	0.1437	0.2046**	0.1411***	-0.0145	0.0607	0.2618***	0.1665***	0.0267	0.0498	0.2334***
	(0.0450)	(0.0291)	(0.1379)	(0.0869)	(0.0479)	(0.0315)	(0.1488)	(0.0808)	(0.0575)	(0.0405)	(0.1516)	(0.0712)
Year = 4	0.0586	-0.0533**	0.2781*	0.1962***	0.0711	-0.0531*	0.1532	0.2784***	0.0918*	-0.0306	0.1373	0.2405***
	(0.0411)	(0.0264)	(0.1439)	(0.0661)	(0.0457)	(0.0276)	(0.1480)	(0.0697)	(0.0539)	(0.0324)	(0.1601)	(0.0717)
ATT	0.0722*	-0.0240	0.1933	0.1697**	0.0807*	-0.0220	0.1110	0.2162***	0.0919*	0.0043	0.1055	0.1817***
	(0.0392)	(0.0277)	(0.1445)	(0.0703)	(0.0440)	(0.0294)	(0.1498)	(0.0712)	(0.0535)	(0.0363)	(0.1516)	(0.0598)
CATT (cohort = 2008)	0.1344	0.0318	0.7931	0.1809*	0.1407	0.0426	0.5369	0.1948*	0.1877	0.0512	0.7032	0.1560
	(0.1145)	(0.0662)	(0.4852)	(0.1077)	(0.1151)	(0.0607)	(0.4164)	(0.1106)	(0.1220)	(0.0676)	(0.4604)	(0.1318)
CATT (cohort = 2010)	0.0375	-0.0340	0.0615	0.0755	0.0507	-0.0384	0.0079	0.1379*	0.0495	-0.0041	-0.0235	0.1557**
	(0.0432)	(0.0321)	(0.1406)	(0.0637)	(0.0501)	(0.0357)	(0.1603)	(0.0803)	(0.0682)	(0.0456)	(0.1542)	(0.0783)
CATT (cohort = 2012)	0.1813*	-0.0525	0.0369	0.4943	0.1652*	-0.0261	0.0567	0.5336*	0.1943**	-0.0190	-0.0646	0.3158*
	(0.1021)	(0.0570)	(0.1442)	(0.4171)	(0.0873)	(0.0591)	(0.1676)	(0.3080)	(0.0913)	(0.0656)	(0.2041)	(0.1642)
Fixed-effects												
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year					Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Year									Yes	Yes	Yes	Yes
Fit statistics												
Observations	8,663	8,663	8,649	3,160	8,663	8,663	8,649	3,160	8,663	8,663	8,649	3,160
R^2	0.77177	0.55473	0.82360	0.91351	0.81770	0.63853	0.84756	0.94704	0.82777	0.65540	0.85504	0.95968
Within R^2	0.00616	0.00528	0.00704	0.01421	0.00726	0.00567	0.00519	0.02563	0.00902	0.00599	0.00676	0.02782

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.30: Direct effects of the LECC on non-funded cluster members

	I				II				III			
Dependent Variables:	Patent(adjusted)	Joint patents	Fixed assets	Employment costs	Patent(adjusted)	Joint patents	Fixed assets	Employment costs	Patent(adjusted)	Joint patents	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variables												
Year = -4	0.0219	-0.0002	-0.2936**	0.1123	-0.0224	-0.0089	-0.3408**	0.0469	0.0302	0.0099	-0.2228	0.0052
	(0.0459)	(0.0255)	(0.1181)	(0.2185)	(0.0442)	(0.0270)	(0.1368)	(0.1614)	(0.0517)	(0.0312)	(0.1462)	(0.1006)
Year = -3	0.0939**	0.0075	-0.1752	0.1129	0.0849*	0.0119	-0.1677	0.1311	0.1297**	0.0290	-0.2374	0.0432
	(0.0429)	(0.0281)	(0.1300)	(0.1264)	(0.0483)	(0.0302)	(0.1386)	(0.1036)	(0.0624)	(0.0372)	(0.1532)	(0.0688)
Year = -2	0.0408	-0.0107	-0.1365	0.0214	0.0429	-0.0136	-0.1434	0.0637	0.0715	0.0082	-0.1063	-0.0382
	(0.0443)	(0.0285)	(0.1354)	(0.1320)	(0.0493)	(0.0305)	(0.1487)	(0.1008)	(0.0619)	(0.0356)	(0.1550)	(0.0742)
Year = 0	0.0443	-0.0021	0.0820	0.1283	0.0470	-0.0025	0.0780	0.1305	0.0386	0.0146	0.0567	0.0741
	(0.0423)	(0.0283)	(0.1262)	(0.1027)	(0.0460)	(0.0311)	(0.1384)	(0.0850)	(0.0559)	(0.0379)	(0.1496)	(0.0687)
Year = 1	0.0581	0.0164	0.0564	0.0930	0.0590	0.0239	0.0091	0.0984*	0.0548	0.0419	0.0304	0.0718
	(0.0402)	(0.0296)	(0.1144)	(0.0575)	(0.0452)	(0.0326)	(0.1243)	(0.0591)	(0.0566)	(0.0425)	(0.1406)	(0.0854)
Year = 2	0.0976**	-0.0187	0.2862**	0.1943**	0.1082**	-0.0177	0.1946	0.2326***	0.1186**	0.0237	0.2100	0.2137***
	(0.0437)	(0.0307)	(0.1336)	(0.0962)	(0.0472)	(0.0321)	(0.1385)	(0.0847)	(0.0582)	(0.0424)	(0.1496)	(0.0782)
Year = 3	0.1310***	-0.0191	0.1437	0.2046**	0.1411***	-0.0145	0.0607	0.2618***	0.1665***	0.0267	0.0498	0.2334***
	(0.0450)	(0.0291)	(0.1379)	(0.0869)	(0.0479)	(0.0315)	(0.1488)	(0.0808)	(0.0575)	(0.0405)	(0.1516)	(0.0712)
Year = 4	0.0586	-0.0533**	0.2781*	0.1962***	0.0711	-0.0531*	0.1532	0.2784***	0.0918*	-0.0306	0.1373	0.2405***
	(0.0411)	(0.0264)	(0.1439)	(0.0661)	(0.0457)	(0.0276)	(0.1480)	(0.0697)	(0.0539)	(0.0324)	(0.1601)	(0.0717)
ATT	0.0722*	-0.0240	0.1933	0.1697**	0.0807*	-0.0220	0.1110	0.2162***	0.0919*	0.0043	0.1055	0.1817***
	(0.0392)	(0.0277)	(0.1445)	(0.0703)	(0.0440)	(0.0294)	(0.1498)	(0.0712)	(0.0535)	(0.0363)	(0.1516)	(0.0598)
CATT (cohort = 2008)	0.1344	0.0318	0.7931	0.1809*	0.1407	0.0426	0.5369	0.1948*	0.1877	0.0512	0.7032	0.1560
	(0.1145)	(0.0662)	(0.4852)	(0.1077)	(0.1151)	(0.0607)	(0.4164)	(0.1106)	(0.1220)	(0.0676)	(0.4604)	(0.1318)
CATT (cohort = 2010)	0.0375	-0.0340	0.0615	0.0755	0.0507	-0.0384	0.0079	0.1379*	0.0495	-0.0041	-0.0235	0.1557**
	(0.0432)	(0.0321)	(0.1406)	(0.0637)	(0.0501)	(0.0357)	(0.1603)	(0.0803)	(0.0682)	(0.0456)	(0.1542)	(0.0783)
CATT (cohort = 2012)	0.1813*	-0.0525	0.0369	0.4943	0.1652*	-0.0261	0.0567	0.5336*	0.1943**	-0.0190	-0.0646	0.3158*
	(0.1021)	(0.0570)	(0.1442)	(0.4171)	(0.0873)	(0.0591)	(0.1676)	(0.3080)	(0.0913)	(0.0656)	(0.2041)	(0.1642)
Fixed-effects												
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year					Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Year									Yes	Yes	Yes	Yes
Fit statistics												
Observations	8,663	8,663	8,649	3,160	8,663	8,663	8,649	3,160	8,663	8,663	8,649	3,160
R^2	0.77177	0.55473	0.82360	0.91351	0.81770	0.63853	0.84756	0.94704	0.82777	0.65540	0.85504	0.95968
Within R^2	0.00616	0.00528	0.00704	0.01421	0.00726	0.00567	0.00519	0.02563	0.00902	0.00599	0.00676	0.02782

