Firms and Innovation

Multinational Strategies, Green Technologies and Government R&D Support

Cristina Rujan

Dissertation
Munich 2024

Max Planck Institute for Innovation and Competition
LMU Munich, Department of Economics

Firms and Innovation

Multinational Strategies, Green Technologies and Government R&D Support

Inaugural-Dissertation
zur Erlangung des Grades

Doctor oeconomiae publicae
(Dr. oec. publ.)

an der
Ludwig–Maximilians–Universität München

2024

vorgelegt von
Cristina Rujan

Referent: Prof. Dietmar Harhoff, Ph.D.
Korreferentin: Prof. Dr. Dr. h.c. Monika Schnitzer
Promotionsabschlussberatung: 10. Juli 2024

Datum der mündlichen Prüfung:	01. Juli 2024
Namen der Berichterstatter:	Prof. Dietmar Harhoff, Ph.D.
	Prof. Dr. Dr. h.c. Monika Schnitzer
	Prof. Claudia Steinwender, Ph.D.

Pentru Kiki

Acknowledgements

There are so many people without whom this dissertation would not have been possible. First, I would like to thank my supervisor, Dietmar Harhoff, for his support and encouraging guidance over the last years. I am particularly grateful for the trust and kindness that he has shown me during the moments when I needed them the most. I am thankful for everything I have learned from him as a supervisor and co-author, and for being inspired by his genuine excitement for the work that he does. I am also grateful to my second supervisor, Monika Schnitzer, for her advice and encouragement over the years, and for including me in her team and providing me with a second academic home, especially during the challenging times of the pandemic. I would like to thank Claudia Steinwender for completing my committee, reading the thesis and providing useful suggestions.

A big thank you goes to all my co-authors for their contributions to our joint projects and for the valuable lessons that they have provided me with over the past years. I would like to highlight Anna Gumpert, who played an instrumental role in my doctoral studies, not only as a co-author, but also as a mentor and role model. It has been a great privilege and pleasure to work with and learn from her. Additionally, I thank Kalina Manova, whose dedication, sustained enthusiasm and openness have been a source of inspiration for me. Her constant words of encouragement have meant so much, especially during the final stages of my dissertation. This work has also benefited from the support of research assistants for which I am particularly grateful. Furthermore, I would like to thank the Research Data and Service Center (RDSC) of the Deutsche Bundesbank and the French Secure Data Access Center (CASD) for providing access to the data that became the foundation of a significant part of this dissertation. At the same time, I am thankful to the MPI administration and IT team, who made it possible for me to access this data. I would also like to acknowledge the kind support of my colleagues Erik Buunk and Mainak Ghosh, who have helped me retrieve parts of the data used in my third chapter.

I am indebted to my colleagues at the MPI, both past and present, for their feedback and contributions that helped me improve and complete this work. I would also like to express my appreciation for the great and friendly atmosphere, for their support and engagement in the activities we have organized together over the years. All of this made a significant difference during an experience that is inherently solitary. Special thanks go to Fabian Gaessler, not only for his generosity in sharing knowledge about patents but also for his availability and general support over the past years. My work has also benefited from conversations with Rainer Widmann and David Heller, to whom I am also deeply thankful. My stay at the MPI has not only taught me how to be a better researcher, but also how to be a better mentor, colleague, communicator, listener and most importantly, a better friend. I am immensely grateful to a few colleagues, now friends, who have provided me with endless moral support and a much-needed safety net in the course of this experience. I would like to thank Zhaoxin Pu for her mentorship, friendship and for being someone I always looked up to. Research stays at the RDSC in Frankfurt would not have been as enjoyable without her. Thanks also go out to Marina Chugunova, whose warmth, deep compassion and mutual love for the arts have given me so much comfort. I am grateful to Elisabeth Hofmeister for her unwavering optimism and confidence in me, and for all the fun times we shared in the last two years. Lastly, I would like to express my deep gratitude to Timm Opitz, with whom I shared this journey as part of the same cohort. He has been there for me consistently through the highs and especially all the lows of this experience. I am grateful for our numerous insightful and stimulating conversations, for his generosity, sense of humor, clarity of mind, and above all, for teaching me how to be braver. I am incredibly proud of the meaningful friendships I am taking away from this journey.

I am fortunate to have received so much love and support from my close friends, who have provided me with plenty of distractions and joyous moments that had nothing to do with my work as a researcher. Sometimes, being an expat added a layer of hardship to the trials of academia. I am profoundly thankful to Federica, Irakli and Laura for being a constant source of encouragement and support, even from a distance. Special thanks go to Ellie, Denis and Joana, not only as flatmates but as close friends navigating life during and soon after a global pandemic. Lastly, I would like to thank my parents. Without their sacrifice, love and support none of this would have been possible.

My academic journey has not only been conducive to professional development, but has also offered numerous opportunities for personal growth and self-discovery. I am particularly grateful to have learnt how much power lies in vulnerability, how it can bring people together and build a sense of community. This is invaluable!

Contents

Preface

Firm innovation is pivotal for technological progress. Economists have long established that technological progress drives not only productivity and economic growth, but also overall social welfare. Without it, current global challenges such as climate change, healthcare, food security and digital transformation would be unsolvable. Investing in research and development (R&D) can enable firms to secure a competitive advantage in such markets. However, the wedge between the private returns to innovation and its social benefits leads firms to underinvest in R&D (Nelson, 1959; Arrow, 1962). Moreover, the profitability of R&D investment may be highest in directions other than those deemed socially desirable. Therefore, in order to tackle the world's current challenges, governments must ensure that the right incentives are in place for firms to pursue innovations that lead to the needed solutions.

The profitability of certain types of R&D is likely to differ depending on firm size (Clancy & Dyèvre, 2023).[1] There is evidence that large firms are more likely to focus on incremental, efficiency-inducing innovation (e.g., Cohen & Klepper, 1996; Akcigit & Kerr, 2018). Their behavior may be driven in this direction by at least two different incentives. First, there is a cost spreading rationale that may encourage larger firms to focus on making their production processes more efficient: the cost of innovation can therefore be spread over their larger output, increasing the returns to R&D (Cohen & Klepper, 1996). Second, large, incumbent firms are less likely to invest in R&D that would result in the replacement of their existing products (Arrow, 1962). In contrast, smaller firms, particularly new ventures, have the potential to generate radical innovation, often challenging existing business models and disrupting markets (e.g., Henderson, 1993; Baumol, 2005; Schneider & Veugelers, 2010). Firms' different incentives as well as their reaction to market opportunities and challenges are relevant for governments aiming to promote innovation as a means to address pressing societal issues.

This dissertation contributes to the understanding of how firms innovate in light of the challenges and opportunities they face. The first chapter delves into the complexity of innovation management within multinational companies (MNCs) who operate in a globalized economy. Traditionally, more attention has been paid to MNCs' role in establishing global value chains through the fragmentation of their production. Yet, the increasing internationalization of R&D documented over the last decades (e.g., OECD, 2008; Hall, 2011) highlights that MNCs have

[1] See Cohen (2010) for a comprehensive overview of the literature on innovation and firm size.

been using the global nature of their operations as an opportunity for optimizing their innovation strategy. This chapter provides insights into how MNCs organize their innovation activity along their global value chains.

The second chapter looks at how large R&D investing firms react to the major societal challenge of climate change. Specifically, it sheds light on firms' contribution to climate change mitigation through their innovation endeavors. Technology development is instrumental in curbing climate change, as the current set of available technologies is insufficient for reaching the ambitious climate-neutrality goal established for mid-century through the adoption of the Paris Agreement (IRENA, 2017). Although the development of climate-related technologies can be seen as a major economic opportunity for companies, without the right incentives in place, their contribution may fall short of what is socially optimal. There are several reasons for this that go beyond the fundamental market failure that characterizes the production of innovation.[2] For example, firms do not internalize the cost of pollution they generate, which dampens their incentive to invest in green innovation. Smaller firms seeking to develop novel technological solutions may face difficulties in securing financing, given the higher risk that is often associated with green technologies (Vysoká et al., 2021). Against this background, the chapter contributes by analyzing how firms with the highest capacity to allocate resources to innovative projects currently react to the global climate crisis. It uncovers recent trends in green technology development by the world's leading R&D investing firms.

The third chapter studies how government R&D support can help young innovative companies overcome financing challenges. In the early stages of their development, such firms face significant challenges in securing access to external financing. This is the case because potential investors cannot easily assess the commercial viability of entrepreneurs' ideas or the likelihood of success of their inventive activities (Hall & Lerner, 2010). Such financial constraints may impede firms from seizing market opportunities, hinder their growth rate and alter their R&D trajectories in the form of delayed or abandoned projects. Aware of the contribution of new ventures in generating radical innovation, shaping new markets and driving technological change, governments around the world have been using a variety of policy tools aiming to support them. This chapter examines how an unconventional government R&D support scheme introducing a reduction in R&D labor cost influences startups' performance and innovation output.

All chapters are empirical in nature, exploiting complex datasets that combine numerous sources of firm-level data with detailed patent information. In the last decades, economists have studied various aspects related to the innovation behavior of firms focusing on their patenting activity.[3]. Patents are an outcome of firms' innovation process that can be systematically measured. Not all innovation efforts result in a patent application, implying that patents

[2] See OECD (2011) for a discussion on market failures that are specific to the market for green innovation.

[3] See e.g., the early work of Griliches (1990), Jaffe, Trajtenberg, and Henderson (1993), and Hall, Jaffe, and Trajtenberg (2001).

only partially reflect firms' R&D activities. However, they incorporate a wealth of information regarding an invention's inherent value, the technological domain it belongs to, as well as the sources of knowledge it builds upon. This renders them a valuable proxy for firm innovation. To summarize, this dissertation studies various facets of firm innovation behavior as reflected in their patenting activity.

Chapter 1 examines how multinational companies organize their global production and innovation. MNCs have been driving two major globalization trends: the fragmentation of production chains and internationalization of R&D. Through their global operations, MNCs are instrumental in enabling cross-border technological transfer, which plays an important role in the global adoption of technological advancements. They can therefore make a significant contribution to addressing current global challenges.

While prior literature has paid particular attention to the global organization of multinational production (e.g., Helpman, Melitz, & Yeaple, 2004; Yeaple, 2013), the rise in the internationalization of R&D highlights that MNCs have been increasingly relying on global R&D networks. This contradicts the traditional belief that skill-intensive activities would be retained in MNCs' home country and low-skill intensive manufacturing would be offshored to less developed countries. We establish novel stylized facts on MNCs' global production and innovation activities using uniquely rich data on the network of foreign affiliates and patenting activity of German multinationals. These stylized facts are rationalized with a theoretical model in which companies jointly choose the location, scale and organization of production and innovation activities, distinguishing between basic and applied R&D. We derive key theoretical predictions that we then test empirically.

For each German MNC, we use the location of its foreign affiliates and patent inventors in order to identify whether innovation is conducted domestically, offshored to a country with an affiliate, or outsourced to a third country. We operationalize the distinction between basic and applied R&D by assessing the proximity of firms' patents to fundamental science, as evidenced by their references to scientific articles (Ahmadpoor & Jones, 2017). We define science-based patents as closely linked to fundamental science, stemming from basic R&D efforts. Non-science-based patents are those more distant from science, considered an outcome of applied R&D.

Notably, MNCs innovate actively and frequently abroad, offshoring innovation to locations both with and without foreign affiliates. Empirical evidence consistent with the theoretical model indicates that larger MNCs innovate more intensively in terms of both the number and quality of patents. In addition, such companies offshore innovation to more countries, both with and without foreign affiliates. We find that MNCs organize their global innovation activities following countries' comparative advantage in certain technology areas, with applied innovation being more likely to be co-located with production compared to basic innovation.

We advance the literature on the innovation activity of multinational firms by considering MNC global production and global innovation strategies in an integrated framework, examining offshore innovation decisions across countries with and without foreign affiliates. We leverage insights from recent innovation literature that highlights the importance of science for private-sector innovation (e.g., Ahmadpoor & Jones, 2017; Krieger, Schnitzer, & Watzinger, 2023) and uncover differences between basic and applied innovation, studying their cost synergies with production. Our findings provide insights that can inform the design of trade and innovation policy and the scope for multilateral agreements. Moreover, they suggest that innovation activities across countries are complements, rather than substitutes. This may alleviate concerns about the impact of offshoring innovation on sending economies.

Chapter 2 studies how the world's leading R&D investing firms contribute to climate change mitigation through their innovation activity. Innovation is a fundamental pillar in the race against climate change, facilitating the transition to a low-carbon economy. The current set of available technologies falls short of meeting the ambitious net-zero target by 2050, with a substantial portion of CO_2 emission reductions expected to stem from technologies currently in the demonstration and prototype phase (IRENA, 2017; IEA, 2023). Therefore, it is important to understand whether the urgency for new technological solutions is reflected in the recent innovation efforts of the world's top R&D investors. These firms are able to allocate resources to riskier green innovative projects, that are likely to entail high fixed costs and longer-term investments. At the same time, given that top R&D investors are typically large multinational companies, they can drive the faster adoption of technological advancements through their global operations.

The chapter provides a comprehensive overview of recent trends in green technology development by top R&D investors, focusing on their most internationally relevant green inventions patented during 2012-2019. The analysis sheds light on their contribution to global green innovation, identifies pioneering sectors, countries, and companies, and offers insights into the nature of green and non-green innovation pursued by these firms. In addition, I pay particular attention to dynamics since the adoption of the Paris Agreement in 2015 and investigate whether firms have shifted their R&D efforts towards green inventions in line with growing international commitments to curbing climate change.