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Overall, while the inclusion of extra fixed effects does improve the explanatory power of the models marginally, it doesn't significantly alter the estimated coefficients for the LECC policy. Thus, the additional fixed effects contribute to better-explaining variations in the outcomes without changing the overall interpretation of the LECC policy effects.

APPENDIX B

Chapter 2

B.1 Descriptive Analysis

B.1.1 Funding concentration

To examine the distribution of R&D funding across German states, I calculate the Gini coefficient [1] and plot Lorenz curves for each state (see Figures B.1, B.3 and B.4). The Lorenz curve illustrates the cumulative share of firms against the cumulative share of funding, with the line of equality represented by a 45-degree, upward-sloping line. In the graphs, the line of equality is depicted in red, while the Lorenz curve is shown in blue. The Gini coefficient, which measures inequality, is the gap between the line of equality and the Lorenz curve.

Overall, the analysis reveals that R&D funding is unevenly distributed across German states. In 2000, the Gini coefficient ranges from 0.63 to 0.73, while in 2020, it varies from 0.56 to 0.82. Saarland and Thüringen exhibit the least unequal distribution of funding, with Gini coefficients of 0.63 and 0.64 respectively in 2000, improving to 0.56 and 0.59 by 2020. Conversely, Hamburg, Schleswig-Holstein, Sachsen-Anhalt, Bremen, Rheinland-Pfalz, Mecklenburg-Vorpommern, and Brandenburg display the highest inequality in funding concentration in 2000, with Gini coefficients ranging from 0.68 to 0.72. While most of these states experience an exacerbation in funding inequality by 2020 (ranging from 0.74 up to 0.86), Bremen, Schleswig-Holstein, and Sachsen-Anhalt demonstrate improvements in funding concentration, with their Gini coefficients decreasing to 0.61, 0.62, and 0.69 respectively.

[1] Gini coefficient is a measure of inequality, which ranges from 0 (or 0%) to 1 (or 100%), with 0 representing perfect equality and 1 representing perfect inequality.

Figure B.1: R&D funding concentration in German states

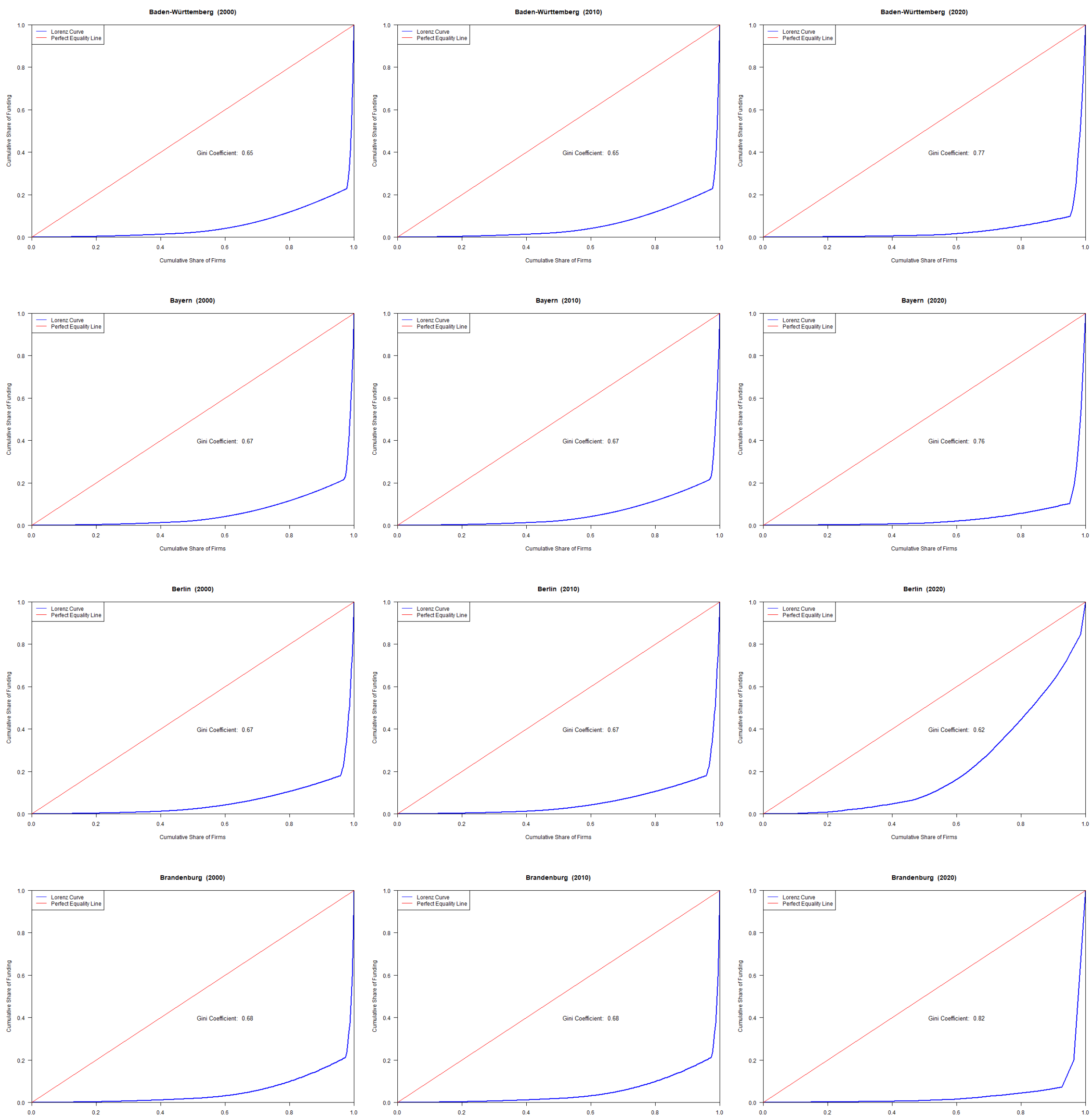


Figure B.2: R&D funding concentration in German states (cont.)