I build on and enhance the work of Amoroso et al. (2021) by linking top R&D investors in the EC-JRC-OECD COR&DIP database to PATSTAT. Worldwide top R&D investors are identified based on their absolute R&D expenditures in 2018. The analysis focuses on triadic patents filed at the US, European and Japanese patent offices that are known to capture higher-value inventions (Dernis & Khan, 2004). Green patents are defined using the EPO's *Y02 tagging scheme*, a standardized classification meant to target technologies that mitigate or provide adaptation to climate change (Angelucci, Hurtado-Albir, & Volpe, 2018).

Top R&D investors account for 67% of all high-value green patents filed between 2012 and 2019. Despite global commitments to the net-zero goal following the adoption of the Paris Agreement, firms have shifted their focus away from the development of green technologies in recent years. This is evidenced by a decrease in both the number of green patents and their share relative to all patents. Firms' green patents tend to receive more citations, be more original and rely on more recent sources of knowledge compared to their non-green patents. Moreover, I do not find consistent evidence that the nature of firms' patents has changed since 2015. The only exception is that recent green patents appear to rely on older knowledge sources compared to those filed before 2015. The puzzling result points to a change in their R&D activity that deserves further exploration.

This chapter offers policy-relevant insights into recent developments in green inventions, as evidenced by the global landscape of green patenting. As global policy goals become increasingly ambitious, the documented negative trends raise concerns and draw attention to the need for a deeper understanding of the factors driving firms' innovation decisions in the realm of green technology. This result highlights a significant disparity between the increased perceived importance of climate change mitigation technologies and the decline in green patented inventions by large R&D investors. Ultimately, it suggests that the private sector's innovation mechanism is not operating at its full potential, which goes against widespread perceptions in policy circles and the public at large.

Chapter 3 examines whether government R&D support targeted at young innovative companies can help them become successful. Policymakers are interested in supporting innovative startups given their fundamental role in generating radical innovations, shaping new markets, fostering job creation, thereby contributing to economic growth and technological progress (Henderson, 1993; Schneider & Veugelers, 2010; Haltiwanger, Jarmin, & Miranda, 2013; Haltiwanger et al., 2016). Government intervention is primarily motivated by the financial challenges faced by such firms in their early stages of development. This chapter investigates how reducing R&D labor costs through exemptions from social security payments during startups' early stages could impact their performance.

The context of our analysis is the French *Jeune Entreprise Innovante* (JEI) scheme, an unconventional form of government R&D support that provides innovative startups with exemptions from paying social security contributions for their R&D personnel. We document the uptake of the JEI status since its introduction in 2004 and emphasize its widespread usage, demonstrated by the granting of over €2 billion in total exemptions to young firms. Most importantly, we shed light on the early impact of JEI benefits on firms' survival, likelihood of acquisition, innovation output, and overall performance as evidenced by the evolution of their employment and sales.

We do so by comparing the outcomes of JEI beneficiaries with those of similar young innovative companies that do not benefit from the scheme. We leverage access to complex French administrative data, along with patent and firm acquisition information from external sources. From a pool of 11,000 JEI beneficiaries, we focus on a subset of firms founded between 2006-2012 that entered the scheme within their first four years and follow them over an eight to twelve-year time period. Evaluating firms' eligibility for the JEI status is empirically challenging given the multiple criteria that they must fulfill. However, the richness of our data enables us to confidently identify a set of similar startups eligible for the scheme in their initial stages, but not benefiting from it.

We provide evidence suggesting that JEI firms are more successful compared to similar young innovative companies not benefiting from the scheme. We find that startups receiving JEI benefits in the initial phase of their development are significantly more likely to survive and to be acquired. Additionally, they are not only more likely to file a patent at ages five to eight, but they are also filing more patents. Finally, we show that there are significant differences in the workforce composition between JEI and non-JEI firms. Specifically, JEI firms not only prioritize hiring more R&D-related personnel but also increase recruitment of non-R&D employees. We interpret this finding as evidence of JEIs' maintained research emphasis, complemented by a focus on scaling up their operations.

This chapter analyzes a rare and unconventional R&D support tool that, to the best of our knowledge, has been understudied with respect to diverse aspects of startup performance. Prior work has mostly focused on labor market outcomes (Hallépée & Houlou Garcia, 2012; Gautier & Wolff, 2020; Quantin, Bunel, & Lenoir, 2021). The relevance of the scheme extends beyond its widespread adoption or the substantial exemptions granted to young firms. Compared to a standard R&D grant program allocated on a competitive basis, such a scheme offers simpler implementation, reduced bureaucratic burden, as well as the provision of more long-lasting support. Our analysis offers valuable policy-relevant insights into startups' utilization of the JEI scheme as well as its effectiveness in helping firms overcome financial challenges during their early development stages.

To summarize, this dissertation sheds light on various novel aspects related to firms' innovation behavior and contributes to the understanding of how firms innovate in light of the opportunities and challenges they face. It offers policy-relevant insights into how multinational companies optimize their innovation strategies in an increasingly globalized economy, how prominent R&D investors contribute to combating climate change through their innovation efforts and how governments can support innovative startups by reducing their R&D labor cost.

1

Multinational Firms and Global Innovation

1.1 Introduction

Multinational companies (MNCs) are at the heart of two key globalization trends: the fragmentation of production chains and the internationalization of technological progress. MNCs are focal in global value chains (GVCs), in that they manage complex production networks across multiple countries and offshore manufacturing stages to both foreign affiliates and independent parties. At the same time, MNCs are also responsible for the vast majority of frontier R&D and cross-border technology transfer. These phenomena have first-order implications for the optimal design of trade and innovation policy in developed and developing countries. They also shape the impact of technological leaps such as automation on the global distribution of production, innovation and adoption, and thereby on economic growth across countries.

While a large literature examines the global organization of MNC production, relatively little is known about the global organization of MNC innovation. Traditional priors might suggest that

*This chapter is based on joint work with Anna Gumpert, Kalina Manova and Monika Schnitzer.

MNCs headquartered in rich countries retain skill-intensive activities at home and offshore low-skill intensive manufacturing to poorer countries. Yet anecdotal evidence points to increasing offshoring of innovation, to both advanced and emerging economies. For example, while leading German car producer BMW has for a long time sourced auto components from China, in 2018 it also unveiled a large research and development center in Shanghai to specialize in digital services, autonomous driving, and automotive design.[1] In 2017, Mercedez-Benz opened its sixth R&D lab in Seattle, primed as a digital hub for cloud computing.[2]

This paper provides one of the first analyses of MNCs' global production and innovation. We establish novel stylized facts using uniquely rich data on the network of production affiliates and patenting activity of German multinationals.[3] We rationalize these facts with a heterogeneous-firm model, in which companies jointly determine the location and scale of production, basic (science-based) innovation and applied (non-science-based) innovation, under asymmetric complementarities across these three activities. Empirical evidence consistent with the model indicates that bigger MNCs innovate more intensively in terms of patent frequency and quality. Such companies also offshore a greater share of their innovation to more countries, including both countries with and without production affiliates. Moreover, MNCs pursue innovation across countries and technology fields following countries' comparative advantage, with applied R&D more likely to be co-located with production than basic R&D.

Our first contribution is to uncover new facts about the global innovation activity of multinational companies. We obtain firm-level data on German MNCs and their worldwide network of production affiliates from the Microdatabase Direct investment (MiDi) of the Deutsche Bundesbank, and match it to patent-level data from PATSTAT Global, the database maintained by the European Patent Office. For each parent company, we use the country locations of its subsidiaries and patent inventors to identify innovation conducted at the headquarter country, offshored to a country with an affiliate, or outsourced to a third country. We distinguish between two types of R&D: basic and applied. In practice, we operationalize this distinction through measuring patents' distance to fundamental science, as proxied by backward citations to scientific articles (Ahmadpoor & Jones, 2017). For this purpose, we define science-based patents as those with a closer connection to fundamental science and thus an outcome of basic R&D activities. Non-science-based patents are those more distant from science, considered an outcome of applied R&D. We also quantify patent quality with the number of forward citations by subsequent patent applications.

We establish three stylized facts about MNCs' patent-generating research and development. First, MNCs innovate actively and frequently abroad. Second, MNCs innovate in multiple locations, and offshore innovation to locations both with and without affiliates. Third, larger

[1] BMW Corporate Communications. Press Release, 15.06.2018.

[2] Day, M. (2017, Nov 14). Mercedes-Benz plans up to 150 software engineers at Seattle R&D Office. *The Seattle Times*.

[3] We use "patenting", "innovation" and "R&D" interchangeably throughout the paper for expositional convenience.

MNCs innovate at higher intensity and quality. Germany provides an ideal economic context to study these patterns, as it is the third biggest exporter (as of 2022), top MNC origin, and a world innovation leader.

Our second contribution is to develop a partial-equilibrium three-country model of MNCs' global production and innovation strategy, motivated by the new stylized facts. In the model, heterogeneous firms jointly choose the location and scale of their output production, basic innovation, and applied innovation to maximize total profits. Firms optimally operate a single manufacturing facility due to economies of scale in production. Each type of R&D may or may not be offshored, to one or multiple countries, with or without a production affiliate in the same location. The returns to innovation are additive across countries within an innovation type and multiplicative across innovation types, with basic innovation increasing future expected profits and applied innovation raising profits immediately by lowering marginal production costs. Location-specific innovation costs are higher abroad, and rise with innovation intensity and local inventor wages. Importantly, there are cost synergies between production and applied R&D, but not with respect to basic R&D. This setup generates complementarities in innovation across types and locations.

This model delivers several key predictions. First, more productive multinationals are more likely to innovate and to innovate more intensively. Second, more productive MNCs are more likely to offshore R&D and to undertake R&D in more countries. Given the geographic concentration of production, this also implies that more productive MNCs have a greater propensity to simultaneously pursue innovation both in locations with and without a manufacturing subsidiary. Third, MNCs are more likely to innovate and to innovate more intensively in countries with lower inventor wages. When inventor wages also vary across technology areas, MNCs locate innovation activity according to countries' comparative advantage. Finally, MNCs are more likely to conduct applied innovation in locations where they operate a production affiliate, compared to basic innovation.

Our third contribution is to provide systematic empirical evidence for the operations of German multinationals that is consistent with the model's predictions. In the absence of direct productivity measures and innovation data, we use global firm sales and observed patent outcomes as model-consistent proxies.

We first confirm that bigger MNCs are more likely to file patents. Conditional on patenting activity, bigger MNCs generate more patents, record more total patent citations, and receive more citations per patent on average. Moreover, these patterns hold for each of type of R&D activity, basic and applied.

We then establish that larger firms are more likely to offshore innovation. We proxy offshore innovation using patents that have at least one inventor located abroad. Along the extensive margin, bigger MNCs have a higher probability of innovating in at least one foreign country.

Along the intensive margin, bigger MNCs develop a greater share of their patented technologies abroad. Larger firms also innovate in more countries on average, and are more likely to pursue R&D both in locations with and without a production subsidiary. All of these findings hold both with respect to total innovation activity and separately within each innovation type.

We next document that MNCs respond to cross-country differences in comparative advantage in innovation across technological fields. We construct an indicator of revealed comparative advantage in innovation at the country-technology area level, based on the total number of patents invented in each country and technology area. We observe that even within an MNC, the pattern of inventor location across countries strongly follows comparative advantage, particularly when we consider citation-weighted patent counts.

Lastly, we demonstrate that multinationals are more likely to co-locate applied R&D with production, compared to basic R&D. Non-science based patents are more likely to be offshored to countries where the MNC operates a production affiliate, compared to science-based patents that are closer to fundamental science.

Our paper advances several strands of literature. At a broad level, we add to a large body of work on the drivers of multinational production activity. Extensive theoretical and empirical analysis has found that firm productivity, cross-country differences in production wages, and economies of scale at both the firm and establishment level are key determinants of MNC production patterns (Yeaple, 2003; Helpman, Melitz, & Yeaple, 2004; Yeaple, 2013). We incorporate these ingredients in a generalized model of the joint production and innovation decisions of multinational firms. We purposefully keep manufacturing choices stylized to highlight the novel interdependence of innovation decisions across locations, as well as between manufacturing and innovation. Our framework can, however, be readily enriched to incorporate more complex production strategies across countries that recent contributions explore (Ramondo, Rappoport, & Ruhl, 2016; Tintelnot, 2017).

We also extend a separate literature on the innovation activity of multinational firms. This line of work traditionally examines R&D at the parent headquarters and its deployment across the firm's affiliate network and consumer markets through production technologies and product design. Evidence indicates that intellectual property rights protection matters for multinationals' production and sales decisions (Javorcik, 2004; Branstetter, Fisman, & Foley, 2006; Bilir, 2014). MNC parents nevertheless earn significant returns on their home-grown innovation abroad, with time-zone differences shaping the creation and diffusion of knowledge within the firm (Bilir & Morales, 2020; Bircan, Javorcik, & Pauly, 2021). Moreover, improved opportunities for production offshoring can generate cost savings that incentivize innovation at headquarters (Branstetter et al., 2021; Bernard et al., 2024).

We advance this agenda in three dimensions. First, we consider MNCs' global production and global innovation strategy in an integrated framework. We examine offshore innovation to both locations with and without production affiliates, uncover important distinctions between basic and applied innovation, and study their different cost synergies with production. In closely related work, Liu (2023) focuses specifically on the benefits of co-locating production and generic innovation in a structural dynamic model of MNC investments, to evaluate the impact of re-shoring policies.[4] Second, we exploit novel margins of MNC patent activity, comparing basic and applied innovation. Unlike standard R&D metrics, patent data enables comprehensive coverage, comparisons between basic and applied innovation, and assessment of innovation quality. And third, our work provides empirical context for recent models that quantify the welfare impact of MNC operations under innovation either by the parent or by both parent and affiliate (Arkolakis et al., 2018; Fan, 2021).

To shed light on the determinants of innovation offshoring, we leverage insights from a recent body of innovation literature that highlights the importance of science for private-sector innovation. Ahmadpoor and Jones (2017) show that patents with a shorter distance-to-science receive more forward citations, suggesting that they are more valuable. Furthermore, Krieger, Schnitzer, and Watzinger (2023) find that patents closer to science are more novel and have higher private value. Building upon these findings, we add a new angle to the literature by distinguishing between science-based and non-science-based patents while exploring firms' motives for innovation offshoring.

The rest of the paper is organized as follows. Section 1.2 introduces the data and novel stylized facts about the global patent activity of German multinational firms. Section 1.3 develops an integrated theoretical model of MNC production and innovation that rationalizes these facts and delivers a series of additional testable predictions. Section 1.4 provides systematic empirical evidence consistent with the model. The last section concludes.

1.2 Data and Stylized Facts

We first establish stylized facts about the global production and innovation activity of multinational firms using uniquely rich new data for Germany. This presents an ideal economic environment in which to study multinational activity, as Germany is the third biggest exporter in the world (as of 2022), hosts many globally engaged firms, and is a global innovation leader.

1.2.1 MNC Production and Innovation Data

We combine administrative firm-level data on German multinational firms from the Micro-database Direct investment (MiDi) of the Deutsche Bundesbank (German Central Bank) with rich patent-level data from PATSTAT Global, the worldwide patent statistical database provided

[4] In a domestic context, Fort et al. (2020) study the innovation behavior of US firms over the long run, and link the increasing importance of former manufacturing firms for US innovation to the fragmentation of production.

by the European Patent Office.

MNC production. We characterize the domestic and foreign production operations of German multinational firms with comprehensive data from MiDi. This dataset covers approximately 15,000 German parents and their global network of affiliates in around 200 host countries over 1999-2016.[5] Since German parent firms are legally obliged to report their foreign investments, MiDi provides high-quality information from the accounting statements of each parent and each subsidiary that is comparable across countries (e.g., sales, employment, balance sheet total, industry affiliation).[6] Therefore, MiDi provides a comprehensive global overview of German parents' domestic and foreign operations, which we use as proxy for MNC production. It allows us to not only locate foreign affiliates, but also to distinguish between domestic and foreign sales generated by the MNCs of interest.

Panel A of Table 1.1 summarizes the substantial variation in multinational production activity across firms and over time. On average, a German parent has 4 affiliates in 3 countries, with standard deviations of 10.8 and 4.8 around these means. The typical multinational generates EUR 254 mil worth of sales revenue at its headquarters and EUR 262 mil across its foreign affiliates, with standard deviations of EUR 2,954 mil and EUR 2,311 mil respectively.

MNC innovation. We characterize MNCs' global innovation activity using detailed information from the patents they file. For this purpose, we rely on PATSTAT Global, a database that contains detailed patent bibliographical data on over 100 million patent documents filed around the world. Patents reflect the outcome of complex invention processes, and therefore provide an imperfect, but informative proxy for the underlying innovation effort. Patent data is not only more complete than typically sparse R&D records, but it also contains additional detailed information such as the location of inventors and the nature of the patented invention.

Firms can in principle secure property rights over a single invention or technology in multiple markets by filing a collection of patent applications within multiple jurisdictions. This is typically referred to as a patent family. Following common practice in the literature, our analysis focuses on patent families and not single patent applications, such that we count each invention only once. In the remainder of the paper, we refer to patent families simply as patents. We restrict the baseline analysis to patent families that include an application filed at the European Patent Office, which we label EP patents. This ensures that we compare firms' patenting

[5] MiDi also comprises private and public households. We exclude those in our analysis. More details on data construction can be found in Appendix A.2.1.

[6] See Drees, Schild, and Walter (2018) for a comprehensive description of the dataset. German parent firms are required to report to the Bundesbank if they directly own at least 10% of the shares or voting rights in a foreign company with a balance sheet total above EUR 3 mil. German parent firms must also report indirect investment relationships and relationships that mix direct and indirect ownership if they hold at least 50% controlling share in a foreign company with a balance sheet total above EUR 3 mil. MiDi carefully tracks changes in affiliate ownership over time. We do not focus on ownership turnover as this is a rare event in our baseline sample.

Table 1.1: Summary Statistics

Panel A. MNC Production			
Sample: All German MNC (N = 10,155)			
	Firm-year level		
Variable	N	Mean	St. dev.
Parent sales, mil. €	73,800	254	2,953.8
Affiliate sales, mil. €	84,701	262	2,311.3
# affiliates	84,701	4	10.84
# host countries	84,701	3	4.78
Panel B. MNC Production			
Sample: Innovating German MNC (N = 2,374)			
	Firm-year level		
Variable	N	Mean	St. dev.
Parent sales, mil. €	22,048	551	4,010.4
Affiliate sales, mil. €	25,712	397	3,549.7
# affiliates	25,712	6	14.06
# host countries	25,712	4	6.15
Panel C. MNC Innovation			
Sample: Innovating German MNC (N = 2,374)			
	Firm level		
Variable	N	Mean	St. dev.
# patents	2,374	148.58	1,464.40
# EP patents	2,374	63.70	538.20
# offshore patents	2,374	21.13	258.34
# EP offshore patents	2,374	10.74	122.84
# EP non-science-based patents (applied)	2,374	44.91	410.79
# EP science-based patents (basic)	2,374	17.59	169.18
# citations*	2,073	176.23	1,601.21
Average # citations*	2,073	1.09	1.35
Share science-based patents (EP)	2,030	0.17	0.26
Share offshore patents (EP)	2,030	0.12	0.24
Share offshore co-located (EP)	2,030	0.04	0.14
Share offshore science-based patents (EP)	2,027	0.03	0.11
Share offshore non-science-based patents (EP)	2,020	0.08	0.18

Table 1.1: Summary Statistics (continued)

Panel D. Patent Characteristics			
Sample: EP patents			
	Patent level		
Variable	N	Mean	St. dev.
# citations	121,762	2.12	4.59
# citations, non-science based (applied)	84,957	1.81	3.55
# citations, science-based (basic)	34,504	2.92	6.47
# citations, domestic	102,029	2.00	4.09
# citations, offshore	19,733	2.76	6.58
# citations, offshore co-located	14,502	2.68	6.71
# citations, offshore not co-located	5,231	3.00	6.21
Science-based (%)	28 %		
Offshore (%)	16 %		
Offshore co-located (%)	12 %		

Notes: This table presents summary statistics for the production and innovation activity of German MNCs in 1999-2016. The sample includes all German MNCs in Panel A, all German MNCs that patent at least once in Panels B and C, and all EP patents by German MNCs in Panel D. Firm-level patent counts are aggregated across all years in 1999-2016. 5-year forward citation counts are computed for the restricted 1999-2011 period to account for truncation effects. Patents are classified into science-based and non-science-based based on backward citations to scientific journal articles. *Data sources:* Research Data and Service Center of the Deutsche Bundesbank, MiDi, 1999-2016, combined with PATSTAT, own calculations.

activity within a single jurisdiction - the one most relevant for German firms, and allows us to exploit additional information necessary for determining the patent type as explained below. For completeness, we also provide summary statistics and robustness checks using the full set of patents families. We abstract away from changes in patent ownership, and focus strictly on patents filed by the MNCs of interest.

We build a comprehensive dataset on the global production and innovation operations of German multinationals by merging MiDi and PATSTAT based on unique firm identifiers.[7] This produces a sample of 10,155 German MNCs, roughly 30% of whom file at least one patent within the 1999-2016 time frame. Our baseline sample of innovating MNCs includes 2,374 firms and their 352,720 patent families.[8] Out of these, 151,227 patent families include an EP application filed at the European Patent Office.

[7]We link MiDi to PATSTAT via the Bureau van Dijk's ORBIS database. The Deutsche Bundesbank Research and Data Center has developed a mapping from MiDi parent firms to Orbis firm identifiers (BvD ID) using supervised machine learning. The Bureau van Dijk's Orbis database in turn provides a crosswalk from Orbis to PATSTAT.

[8]The baseline matched sample does not include all 30% of firms that patent at least for several reasons. First, a firm may appear in MiDi and in PATSTAT in different years, i.e. it may patent in a year with no corresponding MiDi entry. Second, we focus on patents with a single MNC owner, and drop jointly-owned patents whose innovation and filing decisions may reflect economic forces beyond the scope of this paper. Appendix A.2 elaborates on all other cleaning steps in building the baseline patent sample.

Panel B of Table 1.1 presents summary statistics for German innovating MNCs, where we have retained all years that an MNC is active in MiDi as long as it appears at least once in PATSTAT. Innovating German MNCs are on average not only larger in terms of domestic and affiliate sales, but are also present in more host countries and have more foreign affiliates relative to the average MNC described in Panel A.

We use information on the location of the patent inventors to identify where the underlying innovation activity took place. For patents with multiple inventor countries, we assign equal fractions to each one. We define offshore patents as those with at least one inventor located outside of Germany.

While German MNCs' innovation activity is highly concentrated in major innovation hubs, MNCs undertake significant patent-generating R&D across the globe. Figure 1.1A plots the average yearly number of MNCs offshoring innovation to a given foreign country against the country's the GDP per capita. The United States stands out as the top location for offshore innovation by German multinationals. France, Austria and Switzerland, technologically advanced and proximate, are also favored hosts. At the same time, numerous other countries across the income distribution attract non-trivial offshore innovation activity.

A similar pattern emerges when looking at the total number of patents originating from an inventor country over the 1999-2016 period in Figure 1.1B. The majority of German MNC innovation is conducted at home in Germany. Offshore R&D is concentrated in rich, developed Western economies at the technological frontier, with the top hubs the same as in Figure 1.1A. Over the entire period, 19% of all offshore patents originate in the US, whereas France, Austria and Switzerland contribute 8%, 7% and 6%, respectively. Appendix Table A.1 illustrates the overall top-5 foreign innovation hub ranking, together with snapshots for 2000 and 2015.

We distinguish between three types of MNC innovation locations, by combining information on the network of production affiliates and patent inventors: (i) at home in the headquarter country, (ii) offshore co-located, i.e. in a country with an affiliate present, and (iii) offshore not co-located, i.e. in a third country with no affiliate present.[9]

Patent characteristics. PATSTAT Global contains detailed bibliographical data that extends beyond patents' applicants, inventors, and underlying invention. In particular, each patent file records the technological classification of the patented technology, preceding patents and non-patent literature the patent application cites (*backward citations*), and subsequent patents that cite it (*forward citations*). Drawing on techniques from the innovation literature, we exploit this information to categorize the type of innovation activity underlying each EP patent and to evaluate its quality.

[9] Intuitively, (i) and (ii) are likely to be performed within firm boundaries, although they could in principle be subcontracted to independent parties in locations with MNC presence. On the other hand, (iii) is unlikely to be performed in-house, as any foreign subsidiary of relevant size should be visible in MiDi.