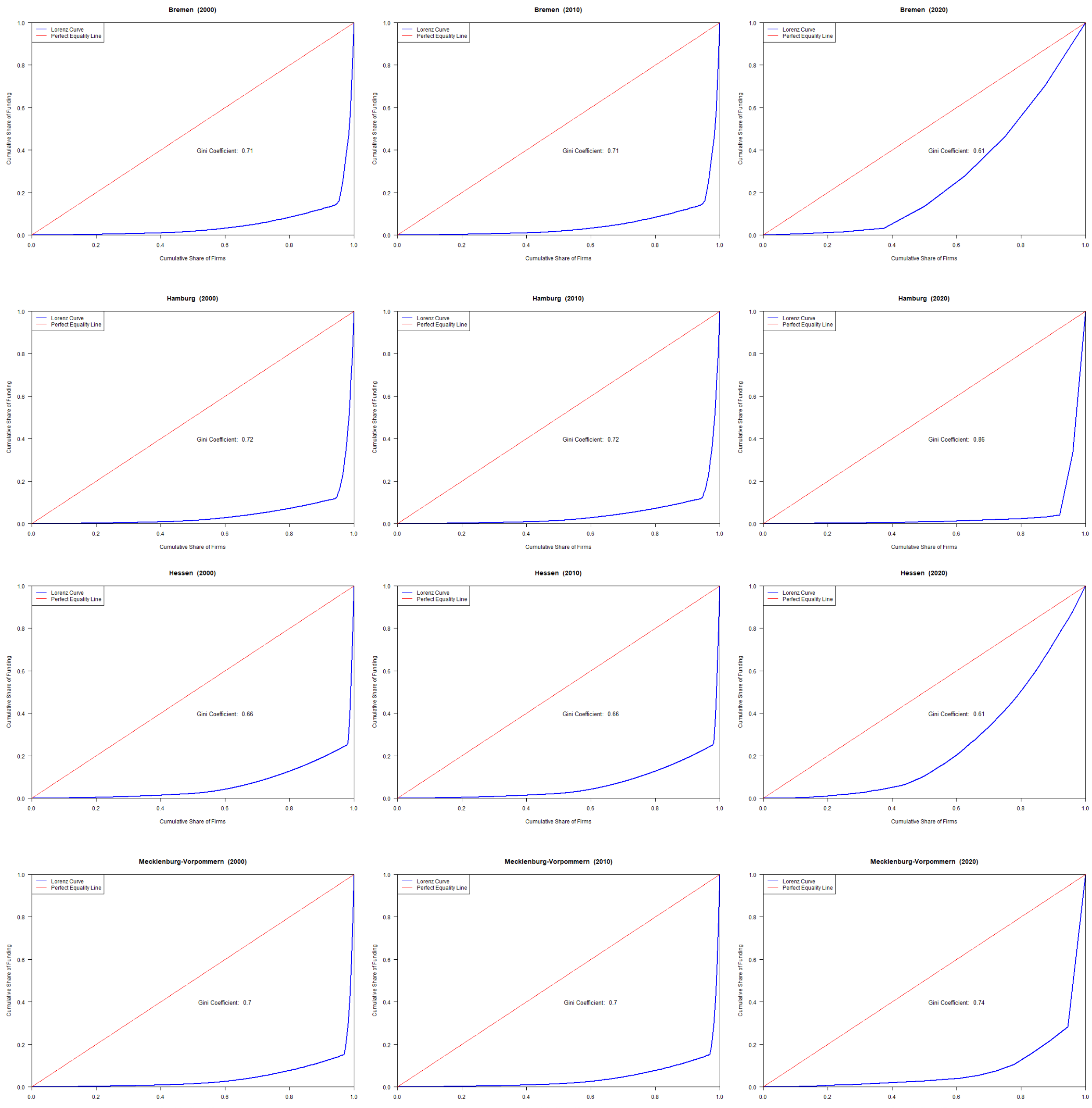


Figure B.3: R&D funding concentration in German states (cont.)

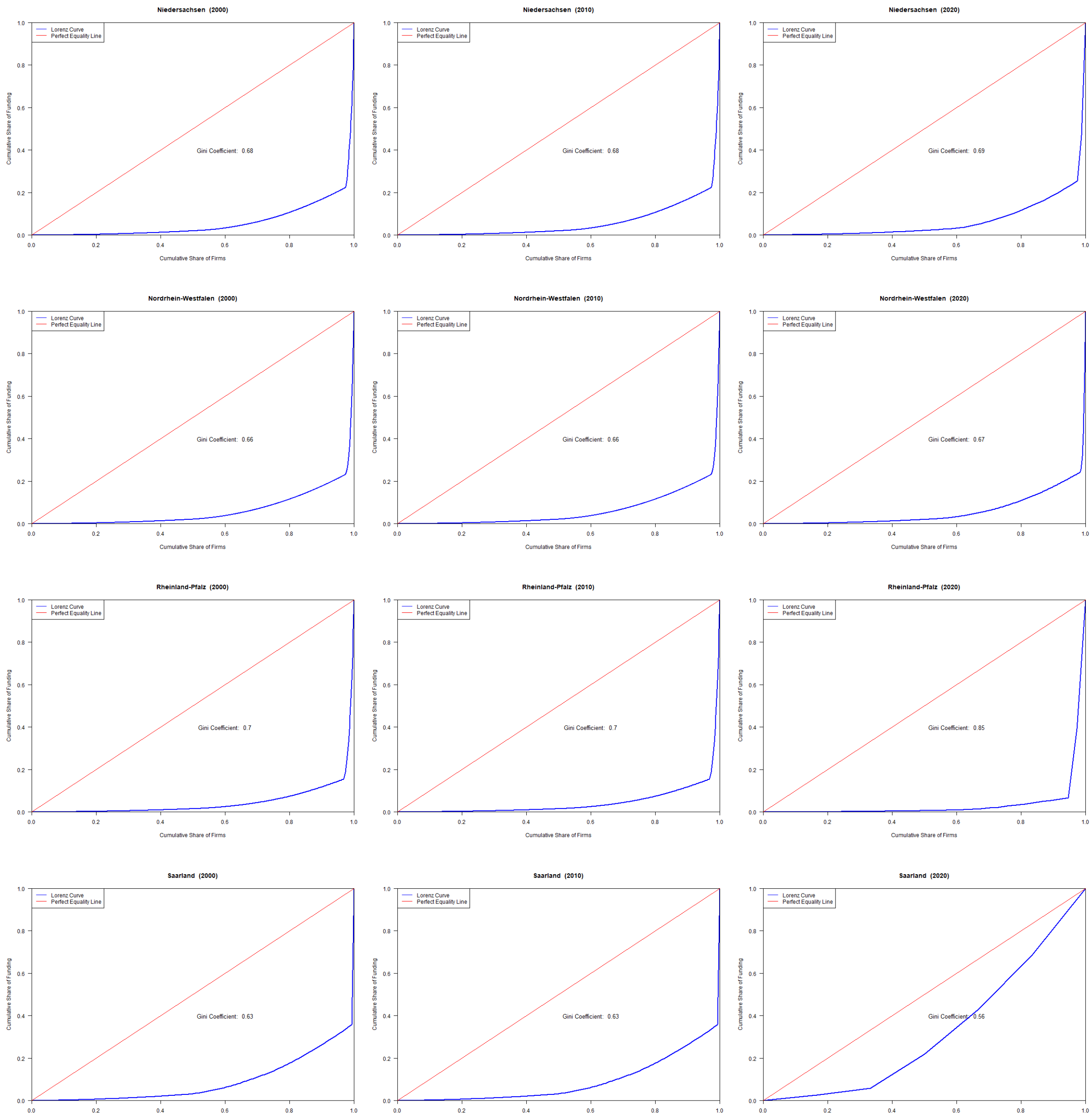
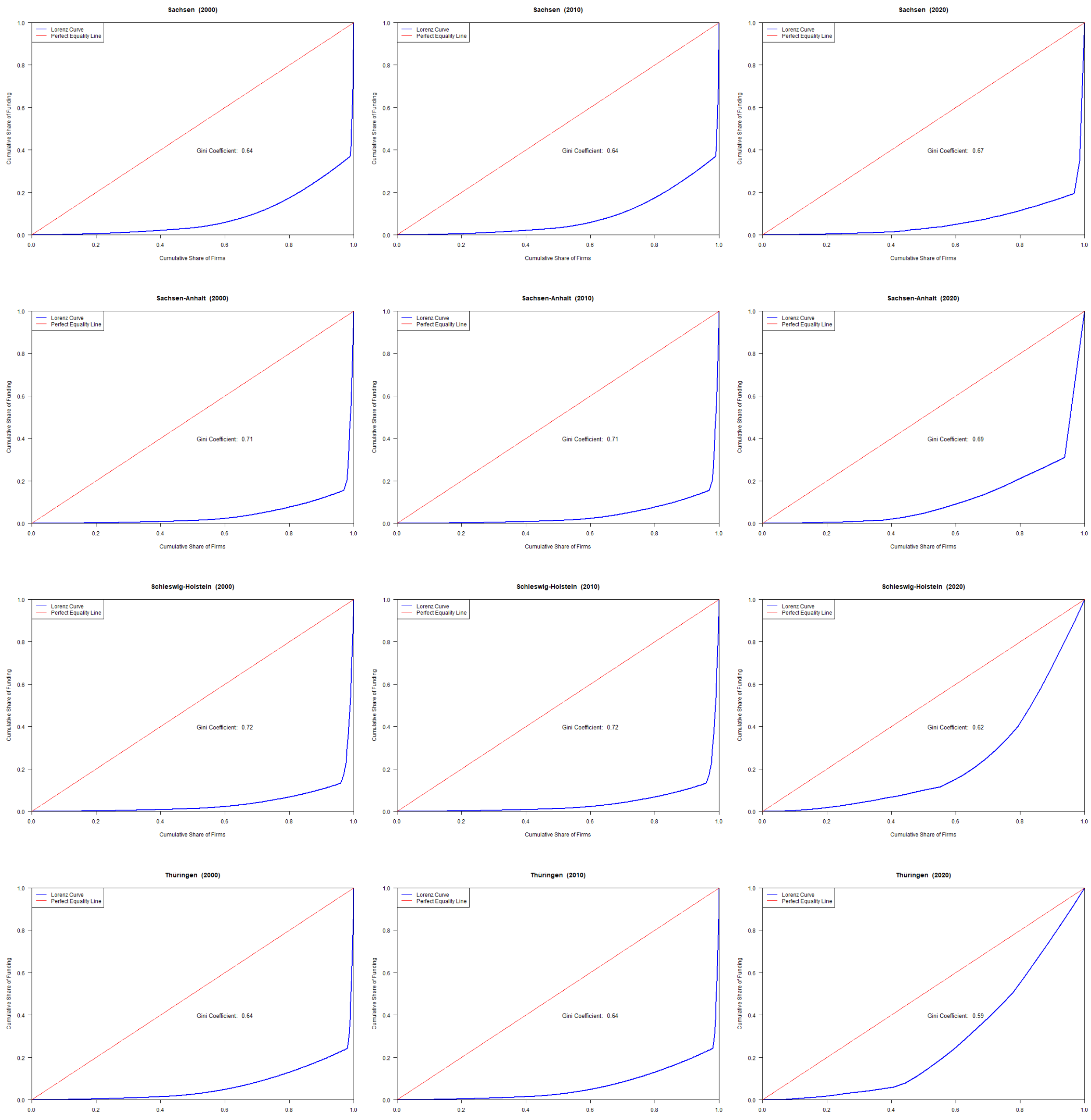