Figure 1.1A: MNC Patenting Activity Across Invention Locations: # MNCs

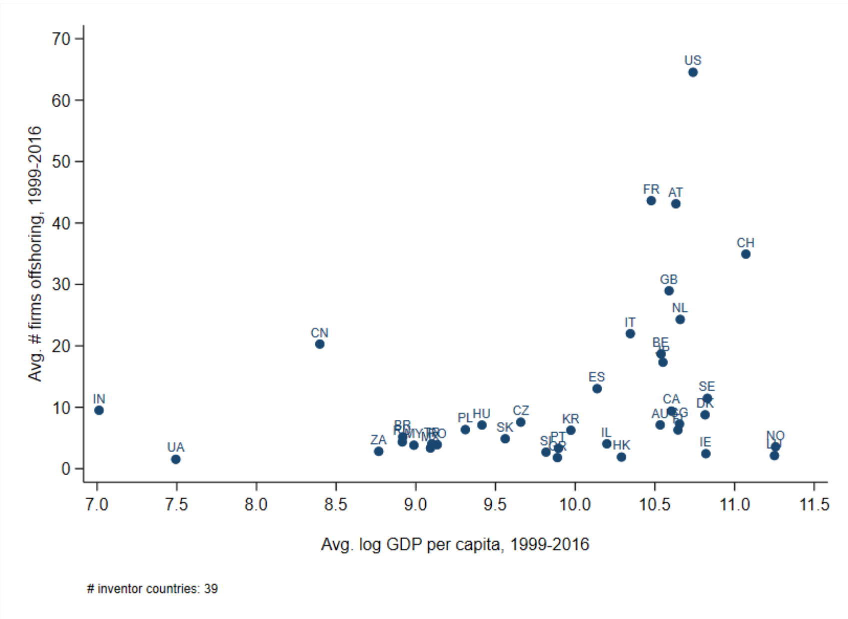
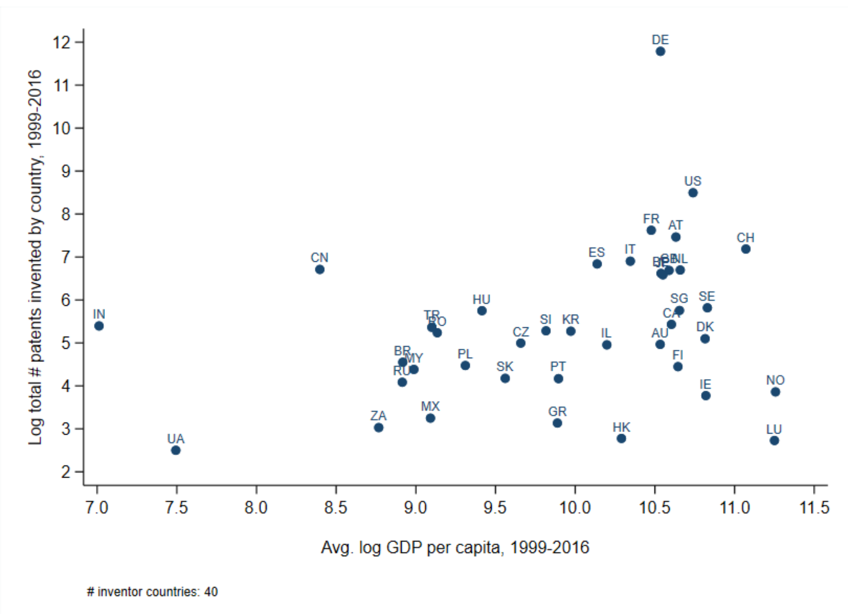


Figure 1.1B: MNC Patenting Activity Across Invention Locations: # MNC Patents



Notes: Figures 1.1A and 1.1B plot the average annual number of firms that offshore innovation and the log total number of patents invented in each country against inventor country's average log GDP per capita in 1999-2016. Patents with inventors from multiple locations are assigned to each country using equal fractions. The sample comprises 39 foreign inventor countries hosting innovation by at least 10 MNCs. *Data sources*: Research Data and Service Center of the Deutsche Bundesbank, MiDi, 1999-2016, combined with PATSTAT and World Bank National Accounts data, own calculations.

We distinguish between two types of patents depending on the kind of research firms have engaged in: science-based ("basic") or non-science-based ("applied"). In particular, we follow Ahmadpoor and Jones (2017) to measure a patent's distance to fundamental science with the minimum backward citation steps to a scientific article, using the Marx and Fuegi (2020) open-access dataset of patent front-page citations. We obtain distance to science scores for all EP patent applications where backward citations are available.[10] A patent that directly cites a scientific article receives a score of 1. A patent that does not itself cite a scientific article, but cites a patent that does so, receives a score of 2 because it is 2 degrees removed from science. This measure thus produces a score of $\{1, 2, 3, \ldots\}$, with lower scores indicating a more proximate connection to fundamental science. We label patents receiving a score of $\{1, 2\}$ as *science-based* and patents receiving a score of at least 3 as *non-science-based*. Since these two patent types conceptually result from basic and applied R&D respectively, we also interchangeably use the labels *basic* patent and *applied* patent for expositional simplicity.[11]

We also quantify each patent's innovation quality with the number of forward patent citations it receives. This methodology follows common practice in the literature, and rests on the premise that innovation of higher quality acts as a stepping stone for more subsequent innovation activity (Harhoff et al., 1999; Hall, Jaffe, & Trajtenberg, 2005). We count a patent family's forward citations received from EP applications within 5 years since its first filing. This standard measure ensures comparability in impact quality across patents filed at different times.

Panel C of Table 1.1 provides an overview of patent activity at the firm level in the matched MiDi-PATSTAT baseline dataset. Since innovation can be time-consuming and patenting sporadic, we collapse the time dimension and aggregate across all years a firm is active in the panel. On average, a patenting German multinational generates roughly 148 patents and 64 EP patents over the 1999-2016 period, where the distribution is extremely skewed with standard deviations of 1,464 and 538 respectively. On average, firms file 18 science-based and 45 non-science-based EP patents. Additionally, they generate 11 EP patents with at least one inventor located abroad. There is significant variation in patent quality across multinationals, with the mean number of citations a firm receives standing at 176 and its standard deviation reaching 1,601.[12]

Panel D of Table 1.1 presents summary statistics at the patent level. Overall, 28% of all EP patents are classified as science-based and 16% are developed abroad, of which 12% in a country with a foreign production affiliate and 4% in a country with no subsidiary. While the

[10] Backward citations can differ across patent applications within the same family depending on the patent office that receives it (Michel & Bettels, 2001). We account for this by computing the measure for European patents alone.

[11] One can also differentiate between product and process innovation by relying on textual analysis of patent abstracts as in Danzer, Feuerbaum, and Gaessler (2020). We document systematic patterns in the data based on the basic-applied distinction, and find little variation of interest by further distinguishing between product and process innovation.

[12] We compute summary statistics that involve citation counts over the truncated panel 1999-2011 given the 5-year forward citation window we take into account.

average patent attracts 2.1 citations, by this measure patent quality is generally higher for basic innovation: While science-based patents receive 2.92 citations on average with a high standard deviation of 6.47, the corresponding metrics for non-science-based patents are about 50% lower at 1.81 and 3.55. This is in line with prior evidence consistent with patents closer to science being more valuable (Ahmadpoor & Jones, 2017; Krieger, Schnitzer, & Watzinger, 2023). At the same time, patent quality also appears to vary across innovation locations. Offshore not co-located patents stand out with a mean citation count of 3, followed closely by offshore co-located patents with a mean of 2.7. Both notably outperform home-grown patents with a mean of only 2 citations. This suggests that MNCs may offshore R&D to locations where there maintain no production operations to tap specific local knowledge and expertise that allow them to develop valuable inventions.

1.2.2 Stylized Facts

The matched MiDi-PATSTAT database is unique in painting a comprehensive picture of the global innovation activity of multinational firms. We begin by establishing three novel stylized facts that emerge from this data.

Fact 1: *MNCs innovate actively and frequently abroad.*

The data reveal that 30% of all German multinational companies file one or more patents during 1999-2016. Moreover, 43% of the firms covered in our baseline matched MiDi-PATSTAT dataset have at least one patent with one or more inventors located outside of Germany.[13] More strictly, roughly 31% of the MNCs have at least one patent with all inventors located abroad. At the patent level, 14% of all patents and 16% of all EP patents in our sample are considered offshore, having at least one inventor abroad, as suggested by Table 1.1 Panel D.

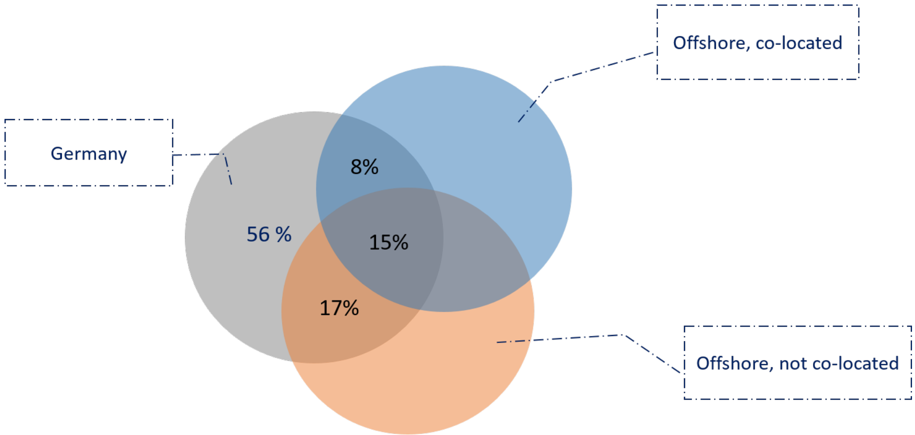
Fact 2: *MNCs innovate in multiple locations and offshore innovation to locations both with and without affiliates.*

We next document the extent to which German multinational firms offshore research activities to foreign countries. The Venn diagram in Figure 1.2 summarizes the global organization of MNC innovation at the firm level over the period of interest. 56% of all innovating multinationals file only patents for inventions that have been developed at home. Some 8% conduct patent-generating innovation both at home and in another country where they maintain a production affiliate, while 17% do so both at home and in a third country with no subsidiary. Fully 15% of firms undertake patented research in all three location types. For consistency,

[13] A subset of patents report multiple inventors, some of which may be located in Germany and some abroad. We consider such patents as invented abroad. Note that in cases with a mixed inventor team, the share of inventors located abroad is non-trivial, averaging at roughly one third. Patents with mixed inventors and those with a full team located abroad are equally represented in our baseline sample.

these figures describe the baseline sample of EP patents, but very similar patterns obtain when considering all patents in Appendix Figure A.1.

Figure 1.2: Location of Global MNC Innovation



Notes: This Venn diagram summarizes the global organization of German MNC patenting activity over 1999-2016. Each segment indicates the share of firms that file EP patents with inventors residing at home in Germany, offshore in a country with an MNC affiliate, and/or offshore in a country with no MNC affiliate. N = 2,030 MNCs. *Data sources*: Research Data and Service Center of the Deutsche Bundesbank, MiDi, 1999-2016, combined with PATSTAT, own calculations.

Table 1.2 provides a complementary summary of the global geography of innovation at the level of the individual patent. 83% of all EP patents filed by German MNCs result from innovation activity within Germany. Of the patents generated abroad, 72% are invented in countries where the MNC runs a production affiliate, with the remaining 28% not co-located with production. Distinguishing between science-based and non-science-based patents reveals that basic research is disproportionately more likely to be offshored and to be offshored to locations without a production affiliate: 23% of all science-based patents are generated abroad, of which 69% in countries with a subsidiary. In comparison, only 15% of non-science-based patents originate abroad, of which 75% in countries with a subsidiary.

Fact 3: *Larger MNCs innovate at higher intensity and quality.*

The binscatters in Figures 1.3A and 1.3B indicate a strong positive relationship between firm size and innovation intensity and innovation quality, respectively. We assign firms into ten bins based on their annual global sales, allowing firms to move across bins over time. Figure 1.3A plots the plots the log average annual number of EP patents per firm in each firm size bin. Similarly, Figure 1.3B shows the average number of 5-year forward citations per EP patent per firm, by size bin. We absorb year fixed effects in order to account for trends in patenting activity and potential concerns regarding citation truncation. Appendix Figures A.2A and A.2B replicate the stark positive relationships in these graphs for the full sample of patents. For com-

Table 1.2: Geography of MNC Global Innovation

	All EP Patents			EP basic			EP applied		
Innovation location	N	%	% within offshore	N	%	% within offshore	N	%	% within offshore
Germany	125,737	83.14		32,339	77.44		90,790	85.15	
Offshore co-located	18,473	12.22	72.47	6,532	15.64	69.34	11,871	11.13	74.99
Offshore not co-located	7,017	4.64	27.53	2,888	6.92	30.66	3,959	3.71	25.01

Notes: This table presents the distribution of inventor locations across all EP patents of German MNCs in 1999-2016. Patents are clasified into science-based patents (basic) and non-science-based ones (applied) based on backward citations to scientific journal articles. Patents can be invented at home, offshore in a country with an MNC affiliate (offshore co-located), and offshore in a country with no MNC affiliate (offshore not co-located). *Data sources*: Research Data and Service Center of the Deutsche Bundesbank, MiDi, 1999-2016, combined with PATSTAT, own calculations.

pleteness, Appendix Figures A.3 and A.4 provide binscatters that group firms by their number of patented inventions. These confirm that MNCs that file more patents tend to generate more cited ideas.

1.3 Theoretical Framework

Motivated by the empirical facts established in Section 2.2, we develop a theoretical model of multinational activity that characterizes the global organization of firms' production and innovation. We adopt a stylized, partial-equilibrium setting in order to transparently illustrate the key economic mechanisms that govern firm decisions, while retaining analytical tractability.

1.3.1 Set-up

Consider a world comprised of three countries: West, East and South. In each country, a continuum of heterogeneous firms produce horizontally differentiated goods which they sell at home and potentially also abroad. Consumers exhibit love of variety, such that the representative consumer in country $j = \{W, E, S\}$ has CES utility $U_j = \left[\int_{i \in \Omega_j} \left(x_{ji}\right)^{\alpha} di\right]^{\frac{1}{\alpha}}$, where x_{ji} is the quantity consumed of variety i, and Ω_j is the set of goods available to j. The elasticity of substitution across products is $\sigma \equiv 1/(1-\alpha) > 1$, with $0 < \alpha < 1$. If total expenditure in country j is R_j, j's demand for variety i is $x_{ji} = R_j P_j^{\sigma-1} p_{ji}^{-\sigma}$, where $P_j = \left[\int_{i \in \Omega_j} \left(p_{ji}\right)^{1-\sigma} di\right]^{\frac{1}{1-\sigma}}$ is the ideal price index, and p_{ji} is the price of good i in market j.

In each country, two types of labor engage respectively in manufacturing consumption goods and in innovation. Firms take the wages of production and innovation workers, w_j and r_j, as exogenously determined in the labor market. This assumption can be microfounded, for example, with the presence of two freely tradeable homogeneous goods, each produced by a

Figure 1.3A: MNC Size and Innovation Intensity

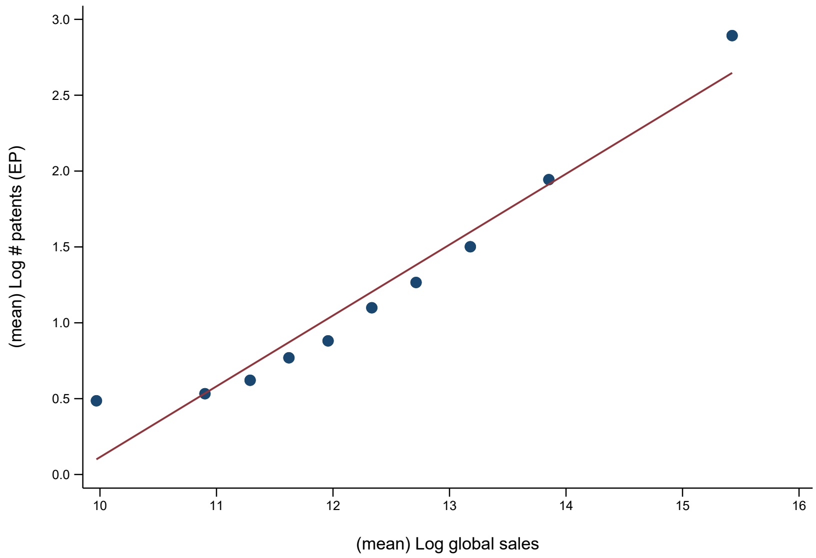
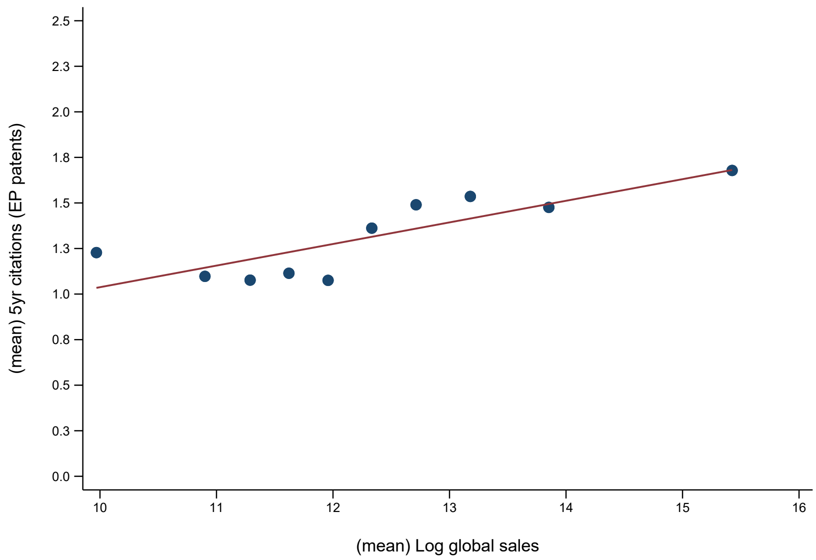


Figure 1.3B: MNC Size and Innovation Quality



Notes: These binscatters plot the log average annual number of EP patents per firm in 1999-2016 and the average number of 5-year forward citations per EP patent per firm in 1999-2011, by firm size bin. German MNCs are assigned to ten bins each year according to their annual global sales. Year fixed effects are absorbed. *Data sources:* Research Data and Service Center of the Deutsche Bundesbank, MiDi, 1999-2016, combined with PATSTAT, own calculations.

different type of labor under constant returns to scale and fixed aggregate labor endowments.

We are interested in understanding the production and innovation decisions of multinational companies. We therefore examine the operations of firms headquartered in West, and interpret the exogenous variation in w_j and r_j as cross-country differences in comparative advantage in production vs. innovation. To focus on meaningful trade-offs in Western firms' profit maximization, we assume that $w_S < w_W$, $w_S < w_E$, $r_W \leqslant r_S$, and $r_E \leqslant r_S$. This ensures that South has absolute and comparative advantage in production compared to both East and West, while East and West have comparative (and potentially absolute) advantage in innovation compared to South.[14]

1.3.2 Production Technology

Western entrepreneurs incur sunk entry costs associated with setting up headquarters. They face ex-ante uncertainty about their production efficiency, and draw productivity $\varphi \in (0, \infty)$ from distribution $G(\varphi)$ upon entry. Firm operations entail fixed costs of headquarter services f^H that must be performed at home. However, production can be offshored, such that the marginal cost of manufacturing in country j is w_j/φ.

Upon observing their productivity draw, firms either exit immediately or commence production and potentially become multinational and/or innovate. Western firms face a trade-off when deciding whether to locate production at home or abroad: Setting up a foreign affiliate implies additional fixed costs f^{FDI} associated with plant equipment, local management, and remote monitoring by headquarters, but it may reduce variable costs if host-country production wages are lower or if there are profitable complementarities with innovation activities.

1.3.3 Innovation Technology

Western firms can choose whether, where, and how much to invest in two types of innovation: basic and applied. Each innovation activity can be performed by headquarters at home, in-house by a foreign production affiliate, and/or at arm's length by a foreign unaffiliated party. Firms can choose to innovate in multiple locations at the same time, with innovation costs additively separable and innovation returns as specified below.

Applied innovation increases profits today and forever. Applied innovation of quality $q_j^A \geqslant 0$ improves production efficiency and lowers marginal production costs to $w_j/\left(1+q_j^A\right)\varphi$. This is qualitatively isomorphic to applied innovation enhancing product appeal and hence demand, for example by improving product quality, marketing competence, or packaging and delivery.

[14]While wages in advanced economies might in practice be higher than in emerging markets for both production and innovation workers, the assumed wage pattern can be seen as accounting for cross-country differences in the quality of innovators and of complementary inputs to innovation outside the model.

Basic innovation raises the probability of higher future profits. Given an exogenous death rate δ, the present discounted value of the future stream of profits for a firm with per-period profits $\pi(\varphi)$ is $\pi(\varphi)/\delta$. We conceptualize basic innovation as higher per-period profits or lower death rate, such that basic innovation of quality $q_j^B \geqslant 0$ boosts the present value of expected profits to $\left(1+q_j^B\right)\pi(\varphi)/\delta$. This is a reduced-form way of introducing dynamic returns to basic innovation, for instance because basic innovation is a prerequisite for subsequent successful applied innovation.

These two types of innovation can be illustrated with an intuitive example: if a pharmaceutical company discovers a new chemical reaction today (basic innovation), this could improve its future chances of developing a more effective drug formulation or a more efficient production process (applied innovation).[15]

Innovation costs increase with innovation quality, and depend on the location and organization of innovation activity. The cost of innovation of quality q in country j is $\mathbf{1}(q_j^{RD} > 0) r_j \left(f_{j,ORG}^{RD} + \frac{\left(q_j^{RD}\right)^\beta}{\beta} \right)$, where r_j is the inventor wage in j, $RD = \{B,A\}$ indicates the type of innovation ($B =$ basic, $A =$ applied), $ORG = \{I,O\}$ denotes whether innovation occurs within firm boundaries ($I =$ in-house, $O =$ outsource), and $\beta > 1$.[16]

We make three assumptions on the cost structure of innovation to build conceptual understanding. First, a Western multinational cannot perform in-house innovation abroad without having first set up a production affiliate. Formally, a firm must incur the fixed subsidiary costs f^{FDI} before that subsidiary can undertake any innovation, and when f^{FDI} is sufficiently high, it would never be optimal to establish pure innovation subsidiaries.

Second, the fixed cost of basic innovation is higher when it is conducted abroad, but is otherwise independent of the Western firm's organizational structure, $f_{W,I}^B = f_{W,O}^B < f_{E,I}^B = f_{E,O}^B = f_{S,I}^B = f_{S,O}^B$. This captures the idea that communication, monitoring and incentive provision require more financial and managerial resources when headquarterts need to supervise basic innovation at a distance and outside the firm's home jurisdiction.

Finally, a Western firm likewise faces higher fixed costs of applied innovation when it is offshored, but lower fixed costs in any location when it is performed in-house and therefore co-located with production, $f_{W,I}^A < f_{E,I}^A = f_{S,I}^A < f_{W,O}^A < f_{E,O}^A = f_{S,O}^A$. This reflects the scope for synergies between owner-operated production and applied research that can arise from frequent interactions between production and sales managers with practical know-how, scientists

[15]In a richer framework, we have considered multi-product firms that draw firm-wide productivity and firm-product specific expertise. Multi-product firms can then pursue applied process innovation to lower marginal production costs across all products and applied product innovation to lower product-specific fixed production costs. Our main theoretical results continue to hold, with more productive firms innovating more intensively and more frequently abroad across all innovation types.

[16]For tractability, we consider a static model that is qualitatively isomorphic to a dynamic model in which sunk innovation costs are captured by constant per-period amortized fixed costs.

with innovation talent, and technicians as two-way design and implementation liaisons.

To fix ideas, take the pharmaceutical example above. The assumptions on the innovation cost function mean that a stand-alone laboratory would be equally equipped to engineer new chemical reactions as a lab attached to a production unit, be it owner-operated or independent. By contrast, the R&D team at an owner-operated manufacturing facility would be best positioned to improve production methods (e.g., reduce gas dissipation) or product design (e.g., combo-vitamin pack), because it can benefit from the knowledge of site managers and easier implementation of test runs.

1.3.4 Firm Problem

Western firms face a multi-dimensional problem: they must choose the optimal location and scale of production, basic and applied innovation to maximize global profits. Optimal decisions are uniquely determined by productivity as the single dimension of firm heterogeneity. However, the model can in principle accommodate various patterns of MNC activity in different segments of the parameter space that govern countries' absolute and comparative advantage in production and innovation. Motivated by the stylized facts above, we make two simplifying assumptions in order to focus on the empirically relevant case and the novel mechanisms of interest. These assumptions yield considerable transparency and tractability with little loss of generality.

First, we abstract away from trade costs, such that all consumers have access to all varieties produced in the world. This implies that firms face the same global demand regardless of where they manufacture, captured by world aggregate expenditure R and a worldwide price index P.

Second, we posit that economies of scale in production are sufficiently strong (i.e. fixed FDI costs f^{FDI} are sufficiently high), such that firms find it optimal to concentrate manufacturing in one location and use it as a platform from which to serve all three markets. Moreover, production wages are sufficiently lower in South than in East to ensure that a Western multinational would always be incentivized to establish its single foreign subsidiary in South. Given this organizational structure, the relevant consideration set of country-specific fixed innovation costs for a multinational headquartered in West becomes $f^B_{W,O} < f^B_{E,O} = f^B_{S,I}$ for basic R&D and $f^A_{S,I} < f^A_{W,O} < f^A_{E,O}$ for applied R&D.

In this environment, a Western firm may choose to remain domestic and produce in-house at home. Such a firm may decide to innovate only in-house at home in W, only at arm's length abroad (in S and/or in E), or both. Alternatively, a Western firm may choose to become multinational and offshore production to an affiliate in S. This multinational may furthermore innovate only at home in W, only abroad (in-house at S and/or at arm's length in E), or both. In other words, firms' innovation strategy can span multiple locations and mix in-house and arm's length R&D. Of note, interdependencies between production and innovation can in principle make it profitable to offshore both, even if offshoring each activity alone might not be desirable.

Upon entry, a Western firm will determine its optimal production and innovation strategy in case it remained domestic and in case it established a foreign affiliate, and go multinational if the latter option is more profitable. With fixed FDI costs, firms above a certain productivity threshold will endogenously sort into multinational activity, consistent with the prior theoretical and empirical literature.

Given our interest in global MNC operations, we henceforth consider the profit maximization problem of a multinational company headquartered in West with a production affiliate in South and no subsidiary in East:

$$
\begin{aligned}
\max_{\mathbf{G}\equiv\left\{p,x,\left\{q_j^B,q_j^A\right\}\right\}} \pi(\varphi) &= \underbrace{\left(1+\sum_j q_j^B(\varphi)\right)\left(p(\varphi)x(\varphi)-\frac{x(\varphi)w_S}{\left(1+\sum_j q_j^A(\varphi)\right)\varphi}\right)}_{\widetilde{\pi}(\varphi)} \\
&\quad \underbrace{-f^H-f^{FDI}-\sum_{RD}\sum_j \mathbf{1}\left[q_j^{RD}(\varphi)>0\right]r_j\left(f_{j,ORG}^{RD}+\frac{\left(q_j^{RD}(\varphi)\right)^\beta}{\beta}\right)}_{F(\varphi)} \\
\text{s.t.}\quad x(\varphi) &= RP^{\sigma-1}p(\varphi)^{-\sigma}.
\end{aligned}
\tag{1.1}
$$

The MNC global strategy is characterized by the location of production (here, South), the output quantity x and price p, and the incidence and quality q_j^{RD} of each innovation activity RD in each location j. We denote this strategy as $\mathbf{G}\equiv\left\{p,x,\left\{q_j^B,q_j^A\right\}\right\}$. Note that innovation costs are additively separable across locations and innovation types. In contrast, innovation returns are not, because applied innovation additively reduces marginal production costs, while basic innovation multiplicatively increases variable profits $\widetilde{\pi}(\varphi)$.