Figure B.4: R&D funding concentration in German states (cont.)



B.1.2 Funding probability

Table B.1: Probability of being funded according to different firm characteristics

Dependent Variable:	Treat										
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Variables											
(Intercept)	-0.0084***	-0.0031***	0.0004	0.0014***	0.0013***	0.0013***	0.0013***	0.0016***	5.08×10^{-5}***	0.0015***	0.0033***
	(0.0003)	(0.0002)	(0.0002)	(3.97×10^{-5})	(3.94×10^{-5})	(3.86×10^{-5})	(3.86×10^{-5})	(4.21×10^{-5})	(7.66×10^{-6})	(4.21×10^{-5})	(0.0008)
Firm size (log(fixed assets))	0.0008***										
	(2.97×10^{-5})										
Firm size (log(number of employees)		0.0030***									
		(0.0001)									
Firm age			0.0004***								
			(7.49×10^{-5})								
log(Patents (simple))				0.0220***							
				(0.0018)							
log(Patents (quality-adjusted))					0.0155***						
					(0.0012)						
Patenting dummy						0.0313***					
						(0.0021)					
MNE							0.0073***				
							(0.0005)				
Basic research								-0.0016***			
								(4.21×10^{-5})			
Applied research									0.1107***		
									(0.0029)		
Secondary industry										0.0017**	
										(0.0008)	
Tertiary industry											-0.0017**
											(0.0008)
Fit statistics											
Observations	738,310	255,597	872,101	877,934	877,934	877,934	877,934	877,934	877,934	877,934	877,934
R^2	0.00185	0.00728	2.8×10^{-5}	0.00457	0.00507	0.00511	0.00136	3.55×10^{-9}	0.10584	1×10^{-5}	1×10^{-5}
Adjusted R^2	0.00185	0.00728	2.69×10^{-5}	0.00457	0.00507	0.00511	0.00136	-1.14×10^{-6}	0.10584	8.9×10^{-6}	8.9×10^{-6}

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

B.2 Robustness Check

B.2.1 Descriptive analysis: Probability of receiving the funding

In this section, I conduct robustness checks to assess the probability of receiving funding through a cross-sectional analysis, with a specific focus on the average year of 2010. I explore this likelihood across different characteristics such as firm size, age, innovativeness, sector, multinational status, and research type, utilizing a logistic regression model.

Table B.2: Logistic regression model for firm size (fixed assets)

	Dependent variable:
	Treat
Firm size (log (fixed assets))	0.417***
	(0.011)
Constant	−12.126***
	(0.166)
Observations	738,310
Log Likelihood	−8,993.045
Akaike Inf. Crit.	17,990.090
Note:	*p<0.1; **p<0.05; ***p<0.01

The results of the logistic regression models for firm size (measured by log(fixed assets) and log(number of employees)) (see Table B.2 and B.3) indicate statistically significant positive associations between firm size and the likelihood of receiving funding. Specifically, for each unit increase in the logarithm of fixed assets, the probability of receiving funding increases by 0.417, and for each unit increase in the logarithm of the number of employees, the probability of receiving funding increases by 0.668. These findings suggest that larger firms are more likely to receive funding.

Table B.3: Logistic regression model for firm size (number of employees)

	Dependent variable:
	Treat
Firm size (log(number of employees)	0.668***
	(0.018)
Constant	−7.940***
	(0.086)
Observations	255,597
Log Likelihood	−3,785.863
Akaike Inf. Crit.	7,575.725
Note:	*p<0.1; **p<0.05; ***p<0.01

Table B.4 further suggests that firm age positively influences access to R&D funding opportunities, indicating that older firms have a higher probability of receiving R&D funding.

Table B.4: Logistic regression model for firm age

	Dependent variable:
	Treat
log (firm age)	0.232***
	(0.047)
Constant	−7.198***
	(0.154)
Observations	872,101
Log Likelihood	−10,182.630
Akaike Inf. Crit.	20,369.260
Note:	*p<0.1; **p<0.05; ***p<0.01

Table B.5 displays the results of the logistic regression model examining the effect of industry type on the likelihood of obtaining R&D funding. The logistic regression model reveals that firms in the secondary industry have a higher likelihood of receiving funding in contrast to the reference group- the tertiary industry.