Optimal Production Conditional on Innovation Strategy

The MNC problem (1.1) can be reduced to first determining the optimal production level and pricing conditional on an innovation strategy and then identifying the optimal innovation strategy. In particular, given $\left\{q_j^B,q_j^A\right\}$, the maximization problem is isomorphic to that of a firm with exogenously set overhead costs, marginal production costs, and actuarial profit factor. Under monopolistic competition and CES consumption preferences, firms therefore optimally charge a constant mark-up $1/\alpha$ above marginal cost, and generate the following output quantity and sales revenues:

$$p\left(\varphi,\left\{q_j^B,q_j^A\right\}\right)=\frac{w_S}{\alpha\left(1+\sum_j q_j^A(\varphi)\right)\varphi}, \tag{1.2a}$$

$$x\left(\varphi,\left\{q_j^B,q_j^A\right\}\right)=RP^{\sigma-1}\alpha^{\sigma}w_S^{-\sigma}\left(1+\sum_j q_j^A(\varphi)\right)^{\sigma}\varphi^{\sigma}, \tag{1.2b}$$

$$r\left(\varphi,\left\{q_j^B,q_j^A\right\}\right)=R\left(P\alpha/w_S\right)^{\sigma-1}\left(1+\sum_j q_j^A(\varphi)\right)^{\sigma-1}\varphi^{\sigma-1}. \tag{1.2c}$$

Note that greater applied innovation directly enables firms to set lower prices and thereby earn higher sales and variable profits. By contrast, basic innovation does not directly affect production choices, but it may do so indirectly through the joint decision that the firm makes over both types of innovation. Note also that conditional on an innovation strategy, more productive firms as usual set lower prices and earn higher sales and profits. We will see below that this advantage gets amplified by the higher innovation intensity they endogenously choose.

Optimal Innovation Strategy

The global production and innovation strategy that maximizes MNC profits can be determined by incorporating the optimal production strategy conditional on innovation activity from equations (1.2a) and (1.2b) into equation (1.1) and solving for the optimal innovation strategy in the reduced firm problem:

$$\begin{aligned}\max_{\left\{q_j^B,q_j^A\right\}}\ \pi(\varphi) \quad = \quad & \underbrace{R\left(P\alpha/w_S\right)^{\sigma-1}\left(1+\sum_j q_j^B(\varphi)\right)\left(1+\sum_j q_j^A(\varphi)\right)^{\sigma-1}\varphi^{\sigma-1}/\sigma}_{\tilde{\pi}(\varphi)} \\ & \underbrace{-f^H-f^{FDI}-\sum_{RD}\sum_j \mathbf{1}\left[q_j^{RD}(\varphi)>0\right]r_j\left(f_j^{RD}+\frac{\left(q_j^{RD}(\varphi)\right)^{\beta}}{\beta}\right)}_{F(\varphi)}.\end{aligned} \tag{1.3}$$

The firm faces a complex choice set with respect to the global organization of its innovation activity. It can in principle choose to conduct each of basic and applied R&D in any subset of the three possible country locations and at varying intensity levels. The global innovation strategy can thus be characterized by a vector of 6 non-negative innovation quality levels, $\left\{q_W^B,q_E^B,q_S^B,q_W^A,q_E^A,q_S^A\right\}$, which are jointly determined by the following set of first-order conditions:

$$\frac{\partial \pi(\varphi)}{\partial q_j^B} = 0 \Longleftrightarrow R(P\alpha/w_S)^{\sigma-1}\left(1+\sum_j q_j^A(\varphi)\right)^{\sigma-1}\varphi^{\sigma-1}/\sigma = r_j\left(q_j^B(\varphi)\right)^{\beta-1},\ q_j^B(\varphi) \geqslant 0, \tag{1.4a}$$

$$\frac{\partial \pi(\varphi)}{\partial q_j^A} = 0 \Longleftrightarrow$$

$$R(P\alpha/w_S)^{\sigma-1}\left(1+\sum_j q_j^B(\varphi)\right)\left(1+\sum_j q_j^A(\varphi)\right)^{\sigma-2}\varphi^{\sigma-1}(\sigma-1)/\sigma = r_j\left(q_j^A(\varphi)\right)^{\beta-1},\ q_j^A(\varphi) \geqslant 0. \tag{1.4b}$$

Although there is no closed-form solution to equations (1.4a)-(1.4b), the optimal innovation strategy exhibits properties that inform the underlying economic mechanisms and allow us to derive comparative statics of interest.

A key feature of the firm problem is that innovation decisions will be interdependent across countries. Consider first applied innovation. From equation (1.4b), the optimal amount of applied innovation in any given location will depend on the global level of applied innovation. This arises because the returns to applied innovation accrue at the firm level, and manifest in lower marginal production costs regardless of where production takes place. Applied innovation will be complementary across locations if $\sigma > 2$ and $\partial^2\pi(\varphi)/\partial q_j^A \partial q_{j'}^A > 0$, substitutable across locations if $1 < \sigma < 2$ and $\partial^2\pi(\varphi)/\partial q_j^A \partial q_{j'}^A < 0$, and independent across locations in the knife-edge case of $\sigma = 2$ and $\partial^2\pi(\varphi)/\partial q_j^A \partial q_{j'}^A = 0$. Estimates of σ in the 3,5 range in the literature suggest that applied R&D is in practice likely complementary across countries within firms.

Equation (1.4b) further implies that applied innovation in any given location - and therefore also globally - will be complementary with the total and regional levels of basic innovation, $\partial^2\pi(\varphi)/\partial q_j^A \partial q_{j'}^B > 0$. This results from basic innovation amplifying variable profits, which rise whenever applied innovation lowers marginal production costs. This means, for example, that any shock that encourages a firm to undertake more basic innovation will induce it to also conduct more applied innovation, and vice versa.

Consider next basic innovation. From equation (1.4a), optimal basic innovation in any one location does not directly depend on basic innovation elsewhere. This occurs because expected profits increase linearly with firm-level global basic innovation. However, optimal local basic innovation rises with total applied innovation and its components, $\partial^2\pi(\varphi)/\partial q_j^B \partial q_{j'}^A > 0$, which are implicit functions of global basic innovation. As a result, there is complementarity in basic innovation intensity across locations, $\partial^2\pi(\varphi)/\partial q_j^B \partial q_{j'}^B > 0$.

Finally, how much basic and applied innovation a firm performs in a given country depends on local conditions and its global levels of basic and applied innovation, but not on the geographic and implicitly organizational (in-house vs. arm's length) composition of these global levels. In particular, while basic and applied innovation are complementary in raising production profits, they incur additively separable costs across locations and innovation types, i.e. $\partial^2 F(\varphi)/\partial q_j^A \partial q_{j\prime}^A = \partial^2 F(\varphi)/\partial q_j^B \partial q_{j\prime}^B = \partial^2 F(\varphi)/\partial q_j^A \partial q_{j\prime}^B = 0$. Since innovation costs depend on innovation wages r_j and organizational structure, the optimal q_j^{RD} will therefore be a function of its type and location and of the total levels of applied and basic innovation, but not directly on the latter's location.

1.3.5 Theoretical Predictions

The integrated model of global production and innovation activity delivers rich predictions for the pattern of MNC operations. We focus here on the novel results for the optimal innovation strategy of multinational firms. These stem from the combination of firm heterogeneity in productivity and the rich structure of innovation costs and returns that depend on its location, integration and quality.

We consider first the incidence and intensity of innovation activity across firms:

Proposition 1. *More productive MNCs are more likely to innovate and to innovate more intensively.*

Proof. See Appendix A.1. □

More productive multinational companies will be incetivized to innovate more actively for two reasons. Along the intensive margin, firm profits are supermodular in productivity and innovation quality of either basic or applied type, $\partial^2 \pi(\varphi)/\partial \varphi \partial q_j^B > 0$ and $\partial^2 \pi(\varphi)/\partial \varphi \partial q_j^A > 0$. Intuitively, applied innovation directly amplifies the advantage of more productive firms via multiplicatively lower marginal production costs, while basic innovation multiplicatively augments variable profits. Along the extensive margin, innovation entails fixed costs that more productive firms can more easily amortize because they earn higher revenues and profits. These extensive and intensive margin patterns are true for each type of innovation activity, basic and applied. They hold for any given location and, aggregating across locations, also for innovation activity at the firm level.

We turn next to firms' optimal location and management of research and development:

Proposition 2. *More productive MNCs are more likely to offshore innovation and to innovate in more countries.*

Proof. See Appendix A.1. □

More productive firms will be more likely to innovate abroad because of economies of scale in both production and innovation. Recall that profits are supermodular in productivity and innovation intensity, while innovation in any given location entails fixed costs. *Ceteris paribus*, there will thus be a minimum productivity cut-off $\varphi^*_{j,RD}$ above which innovation of type RD in country j becomes profitable. Since fixed costs abroad are higher than at home ($f^B_{W,O} < f^B_{E,O} = f^B_{S,I}$ for basic R&D and $f^A_{S,I} < f^A_{W,O} < f^A_{E,O}$ for applied R&D due to co-location advantage), this productivity threshold will tend to be higher for offshore innovation (except potentially for applied innovation at home in West vs. at the affiliate in South).

By the same logic, more productive MNCs will also be more likely to offshore innovation to more countries. Productivity cut-offs $\varphi^*_{j,RD}$ will generally vary across countries depending on local fixed innovation costs, local inventor wages, and model parameters that govern consumer demand and production and innovation technologies. More productive multinationals will clear the minimum threshold for more innovation locations, for at least one of the two types of R&D.

Recall that we consider the innovation strategy of multinationals headquartered in West that operate a subsidiary in South. If such a multinational opts to innovate in both East and South, it would be conducting both in-house and arm's length innovation abroad. Proposition 2 thus implies the following corollary:

Corollary 1 *More productive MNCs are more likely to innovate both in locations with and in locations without a production affiliate.*

The model also speaks to the variation in MNC innovation activity across countries based on their comparative advantage in innovation:

Proposition 3. *An MNC is more likely to innovate and to innovate more intensively in countries with lower inventor wages.*

Proof. See Appendix A.1. □

When deciding where to undertake innovation, firms find it advantageous to choose locations with lower inventor costs. Along the extensive margin, there is a maximum inventor wage r_j, above which innovating is not profitable because of the fixed innovation costs. Along the intensive margin, firms optimally pursue higher-quality R&D in countries with lower r_j. This can be readily observed from the first-order conditions (1.4a) and (1.4b). Consider equation (1.4a) for basic innovation. The left-hand side contains only variables at the firm level, including total basic innovation, while the right-hand side increases with both r_j and $q^B_j(\varphi)$. Hence $\partial q^B_j(\varphi)/\partial r_j < 0$. The analogous result for applied innovation, $\partial q^A_j(\varphi)/\partial r_j < 0$, follows from equation (1.4b).

While our baseline model considers an economy with a single manufacturing sector and a single innovation sector, the analysis can be extended to a world with multiple manufacturing sectors that map to multiple innovation sectors based on their relevant technological area. If inventor wages vary both across countries and sectors, Proposition 3 would imply that innovation activity responds to countries' comparative advantage in innovation across sectors:

Corollary 2 *An MNC is more likely to innovate and to innovate more intensively in a given sector in countries with lower innovator wages in that sector.*

Finally, the model has implications for the co-location of production and innovation activities, and thereby for the internalization of innovation activity abroad.

Proposition 4. *Applied innovation is more likely to be co-located with production than basic innovation.*

Proof. See Appendix A.1. □

Innovation technology is such that there are synergies between applied innovation and production when performed in the same facility. In particular, the fixed costs of applied innovation are strictly lower when it is co-located with production. All else constant, this implies that a multinational will be more likely to find it profitable to pursue applied R&D in countries where it also operates a manufacturing affiliate.[17]

1.4 Empirical Evidence

The theoretical framework above can rationalize the stylized facts documented in Section 1.2.2 for the global organization of MNC innovation activity. Through the lens of the model, MNCs have an incentive to perform R&D both at home and abroad, in order to increase current and/or future profits (*Fact 1*). Moreover, the variation in innovation activity across firms can be attributed to more productive MNCs choosing to innovate at greater intensity, which manifests in the data as both more patents and higher average patent quality as measured by patent citations (*Fact 3*).

The model also suggests that MNCs may offshore innovation to benefit from cross-country differences in inventor wages. They can furthermore choose whether or not to co-locate each type of foreign innovation with foreign production depending on the associated innovation

[17]The cost synergies between applied innovation and production may also manifest in lower variable innovation costs, such that total applied innovation costs are $f^A_{j,ORG} + \mu^A_{j,ORG} \frac{\left(q^{RD}_j\right)^\beta}{\beta}$, where $f^A_{S,I} < f^A_{W,O} < f^A_{E,O}$ as in the baseline and $\mu^A_{S,I} < \mu^A_{W,O} = \mu^A_{E,O} = 1$. If so, both the incidence and the quality of applied innovation would be higher when co-located with production.

costs and returns. The observed distribution of offshoring and co-location of innovation activity across multinationals can thus be attributed to heterogeneous firm productivity, combined with variation across innovation types within firms (*Fact 2*). For example, the least productive MNCs may opt to innovate only at home, while the most productive MNCs may undertake applied innovation in locations with an affiliate (where production wages are low and there are synergies with production) and basic innovation in locations without an affiliate (where inventor wages are low and synergies with production are irrelevant).

In this section, we show that the model's predictions find strong empirical support in the global operations of German multinationals that go beyond rationalizing *Facts 1-3*.

1.4.1 Estimation Approach

We evaluate Propositions 1-4 and Corollaries 1-2 in the data by estimating variants of three empirical specifications at different levels of aggregation:

$$I_{ft} = \alpha + \beta\varphi_{ft} + \delta_s + \delta_t + \varepsilon_{ft}, \tag{1.1}$$

$$I_{fact} = \alpha + \beta\varphi_{ft} + \gamma_{RCA}RCA_{act} + \delta_a + \delta_c + \delta_t(+\delta_f) + \varepsilon_{fact}, \tag{1.2}$$

$$I_{fpt} = \alpha + \beta\varphi_{ft} + \gamma_{RD}D_{RD=A} + \delta_a + \delta_t(+\delta_f) + \varepsilon_{fpt}. \tag{1.3}$$

In regression (1.1), the outcome variable I_{ft} reflects various aspects of multinational firm f's innovation activity in year t, such as an indicator for having any patents, the (log) number of patents, and the (log) average number of citations per patent. The main variable of interest on the right-hand side, φ_{ft}, is a proxy for parent-firm productivity. We condition on year fixed effects, δ_t, to absorb fluctuations in aggregate supply and demand conditions. For instance, δ_t would capture any changes in Germany in production and innovation wages, tax regime, or trade and investment promotion policies. We further account for observed and unobserved sector characteristics that may govern innovative activity with a full set of 23 sector fixed effects, δ_s, based on the primary industry of activity of each parent company in the 2-digit NACE 2.0 classification. These subsume, for example, cross-sector differences in factor intensities, technological scope for fragmenting and offshoring manufacturing and R&D, synergies between production and applied innovation, and innovation costs and returns more broadly. We conservatively cluster errors ε_{ft} by firm, to allow for correlated shocks within firms over time.

We follow common practice in the literature, and use (log) global firm sales as our baseline proxy for a multinational firm's productivity φ_{ft}. As in standard heterogeneous-firm models of international trade and investment, in our framework too global firm sales are monotonic in firm productivity. The main advantage of using this proxy is that it poses minimal data requirements and is not subject to potential estimation biases in constructing productivity measures from accounting statements. In particular, while rich in many dimensions, the MiDi data on German MNCs is not sufficiently detailed to permit rigorous total factor productivity estimates.

Through the lens of the model, innovation activity and total revenues are joint outcomes of the firm's profit maximization problem. We therefore interpret coefficient β as a conditional correlation consistent with the model's predictions, rather than the causal effect of productivity underlying it. As we demonstrate below, the results are robust to using (log) sales per worker at the MNC headquarters as an alternative indicator of labor productivity.

In regression (1.2), we unpack firms' patent activity to analyze outcomes I_{fact} that characterize firm f's innovation within technological area a in country c at time t. These include the (log) number of patents and (log) number of citation-weighted patents generated. Given the sparsity in firm patenting activity in the data, we consider 3 non-overlapping 5-year periods t (2002-2006, 2007-2011, 2012-2016), where the outcome variable aggregates up patenting activity within the period of interest.[18] We map each patent to one of 34 technology areas following Schmoch (2008), in order to assess the heterogeneity in innovation activity across firms and countries within technology areas. In addition to parent firm productivity, φ_{ft}, we now also consider countries' revealed comparative advantage by technology area and year, RCA_{act}, as constructed below. Country fixed effects, δ_c, condition on the overall institutional and economic environment in a given country, to isolate the role of differences in innovation conditions across technology areas within countries. Year and technology area fixed effects, δ_t and δ_a, in turn account for supply and demand factors analogously to year and sector fixed effects in regression (1.1). More stringent versions of equation (1.2) add firm fixed effects, to further control for time-invariant firm characteristics that shape patent activity irrespective of technology area or country of invention. We continue to cluster errors ε_{fact} at the firm level.

We construct a novel measure of countries' revealed comparative advantage, RCA_{act}, that conceptually maps to inventor wages in the model. This measure aims to capture countries' capacity to enable patent-generating innovation in different technological classes. We define RCA_{act} as the number of patents generated in technology area a in country c at time t, as a percent share of all patents originating in that country and period. Scaling by the total number of patents ensures that the variation in RCA_{act} across countries is not driven by country size, and implicitly also subsumes cross-country differences in absolute advantage in innovation.

To build an informative RCA_{act} measure, we first identify all patent families in PATSTAT that contain patent applications filed on three continents, i.e. with at least three of the top five leading patent authorities in the world. In particular, we consider patent families that include at least one application each at the European Patent Office (EPO); at the United States Patent and Trademark Office (USPTO); and at the Japan Patent Office (JPO), Korean Intellectual Property Office (KIPO) or China National Intellectual Property Administration (CNIPA).

[18] We restrict the panel to 2002-2016 in this specification to feature three periods of equal duration.

This ensures some degree of comparability in quality across patents, as only higher-quality inventions are generally patented in multiple jurisdictions (Harhoff, Scherer, & Vopel, 2003; de Rassenfosse et al., 2013). We assign each patent to its inventor country or countries, using fractional counts as explained above.[19] To avoid circularity, we exclude patent families with German applicants. Given this definition of RCA_{act} and the inclusion of country and technology area fixed effects, specification (1.2) thus identifies how the innovation strategy of multinational firms responds to a country's comparative advantage in innovation in a given technology area, relative to other countries and technology areas.

Finally, in regression (1.3), we examine innovation outcome I_{fpt} at the most granular level of individual patents, indexed by the firm-patent-year triplet fpt. In these specifications, I_{fpt} is a binary indicator for offshore innovation being co-located with production. In addition to parent-firm productivity, φ_{ft}, the main right-hand side variable of interest is a dummy for non-science-based patents representing applied R&D, $D_{RD=A}$, with science-based patents (i.e. basic R&D) the excluded category. Year, technology area and firm fixed effects, δ_t, δ_a and δ_f, control for firm idiosyncrasies and exogenous variation in innovation conditions across time and technology areas. We once again cluster errors ε_{fpt} by firm, this time to accommodate correlated shocks to research and development operations across time and space within firms.

1.4.2 Innovation Intensity

We first provide evidence that innovation activity varies systematically with total firm sales, in a way consistent with Proposition 1 that more productive firms are more likely to innovate and to innovate more intensively. To examine the extensive margin, we estimate specification (1.1) in the full panel of German multinationals in MiDi, where we set the outcome variable to a binary indicator for any patenting activity by firm f in year t. To evaluate the intensive margin, we then consider the log number of patents, the log number citation-weighted patents, and the average log number of citations per patent by firm-year. Conceptually, these variables can be seen as proxying the quantity and quality of innovation, respectively, which are isomorphically captured by q in the model.

Panel A of Table 1.3 establishes that larger MNCs pursue systematically more innovation activities. Column 1 first indicates that they have a significantly higher probability of filing any patents in a given year. In Columns 2-4, we further observe that bigger MNCs own more patents that are cited more frequently on average. All coefficient estimates are highly statistically significant at the 1% level. The economic magnitudes are also sizeable: the estimates imply that doubling firm size is associated with 4 percentage points higher probability of patent-generating innovation, and, conditional on patenting, approximately 50% more patents and 2.4% more citations per patent.

[19] Appendix A.2.4 elaborates on the construction of the RCA measure.

Table 1.3: Innovation Intensity

Dependent variable	(1) any patent (0/1)	(2) log # patents	(3) log # citation weighted patents	(4) avg log # citations
Panel A. EP patents				
Log global sales	0.039***	0.495***	0.490***	0.024***
	(0.002)	(0.027)	(0.033)	(0.004)
# MNC-years	68,999	9,545	6,180	9,545
Panel B. EP Basic (science-based)				
Log global sales	0.101***	0.407***	0.404***	0.017*
	(0.006)	(0.037)	(0.047)	(0.007)
# MNC-years	9,007	3,986	2,543	3,986
Panel C. EP Applied (non-science-based)				
Log global sales	0.019***	0.483***	0.475***	0.023***
	(0.003)	(0.028)	(0.034)	(0.004)
# MNC-years	9,007	8,217	5,227	8,217
Sector FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: This table examines the relationship between firm size and innovation intensity for German MNCs in 1999-2016, based on equation (1.1). Panel A includes all MNCs and all EP patents of innovating MNCs. Panels B and C restrict the sample to all innovating MNCs and their science-based and non-science-based EP patents, respectively. Patents are classified into science-based (basic) and non-science-based (applied) based on backward citations to scientific journal articles. The outcome variable is an indicator for any patents in Column 1, the log number of patents in Column 2, the log number of citation-weighted patents in Column 3, and the average log number of citations per patent in Column 4. Standard errors clustered by firm. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Data sources:* Research Data and Service Center of the Deutsche Bundesbank (MiDi, 1999-2016) and PATSTAT, own calculations.

Panels B and C of Table 1.3 confirm that larger multinationals have superior innovative performance within each R&D type. In particular, we repeat the regression analysis in Panel A separately for science-based and non-science-based patenting, in the subsample of patenting multinationals. Since not every innovating company owns patents of both types, the number of observations varies across specifications. We find that the probability of filing any patent is an order of magnitude more sensitive to firm size for basic R&D than applied R&D. Conditional on some innovative activity, by contrast, both the frequency and the quality of patenting is equally elastic with respect to firm size across the two innovation types.

While our baseline analysis covers EP patents in order to ensure patent comparability and have information on patent type, stable results hold when we broaden the sample to consider all patents in Appendix Table A.2. Separately, we also observe qualitatively similar patterns in Appendix Table A.3 when we proxy firm productivity with the parent headquarters' log sales per employee instead of firm size. As a caveat, some extensive-margin coefficients turn negative. We attribute this to the lack of precision of this labor productivity proxy, especially in light of how outsourcing production and innovation might influence employment levels and composition at headquarters.

1.4.3 Innovation Offshoring

We next establish that larger multinationals are more likely to innovate abroad, to undertake innovation in more foreign countries, and to offshore a greater share of their total innovation activity. These findings are in line with Proposition 2, and consistent with the presence of both high returns and sizeable fixed costs associated with offshore R&D.

We first analyze whether a firm conducts any offshore innovation by estimating specification (1.1) in the panel of patent-active MNCs, with an indicator for at least one patent originating abroad as the outcome of interest.[20] We present the results in Panel A of Table 1.4. Column 1 establishes that bigger MNCs are disproportionately more likely to file patents for inventions developed abroad. Columns 3 and 5 provide consistent evidence for the incidence of offshore science-based and non-science-based patents, respectively. All estimates are highly statistically and economically significant: A doubling of global sales is associated with approximately 10 percentage points higher probability of pursuing innovation of either type outside of Germany.

We then consider the intensive margin of offshore innovation, and study how the share of offshore patents varies with firm size in Panel B of Table 1.4. Larger multinational companies do not simply scale up domestic and offshore R&D activity proportionately. Instead, they generate a bigger share of their patent portfolio in foreign locations. On average, a firm double the size would develop 1.8 percentage points more of its patents abroad, as seen in Column 1. Columns 3 and 5 demonstrate that this pattern is almost identical for science-based and non-science-based patents.

Panel C of Table 1.4 confirms that larger MNCs offshore R&D to more foreign locations, by setting the outcome variable in equation (1.1) to the number of foreign inventor countries that firms' patents originate from. Odd columns establish this result first for all patents and then separately for science-based and non-science-based patents. A doubling of firm size corresponds to 0.8-1 more offshore innovation locations. Even columns explore the extent to which this reflects bigger multinationals operating more production subsidiaries worldwide, by expanding the specification to include the number of host countries in an MNC's affiliate network. Multinationals that produce in more countries are also more likely to innovate in more countries, with a somewhat larger elasticity for non-science-based patents than for science-based patents. At the same time, while the coefficient on firm size falls by approximately 40% for basic R&D and approximately 50% for applied R&D, it remains highly significant. In other words, bigger multinationals pursue patent-generating R&D in more countries even controlling for the number of their production locations.

[20] As discussed earlier, we label patents as foreign-invented if at least one of its inventors resides outside of Germany.

Table 1.4: Innovation Offshoring

	EP patents		Basic EP patents		Applied EP patents	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Dependent variable: Any offshore patent (0/1)						
Log global sales	0.115***		0.090***		0.106***	
	(0.005)		(0.009)		(0.006)	
# MNC-years	9,545		3,986		8,217	
Panel B. Dependent variable: Share offshore patents						
Log global sales	0.018***		0.017**		0.017***	
	(0.004)		(0.005)		(0.004)	
# MNC-years	9,545		3,986		8,217	
Panel C. Dependent variable: # foreign inventor countries						
Log global sales	0.953***	0.522***	0.869***	0.510***	0.824***	0.401***
	(0.208)	(0.122)	(0.212)	(0.130)	(0.216)	(0.103)
# affiliate countries		0.094*		0.069+		0.088**
		(0.040)		(0.037)		(0.032)
# MNC-years	2,920	2,920	1,327	1,327	2,309	2,309
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table examines the relationship between firm size and offshore innovation activity for innovating German MNCs, based on equation (1.1). The dependent variable is an indicator for any foreign-invented patents in Panel A, the share of patents invented abroad in Panel B, and the number of host countries for foreign-invented patents. The sample includes all EP patents in Columns 1-2, all science-based EP patents in Columns 3-4, and all non-science-based EP patents in Columns 5-6. Patents are classified into basic (science-based) and applied (non-science-based) based on backward citations to scientific journal articles. Standard errors clustered by firm. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Data sources: Research Data and Service Center of the Deutsche Bundesbank (MiDi, 1999-2016) and PATSTAT, own calculations.

Appendix Table A.4 shows that the geographic composition of MNCs' innovation activity varies systematically not only with firm size, but also with headquarters' log sales per worker, as a proxy for labor productivity. In particular, similar results emerge for the propensity to offshore R&D and the share of offshore patents across all patents and within patent type.