In Table B.6, I delve deeper into a more detailed categorization of industries by using the NACE 4-digit code. The results suggest that industries such as Education, Information and Communication, Manufacturing, and Professional, Scientific, and Technical Activities have positive coefficients, indicating a higher probability of receiving funding, while other industries show mixed or negligible effects.

Table B.5: Logistic regression model for industry

	Dependent variable:
	Treat
Secondary industry	0.754***
	(0.260)
Constant	−6.468***
	(0.027)
Observations	877,934
Log Likelihood	−10,206.650
Akaike Inf. Crit.	20,417.300
Note:	*p<0.1; **p<0.05; ***p<0.01

Table B.6: Logistic regression model for industry (NACE1)

	Dependent variable:
	Treat
Activities of households as employers	−0.00000
	(3,187.030)
Administrative and Support Service Activities	13.509
	(127.749)
Arts, Entertainment, and Recreation	11.200
	(127.752)
Construction of Buildings	11.788
	(127.749)
Education	15.338
	(127.749)
Financial and Insurance Activities	12.185
	(127.749)
Human Health and Social Work Activities	14.178
	(127.749)
Information and Communication	15.185
	(127.749)
Manufacturing	15.174
	(127.749)
Other Service Activities	14.526
	(127.749)
Professional, Scientifi,c and Technical Activities	14.100
	(127.749)
Public Administration and Defense: Compulsory Social Security	15.291
	(127.749)
Real Estate Activities	12.184
	(127.749)
Transportation and Storage	13.961
	(127.749)
Water Supply	14.572
	(127.749)
Wholesale and Retail Trade	13.063
	(127.749)
Constant	−20.566
	(127.749)
Observations	876,793
Log Likelihood	−9,570.891
Akaike Inf. Crit.	19,175.780
Note:	*p<0.1; **p<0.05; ***p<0.01

Table B.7 displays the logistic regression results for patenting, indicating a significant positive relationship between the presence of patents (measured by a patenting dummy) and the likelihood of receiving funding. Moreover, funding probability is also positively associated with the simple and quality-adjusted number of patents (see Table B.7). These results suggest that firms with an earlier patenting history and a higher number of patents are more likely to get the funding.

Table B.7: Logistic regression model for patent outcome

	Dependent variable:		
	Treat		
	(1)	(2)	(3)
log(Patents (simple))	1.227***		
	(0.040)		
log(Patents (quality-adjusted))		0.994***	
		(0.027)	
Patenting dummy			3.253***
			(0.073)
Constant	−6.524***	−6.548***	−6.642***
	(0.028)	(0.028)	(0.030)
Observations	877,934	877,934	877,934
Log Likelihood	−9,952.465	−9,879.364	−9,698.093
Akaike Inf. Crit.	19,908.930	19,762.730	19,400.190
Note:	*p<0.1; **p<0.05; ***p<0.01		

I further examine the relationships between research type and funding likelihood (see Table B.8). While basic research does not show a statistically significant effect, applied research exhibits a significant positive relationship with funding implying that firms engaged in applied research are more likely to obtain the funding.

Table B.8: Logistic regression model for research type

	Dependent variable:	
	Treat	
	(1)	(2)
Basic research	−4.103	
	(84.477)	
Applied research		7.805***
		(0.154)
Constant	−6.463***	−9.887***
	(0.027)	(0.151)
Observations	877,934	877,934
Log Likelihood	−10,209.990	−4,639.728
Akaike Inf. Crit.	20,423.970	9,283.455
Note:	*p<0.1; **p<0.05; ***p<0.01	

Finally, Table B.9 indicates a statistically significant positive relationship between being a multinational firm (MNE) and the likelihood of receiving funding suggesting that MNEs are more likely to secure funding compared to non-multinational firms.

Table B.9: Logistic regression model for being a multinational

	Dependent variable:
	Treat
MNE	1.923^{***}
	(0.065)
Constant	-6.679^{***}
	(0.031)
Observations	877,934
Log Likelihood	−9,907.244
Akaike Inf. Crit.	19,818.490
Note:	*p<0.1; **p<0.05; ***p<0.01

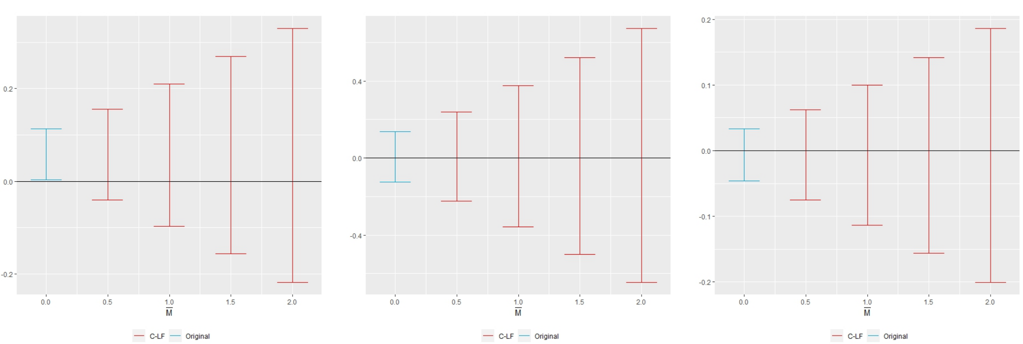
To sum up, the results from the logistic regression model are consistent with those from the OLS regressions (see Section 2.3), indicating a coherent pattern across different analytical approaches.

B.2.2 Honest DID: Relative magnitudes restriction

As a robustness check, I implement the relative magnitude restriction of the Honest Difference-in-Differences (HDD) approach proposed by Rambachan and Roth (2023). This constraint within the Honest DID framework is particularly targeted at the period immediately following the treatment. It guarantees that the confounding factors responsible for non-parallel trends post-treatment do not exceed significantly in magnitude compared to those pre-treatment.

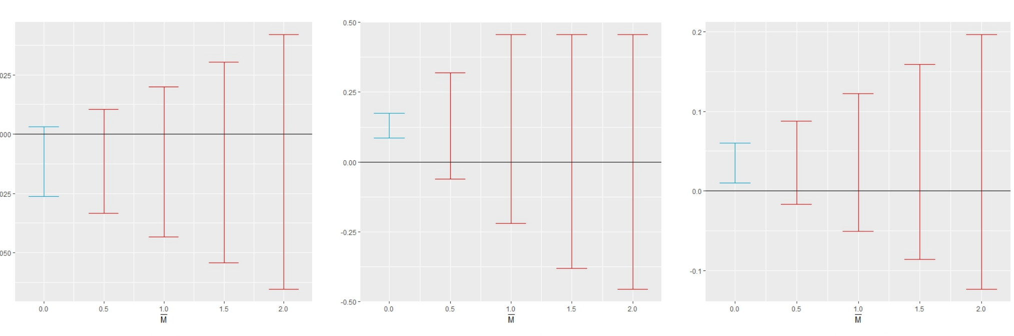
Large Firms: The results depicted in Figure B.6 illustrate the influence of R&D funding on firm performance and innovation outcomes during the initial post-treatment phase, accompanied by robust confidence intervals assessing trend differences. Referencing Table B.10 for coefficient details, the examination of large firms reveals a positive impact on the quality-adjusted number of patents according to the original TWFE estimate, with confidence intervals indicating effects distinct from zero. However, for fixed assets and employment costs, no significant impacts are discerned, as evidenced by confidence intervals encompassing null effects. Nonetheless, imposing the restriction $\overline{M} = 1$, wherein post-treatment violations of parallel trends are capped at the maximal pre-treatment violation, results in immediate significance diminishment (with a "breakdown value" of 0.5), indicating a significant pre-trend influencing the outcome.

Figure B.6: Sensitivity analysis using relative magnitudes restriction (large firms)



SMEs: The impact of R&D funding on SMEs is depicted in Figure B.8, focusing on outcomes one-year post-treatment.

Figure B.8: Sensitivity analysis using relative magnitudes restriction (SMEs)



The results indicate insignificant effects on innovation outcomes while revealing positive and significant impacts on fixed assets and employment costs (as indicated by the blue line). However, when allowing for deviations from parallel trends ($\overline{M} = 1$), the significant effect diminishes, with a critical value of 0.5 (as shown by the red FLCIs). Further coefficient details are available in Table B.11.

To sum up, similar to the sensitivity analysis with the smoothness restriction (see Section 2.5), imposing the relative magnitudes restriction also diminishes significant effects. This suggests the presence of pre-existing trends that continue to influence the outcomes.

Table B.10: Sensitivity analysis using relative magnitudes restriction (large firms)

lb	ub	method	Delta	Mbar	Model
-0.0406838	0.1554900	C-LF	DeltaRM	0.5	Patents (adjusted)
-0.0975296	0.2101066	C-LF	DeltaRM	1.0	Patents (adjusted)
-0.1566046	0.2691816	C-LF	DeltaRM	1.5	Patents (adjusted)
-0.2179089	0.3293713	C-LF	DeltaRM	2.0	Patents (adjusted)
-0.2237740	0.2398536	C-LF	DeltaRM	0.5	Fixed assets
-0.3577704	0.3765299	C-LF	DeltaRM	1.0	Fixed assets
-0.4998066	0.5212461	C-LF	DeltaRM	1.5	Fixed assets
-0.6445228	0.6713221	C-LF	DeltaRM	2.0	Fixed assets
-0.0748208	0.0618788	C-LF	DeltaRM	0.5	Employment costs
-0.1136467	0.0998958	C-LF	DeltaRM	1.0	Employment costs
-0.1557081	0.1419572	C-LF	DeltaRM	1.5	Employment costs
-0.2001961	0.1856364	C-LF	DeltaRM	2.0	Employment costs

Table B.11: Sensitivity analysis using relative magnitudes restriction (SMEs)

lb	ub	method	Delta	Mbar	Model
-0.0334969	0.0103645	C-LF	DeltaRM	0.5	Patents (adjusted)
-0.0434108	0.0199780	C-LF	DeltaRM	1.0	Patents (adjusted)
-0.0542260	0.0304927	C-LF	DeltaRM	1.5	Patents (adjusted)
-0.0653415	0.0419087	C-LF	DeltaRM	2.0	Patents (adjusted)
-0.0615765	0.3197416	C-LF	DeltaRM	0.5	Fixed assets
-0.2203070	0.4556660	C-LF	DeltaRM	1.0	Fixed assets
-0.3826864	0.4556660	C-LF	DeltaRM	1.5	Fixed assets
-0.4556660	0.4556660	C-LF	DeltaRM	2.0	Fixed assets
-0.0166641	0.0879354	C-LF	DeltaRM	0.5	Employment costs
-0.0505052	0.1222892	C-LF	DeltaRM	1.0	Employment costs
-0.0863972	0.1586939	C-LF	DeltaRM	1.5	Employment costs
-0.1233147	0.1961242	C-LF	DeltaRM	2.0	Employment costs

B.2.3 Parametric event study

B.2.3.1 Step 1: Adding non-linear time trend

To account for potential non-linear changes in the treatment effect over time, I incorporate the squared term (Treat × Time to treatment)^2 into the event study regression. This addition allows for a more flexible model, better suited to capture intricate variations in the dynamics of the treatment effect as time progresses before and after funding.

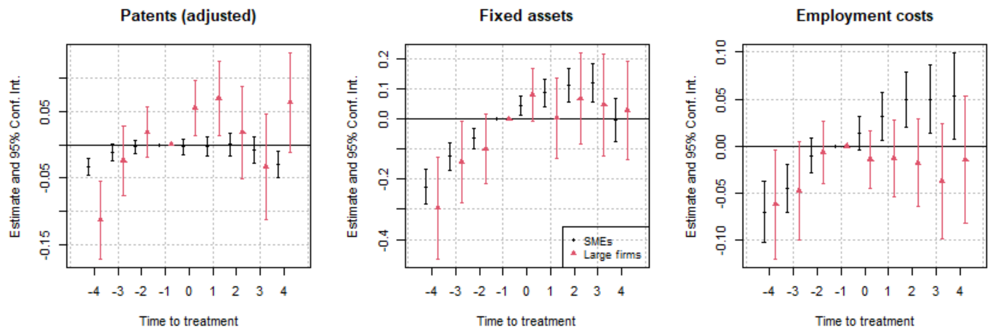
The findings from Table B.12 indicate a significant effect of the policy on the fixed assets and employment costs of SMEs, with no discernible impact on large firms. However, compared to the initial model using the (Treat × Time to treatment) interaction, the coefficients for both treatment periods and interaction terms are smaller in this analysis. Additionally, the significant negative effects of the squared term (Treat × Time to treatment)^2 imply a non-linear association between R&D funding and outcomes over time. Refer to Figure B.9 for a visual representation of these findings.

Table B.12: The direct impact of R&D funding on firm-level & innovation outcomes

	SMEs			Large firms		
Dependent Variables:	Patent (quality-adjusted)	Fixed assets	Employment costs	Patent (quality-adjusted)	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Treat × Relative period = -4	-0.0330***	-0.2258***	-0.0698***	-0.1130***	-0.2963***	-0.0618***
	(0.0056)	(0.0224)	(0.0124)	(0.0233)	(0.0659)	(0.0214)
Treat × Relative period = -3	-0.0114*	-0.1234***	-0.0453***	-0.0237	-0.1414*	-0.0475**
	(0.0063)	(0.0247)	(0.0130)	(0.0259)	(0.0725)	(0.0237)
Treat × Relative period = -2	-0.0031	-0.0642**	-0.0100	0.0185	-0.0998	-0.0066
	(0.0064)	(0.0253)	(0.0123)	(0.0257)	(0.0772)	(0.0217)
Treat × Relative period = 0	-0.0034	0.0435*	0.0142	0.0548**	0.0804	-0.0139
	(0.0070)	(0.0250)	(0.0134)	(0.0274)	(0.0723)	(0.0251)
Treat × Relative period = 1	-0.0023	0.0881***	0.0315**	0.0691**	0.0038	-0.0129
	(0.0076)	(0.0256)	(0.0147)	(0.0295)	(0.0756)	(0.0265)
Treat × Relative period = 2	0.0006	0.1125***	0.0498***	0.0181	0.0674	-0.0179
	(0.0081)	(0.0263)	(0.0140)	(0.0328)	(0.0753)	(0.0255)
Treat × Relative period = 3	-0.0080	0.1209***	0.0498***	-0.0333	0.0464	-0.0371
	(0.0086)	(0.0281)	(0.0159)	(0.0374)	(0.0786)	(0.0293)
Treat × Relative period = 4	-0.0301***	-0.0021	0.0527***	0.0629**	0.0286	-0.0145
	(0.0074)	(0.0250)	(0.0146)	(0.0284)	(0.0672)	(0.0237)
Treat × Time to treatment	-0.0075***	0.0358***	0.0025*	-0.0137***	0.0004	0.0037**
	(0.0004)	(0.0017)	(0.0013)	(0.0016)	(0.0034)	(0.0016)
(Treat ×Time to treatment)^2	-0.0003***	-0.0010***	-0.0003***	-0.0006***	-0.0005***	-0.0003***
	(1.76×10^{-5})	(4.72×10^{-5})	(3.42×10^{-5})	(4.92×10^{-5})	(7.19×10^{-5})	(3.54×10^{-5})
Fixed-effects						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	428,073	385,055	108,104	87,822	72,964	64,366
R^2	0.59030	0.81689	0.90473	0.73030	0.81247	0.81502
Within R^2	0.00585	0.00548	0.00493	0.00841	0.00238	0.00334

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Figure B.9: Direct impact of R&D funding on firm-level and innovation Outcomes



B.2.3.2 Step 2: Pre-period analysis

The results in Table B.13 illustrate the impact of R&D funding on firm-level outcomes during the pre-treatment period. This finding suggests that during the pre-treatment period, there is evidence of a positive relationship between R&D funding and firm-level outcomes, such as quality-adjusted patents, fixed assets, and employment costs for both SMEs and large firms. Moreover, the presence of significant positive coefficients for both the interaction term (Treat $\times$ Time to treatment) and its squared term (Treat $\times$ Time to treatment)^2 implies that the impact of funding on these outcomes may vary non-linearly over time.

Table B.13: Parametric event study in pre-treatment period

	SMEs			Large firms		
Dependent Variables:	Patent (quality-adjusted)	Fixed assets	Employment costs	Patent (quality-adjusted)	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Treat $\times$ Time to treatment	0.0018*	0.1318***	0.0313***	0.0235***	0.0920***	0.0249***
	(0.0011)	(0.0049)	(0.0032)	(0.0043)	(0.0126)	(0.0045)
(Treat $\times$Time to treatment)^2	3.67×10^{-5}	0.0037***	0.0011***	0.0009***	0.0036***	0.0006**
	(5.79×10^{-5})	(0.0003)	(0.0002)	(0.0002)	(0.0006)	(0.0003)
Fixed-effects						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	353,807	322,198	81,281	66,692	56,505	48,767
R^2	0.60605	0.81357	0.91304	0.72172	0.80395	0.78955
Within R^2	6.63×10^{-5}	0.00514	0.00326	0.00086	0.00160	0.00231

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

B.2.3.3 Step 3: Detrending

The results in Table B.14 reveal significant negative effects of the squared interaction term (Treat × Time to treatment)^2 on the outcomes for both SMEs and large firms even after detrending. This suggests that the relationship between R&D funding and firm-level outcomes exhibits a non-linear pattern overtime before the treatment period. Specifically, as the time to treatment progresses, the effect of R&D funding on quality-adjusted patents, fixed assets, and employment costs diminishes at an accelerating rate, as indicated by the negative coefficients of the squared interaction term.

Table B.14: Parametric event study: Detrending

	SMEs			Large firms		
Dependent Variables:	Patent (quality-adjusted)	Fixed assets	Employment costs	Patent (quality-adjusted)	Fixed assets	Employment costs
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Treat × Relative period = -4	-0.0330***	-0.2258***	-0.0698***	-0.1130***	-0.2963***	-0.0618***
	(0.0056)	(0.0224)	(0.0124)	(0.0233)	(0.0659)	(0.0214)
Treat × Relative period = -3	-0.0114*	-0.1234***	-0.0453***	-0.0237	-0.1414*	-0.0475**
	(0.0063)	(0.0247)	(0.0130)	(0.0259)	(0.0725)	(0.0237)
Treat × Relative period = -2	-0.0031	-0.0642**	-0.0100	0.0185	-0.0998	-0.0066
	(0.0064)	(0.0253)	(0.0123)	(0.0257)	(0.0772)	(0.0217)
Treat × Relative period = 0	-0.0034	0.0435*	0.0142	0.0548**	0.0804	-0.0139
	(0.0070)	(0.0250)	(0.0134)	(0.0274)	(0.0723)	(0.0251)
Treat × Relative period = 1	-0.0023	0.0881***	0.0315**	0.0691**	0.0038	-0.0129
	(0.0076)	(0.0256)	(0.0147)	(0.0295)	(0.0756)	(0.0265)
Treat × Relative period = 2	0.0006	0.1125***	0.0498***	0.0181	0.0674	-0.0179
	(0.0081)	(0.0263)	(0.0140)	(0.0328)	(0.0753)	(0.0255)
Treat × Relative period = 3	-0.0080	0.1209***	0.0498***	-0.0333	0.0464	-0.0371
	(0.0086)	(0.0281)	(0.0159)	(0.0374)	(0.0786)	(0.0293)
Treat × Relative period = 4	-0.0301***	-0.0021	0.0527***	0.0629**	0.0286	-0.0145
	(0.0074)	(0.0250)	(0.0146)	(0.0284)	(0.0672)	(0.0237)
Treat × Time to treatment	-0.0075***	0.0358***	0.0025*	-0.0137***	0.0004	0.0037**
	(0.0004)	(0.0017)	(0.0013)	(0.0016)	(0.0034)	(0.0016)
blue!10 (Treat × Time to treatment)^2	-0.0004***	-0.0047***	-0.0014***	-0.0015***	-0.0040***	-0.0009***
	(1.76×10^{-5})	(4.72×10^{-5})	(3.42×10^{-5})	(4.92×10^{-5})	(7.19×10^{-5})	(3.54×10^{-5})
Fixed-effects						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	428,073	385,055	108,104	87,822	72,964	64,366
R^2	0.58967	0.81388	0.90299	0.72585	0.80804	0.81090
Within R^2	0.00662	0.02905	0.04598	0.03276	0.04631	0.01975

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

This robustness check provides additional evidence supporting the presence of pre-existing trends in R&D funding. By examining the squared interaction term in the event study regression, I find consistent evidence indicating that the effect of funding on firm-level outcomes diminishes more rapidly as I move further away from the treatment period. This suggests a non-linear relationship between the timing of the treatment and its impact on firm outcomes, reaffirming the importance of accounting for pre-trends in the analysis.

APPENDIX C

Chapter 3

C.1 Cluster Definition

Table C.1: Geographical classification used in this paper

	NUTS1	Region	Share of funded firms
1	DE1	Baden-Württemberg	0.0208890
2	DE2	Bayern	0.0138296
3	DE3	Berlin	0.0170714
4	DE4	Brandenburg	0.0175150
5	DE5	Bremen	0.0233157
6	DE6	Hamburg	0.0107204
7	DE7	Hessen	0.0127690
8	DE8	Mecklenburg-Vorpommern	0.0203811
9	DE9	Niedersachsen	0.0148482
10	DEA	Nordrhein-Westfalen	0.0148137
11	DEB	Rheinland-Pfalz	0.0132487
12	DEC	Saarland	0.0187119
13	DED	Sachsen	0.0325534
14	DEE	Sachsen-Anhalt	0.0232295
15	DEF	Schleswig-Holstein	0.0106457
16	DEG	Thüringen	0.0302741

Table C.2: Industrial classification used in this paper

	Industry	Share of funded firms
1	Manufacturing	0.0163587
2	Manufacturing	0.0663014
3	Manufacturing	0.0419899
4	Construction	0.0072075
5	Wholesale trade and retail	0.0093676
6	Transport and communication	0.0101057
7	Real estate activities	0.0129624
8	Education	0.0196599
9	Other service activities	0.0218824

C.2 Generating Inverse Probability Weights

On the firm level, we perform logistic regressions for each cluster separately. For each cluster, we use the same set of covariates with pre-treatment values. The set of covariates includes firm age, number of employees, and fixed assets. We regress the treatment dummy d_{ict} on the set of covariates to generate the predicted treatment values $\hat{d_{ict}}$. We obtain the inverse probabilities by calculating the $\frac{1}{\hat{d_{ict}}}$ for treated firms and $\frac{1}{1-\hat{d}_{ict}}$ for non-treated firms. We use these weights afterward in the inverse probability-weighted OLS regressions (Equation 3.3 on the firm-level for each cluster separately.

C.3 Generalized Propensity Score Matching

For the methodology on Generalized Propensity Score Matching, we follow Fong et al. (2018). We start by averaging the firm-level variables across each cluster and defining the treatment variable as the proportion of treated firms within the cluster. This proportion is the average of the treatment dummy on the firm level. The distribution of the treatment variable is skewed to the left, many clusters do not receive any funding or only a few firms receive funding. Therefore, we use the Box-Cox transformation in equation C.1 to transform the distribution of the treatment variable closer to normal. We find the most suitable transformation by computing the distribution correlation between the Box-Cox transformed treatment and the normal distribution. We find the correlation maximizing λ searching on a grid from 150 to -150, where the best λ is at -49.

$$\frac{(p_r + 1)^{\lambda} - 1}{\lambda} \tag{C.1}$$

We then fit four different models, a Maximum Likelihood Estimator (MLE), Generalized Boosted Regression Modelling (GBM), the regular Covariate Balancing Generalized Propensity Score (CBPS), and the non-parametric version of the Covariate Balancing Generalized Proensity Score (npCBPS). Figure C.1 shows the absolute Pearson correlation between each covariate and the transformed treatment variable. The Maximum Likelihood Estimator improves covariate balancing, while GBM, CBPS and the non-parametric CBPS perform better. The CBPS esitmates have the smallest absolute correlation of the covariates. GBM and npCBPS perform slightly worse and come with higher computational effort. Therefore, we use CBPS to estimate Generalized Propensity Scores on the cluster-level. CBPS iteratively optimizes covariate balancing and the treatment prediction resulting in a good model fit and covariate balance while retaining computational feasibility.

Additionally, we perform the generalized matching procedure on the cluster level with different sets of covariates. To reflect the covariates for the estimations with fixed assets and number of employees as outcomes, we exclude the fixed assets or number of employee variables, respectively, from the generalized propensity score estimation and calculate covariate balancing statistics for the different estimators. Figure C.2 shows that for all sets of covariates, the CBPS estimator minimizes covariate correlation. Therefore, we use CBPS throughout our analysis for all cluster-level generalized propensity score estimations.

Figure C.1: Covariate balancing propensity score estimation

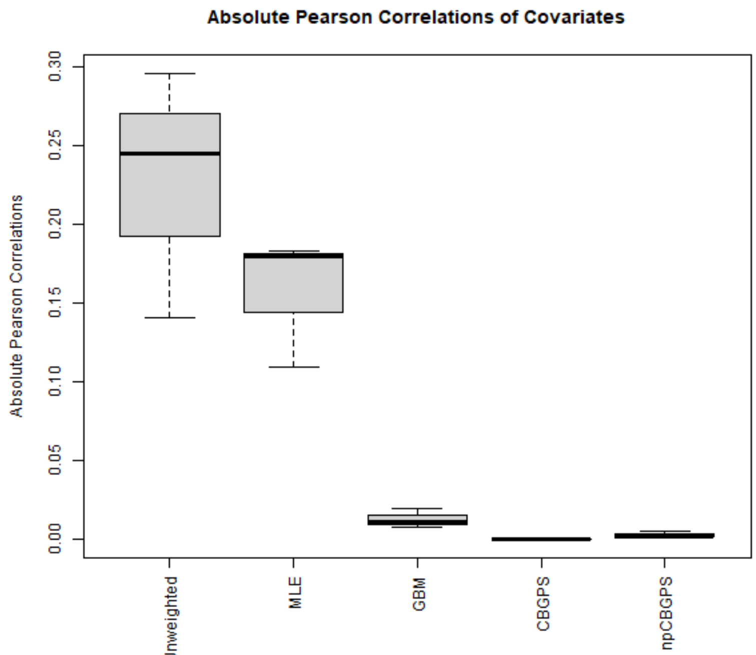
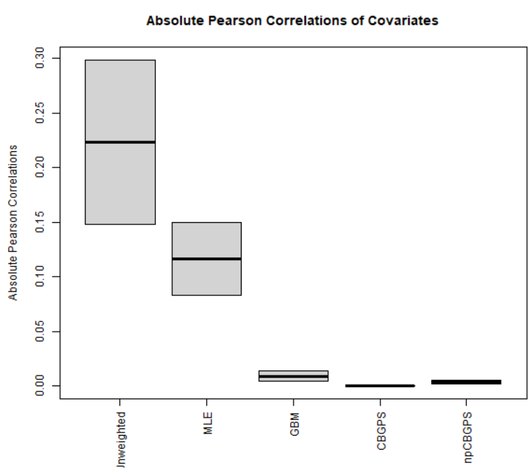
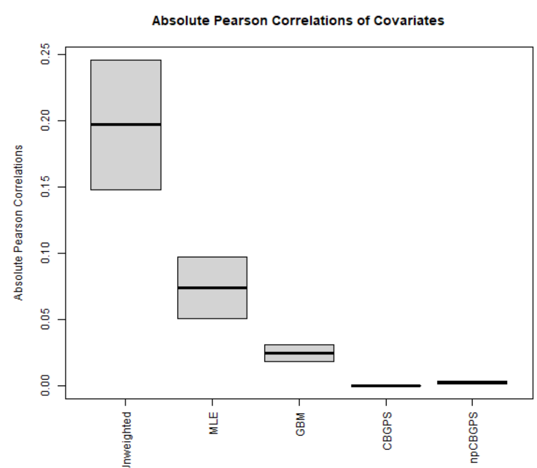


Figure C.2: Covariate balancing statistics: Pearson correlation of covariates



(a) Covariate balancing statistics, excluding fixed assets



(b) Covariate balancing statistics, excluding number of employees

C.4 Bootstrapped Standard Errors

As our analysis includes calculating potential outcomes and involves a multi-step methodology, we bootstrap the standard errors for the confidence intervals shown throughout the paper to assess the variability or uncertainty in our treatment effect estimates. We randomly select samples of firms from our dataset and calculate treatment effects for each sample. We do this process 180 times, creating 180 bootstrap replications. For each bootstrap replication, we calculate different types of treatment effects, such as the direct effect $\bar{\gamma}_{pp}^{10}$, the indirect effect $\bar{\gamma}_{p0}^{00}$, and the overall effect $\bar{\gamma}_{p0}^{10}$. These effects are estimated for various levels of the proportion of treated firms. We follow Girma et al. (2014) for calculating the bootstrap standard errors according to equation (C.2). We simplify the notation by using $\bar{\gamma}$ as a placeholder for the different types of treatment effects $\bar{\gamma}_{pp}^{10}$, $\bar{\gamma}_{p0}^{00}$ and $\bar{\gamma}_{p0}^{10}$. $\overline{\bar{\gamma}_B}$ is the mean of the obtained treatment effects across the bootstrap replications, while $\bar{\gamma}_B$ is the treatment effect obtained from bootstrap replication B.

$$se(\bar{\gamma}) = \sqrt{\frac{1}{180}\sum_{B=1}^{180}(\bar{\gamma}_B - \overline{\bar{\gamma}_B})^2} \tag{C.2}$$