Through the lens of the model, these patterns are consistent with the presence of synergies between offshore innovation and production, especially for applied innovation, as well as with pull factors to undertake R&D even in locations without a production base, especially for basic innovation. The findings are also indicative of firms facing fixed innovation costs at the country level, which MNCs can more easily amortize if operating at a larger scale.

Finally, we evaluate the implication of Corollary 1 that bigger multinationals are more likely to pursue research and development in locations both with and without a production affiliate. To this end, we estimate a multinomial logit regression on the set of MNCs that develop patents abroad. The outcome is a categorical variable that distinguishes between three mutually exclusive strategies for offshore innovation at the firm-year level: (1) any offshore not co-located R&D (i.e. at least one patent with inventors located in a country with no affiliate), (2) any offshore co-located R&D (i.e. at least one patent with inventors located in a country with an affiliate), and (3) both co-located and not co-located offshore R&D (i.e. at least one offshore patent with inventors in each type of location). We regress this outcome variable on firm size, conditioning on year and sector fixed effects and clustering by firm as above.

The results in Columns 1-2 of Table 1.5 indicate that larger multinationals indeed have a greater propensity to offshore R&D both in countries with and without a production subsidiary, compared to inventing in either location type alone. Columns 3-4 confirm that this is not driven by bigger MNCs maintaining production facilities in more host countries. The analysis also reveals that among firms with a single mode of offshore R&D (either only co-located or only not co-located), larger multinationals are more likely to co-locate foreign invention and production. This pattern can be fully attributed to their greater number of subsidiary host countries. Appendix Table A.5 documents similar patterns when we instead consider firms' headquarter labor productivity (proxied by log sales per worker) in place of firm size. Of note, this measure of labor productivity drops both in magnitude and statistical significance when we condition on the number of affiliate locations. We expect this relates to the endogeneity of offshored production and employment retained at headquarters.

1.4.4 Innovation Comparative Advantage

We next demonstrate that MNC innovation activity responds to cross-country differences in comparative advantage across technology areas, as per Proposition 3 and Corollary 2. In particular, within a given technology area, firms develop systematically more patents and receive more patent citations as a marker of innovation quality in countries with a strong revealed comparative advantage for innovation in that technology area.

To explore this pattern, we estimate specification (1.2) at the firm-technology area-country-period level, using the log number of patents and the log number of citation-weighted patents as the outcomes of interest. We use the latter as a more comprehensive measure of the quality-weighted extent of innovation activity. When aggregating at this level of analysis, each patent is assigned to one main technology area and, if relevant, split equally among inventor countries as explained above.

Table 1.6A presents robust evidence that economies with stronger RCA_{act} in a given technology field attract significantly more innovation activity by German multinationals in that field. Columns 1 and 4 establish this baseline result conditioning on a full set of country, technology

Table 1.5: Mixed Innovation Offshoring

Base level: Any offshore not co-located patent				
	(1)	(2)	(3)	(4)
Any offshore co-located patent				
Log global sales	0.433***	0.454***	−0.019	−0.028
	(0.052)	(0.053)	(0.055)	(0.061)
# affiliate countries			0.110***	0.118***
			(0.015)	(0.017)
Both co-located and not co-located offshore patents				
Log global sales	0.663***	0.732***	0.390***	0.389***
	(0.061)	(0.069)	(0.088)	(0.095)
# affiliate countries			0.097***	0.107***
			(0.016)	(0.018)
Year FE	Yes	Yes	Yes	Yes
Sector FE	No	Yes	No	Yes
# MNC-years	2,931	2,925	2,931	2,925

Notes: This table examines the relationship between firm size and the choice of offshore innovation locations for German MNCs with offshore innovation, based on a multinomial logit regression. The dependent variable takes the value 1 if the firm has any offshore patents invented in a country without an affiliate, value 2 if it has any offshore patents invented in a country with an affiliate, and value 3 if it has both co-located and not co-located offshore patents. Standard errors clustered by firm. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Data sources:* Research Data and Service Center of the Deutsche Bundesbank (MiDi, 1999-2016) and PATSTAT, own calculations.

area, and period fixed effects, such that the main coefficient of interest is identified from the variation in comparative advantage within a country across technology areas and within an area across countries.

We next establish that countries' comparative advantage for innovation shapes the allocation of offshore R&D activity even within firms across space. We first condition on firm size alone in Columns 2 and 5. In Columns 3 and 6, we then add a full set of MNC firm fixed effects, and additionally control for the presence of a subsidiary in a given location. These specifications account for the variation in innovation incidence with firm size and other firm unobservables, as well as for potential benefits from co-locating production and innovation. The estimated impact of RCA_{act} remains highly statistically significant, and its economic magnitude is largest under the most stringent specification with firm fixed effects. The estimates suggest that a unit increase in RCA_{act} in a given location would attract roughly 2% more of a firm's patents in a given technology area.

Finally, we distinguish between science-based and non-science-based patents in Table 1.6B and find consistent results that revealed comparative advantage is a strong driver of the loca-

Table 1.6A: Innovation Comparative Advantage

Dependent variable	log # patents			log # cit. weighted patents		
	(1)	(2)	(3)	(4)	(5)	(6)
RCA	0.014***	0.017***	0.023***	0.018**	0.023***	0.027***
	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.007)
Avg. log global sales		0.094***			0.117***	
		(0.021)			(0.018)	
Affiliate country = 1			0.122***			0.068
			(0.032)			(0.064)
Observations	8,741	7,762	7,475	4,970	4,404	4,167
Tech area FE	Yes	Yes	No	Yes	Yes	No
Period FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes

Notes: This table examines the relationship between countries' revealed comparative advantage for innovation in a given technology area and offshore innovation activity by German MNCs across countries and technology areas, based on equation (1.2). Data is aggregated into three non-overlapping five-year periods (2002-2006, 2007-2011, 2012-2016). The outcome variable is the log number of patents in Columns 1-3, and the log number of citation-weighted patents in Columns 4-6. Standard errors clustered by firm. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Data sources:* Research Data and Service Center of the Deutsche Bundesbank (MiDi, 1999-2016) and PATSTAT, own calculations.

tion MNCs each type of R&D. Additionally, repeating the analysis for citation-weighted patent counts separately by innovation type produces similarly strong results for each type. The coefficient estimate on the dummy for the presence of a production affiliate is higher for applied than for basic R&D. This is consistent with the presence of stronger synergies in co-locating applied innovation in close proximity to production activity and anticipates the next set of results.

1.4.5 Innovation Co-location

We conclude by assessing multinationals' strategy with respect to co-locating foreign production and innovation activities in line with Proposition 4. We document that conditional on offshoring innovation, firms are systematically more likely to co-locate applied R&D with production than basic R&D. This is consistent with proximity to manufacturing experience being more synergistic with applied innovation, for instance if close interactions between production managers and scientists can enable cheaper and more effective applied research.

We analyze co-location strategies at the patent level in Table 1.7. For each patent, we construct a binary variable equal to 1 if the patent belongs to a parent company with an active production affiliate in the country of invention. We use this indicator as the outcome of interest in estimating equation (1.3). Columns 1-2 suggest that non-science based patents are

Table 1.6B: Innovation Comparative Advantage by Patent Type

Panel A: Basic (science-based patents)

Dependent variable	log # patents			log # cit. weighted patents		
	(1)	(2)	(3)	(4)	(5)	(6)
RCA	0.009^{+}	0.014^{**}	0.022^{***}	0.013^{*}	0.019^{**}	0.026^{**}
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.008)
Avg. log global sales		0.084^{***}			0.110^{***}	
		(0.015)			(0.020)	
Affiliate country = 1			0.104^{*}			0.016
			(0.041)			(0.076)
Observations	4,005	3,677	3,535	2,490	2,279	2,156

Panel B: Applied (non-science-based patents)

Dependent variable	log # patents			log # cit. weighted patents		
	(1)	(2)	(3)	(4)	(5)	(6)
RCA	0.013^{**}	0.014^{**}	0.019^{***}	0.018^{*}	0.025^{**}	0.026^{**}
	(0.005)	(0.005)	(0.005)	(0.008)	(0.009)	(0.009)
Avg. log global sales		0.082^{**}			0.104^{***}	
		(0.027)			(0.018)	
Affiliate country = 1			0.106^{*}			0.077
			(0.042)			(0.070)
Observations	5,968	5,343	5,050	3,126	2,785	2,567
Tech area FE	Yes	Yes	No	Yes	Yes	No
Period FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes

Notes: Table 1.6B replicates the analysis in Table 1.6A separately for science-based patents in Panel A and for non-science-based patents in Panel B. Patents are classified in each type based on backward citations to scientific journal articles. Standard errors are clustered at firm level. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Data sources:* Research Data and Service Center of the Deutsche Bundesbank (MiDi, 1999-2016) and PATSTAT, own calculations.

Table 1.7: Innovation Co-location

	co-located offshore patent (0/1)					
	(1)	(2)	(3)	(4)	(5)	(6)
Non-science-based patent (0/1)	0.028	0.012^{+}			0.039^{**}	0.013^{*}
	(0.025)	(0.006)			(0.015)	(0.006)
Log global sales			0.113^{***}	0.111^{***}	0.113^{***}	0.111^{***}
			(0.008)	(0.026)	(0.008)	(0.026)
Tech. area FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes
# patents	22,051	21,819	21,997	21,765	21,997	21,765

Notes: This table examines the propensity of German MNCs to co-locate offshore science-based and non-science-based patent invention with a production affiliate, based on equation (1.3). Patents are classified based on backward citations to scientific journal articles. The outcome variable is an indicator for a patent being invented in a country where the MNC has an affiliate. Standard errors clustered by firm. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.*Data sources:* Research Data and Service Center of the Deutsche Bundesbank (MiDi, 1999-2016) and PATSTAT, own calculations.

more likely to be developed alongside production operations than science-based patents. This relationship is moreover statistically significant in the more stringent specification that exploits the variation across patents and locations within firms by conditioning on firm fixed effects.

Since firm size and R&D type jointly determine MNCs' patent location decisions, we examine their combined role in the remainder of Table 1.7. Columns 3-4 first confirm that larger multinationals are more likely to develop their new technologies in host countries with an active subsidiary. This result is related to, but goes beyond the predictions of Proposition 2 and Corollary 1 that more productive firms are more likely to innovate in multiple locations, including countries with and without production affiliates. Once we control for this firm size effect in Columns 5-6, we find strong evidence that applied innovation is systematically more likely to be co-located with production than basic innovation, even when looking across patents within firms.

1.5 Conclusion

Multinational companies play a central role in both global value chains and frontier R&D. We provide one of the first integrated analyses of MNCs' global production and innovation strategy. We establish novel stylized facts using uniquely rich data on the network of production affiliates and patents of German multinationals. We rationalize these facts with a heterogeneous-firm model in which companies jointly choose the location, scale and integration of manufacturing, basic innovation and applied innovation. Empirical evidence consistent with the model indicates that more productive MNCs innovate more intensively in terms of the number and quality of patents. Such companies also offshore more innovation to more countries, spanning both countries with and without a production affiliate. Finally, MNCs pursue innovation across countries and technology classes following countries' comparative advantage, with applied innovation more likely to be co-located with production than basic innovation.

Our findings open the door to various avenues for future research. Richer information on the inputs and outputs of innovation activities, such as data on both R&D investment and successful patenting, can provide a more holistic understanding of the factors governing MNC operations. Also of interest is the role of intellectual property rights protection and general contract enforcement for the location and integration of MNC production and innovation.

It is likewise important to evaluate the implications of MNCs' globalized production and innovation for the design of trade and innovation policy. These implications will inform the scope for multilateral agreements, especially as developed and developing countries occupy different segments of global value chains and engage differently in technological innovation and adoption. For example, our work points to complementarity rather than substitutability in innovation activity across countries, which may alleviate concerns about the impact of offshoring innovation on sending economies. MNC operations may also shape the impact of technological leaps such as automation on the global distribution of production, innovation and adoption, and thereby on economic growth across countries.

2

Top R&D Investors and the Net-Zero Transition

Recent Trends in Green Patenting

2.1 Introduction

Climate change mitigation has become one of humanity's greatest challenges. Despite the increasing number of concerning pieces of evidence, political discussions, or commitments to targets, the current level of action seems to be lagging behind. Reaching the ambitious net-zero goal by mid-century can not be achieved through political engagement alone, but requires significant investment in R&D, technology development and diffusion. The urgent deployment of currently available low-carbon technologies is hindered by their lack of competitiveness compared to existing fossil-fuel based solutions (IEA, 2023). On the other hand, the net-zero target is currently unattainable when considering only the existing set of technologies (IRENA, 2017). The International Energy Agency (IEA) predicts that by 2050, half of the global CO_2 emission reductions will come from technologies that are currently only in the demonstration or prototype phase (IEA, 2023). Innovation is therefore a critical pillar through which the current global climate crisis can be addressed.

This paper provides a comprehensive overview of recent trends in green technology development by the world's leading R&D investing firms. In particular, the analysis focuses on firms' most internationally relevant green inventions as reflected in their patenting activity over the period 2012-2019. The paper sheds light on the overall contribution of top R&D investors to global green innovation efforts, identifies the sectors, countries and companies that are pioneering in this area, and provides insights into the nature of green and non-green innovation pursued by these firms. In addition, the analysis explores dynamics since the adoption of the Paris Agreement in 2015, the most important international coordination effort towards curbing climate change.[1] Specifically, I investigate whether these innovative firms have shifted their R&D efforts towards green inventions in light of the growing international commitment towards the net-zero goal entailed by the Paris Agreement.

Top R&D investors play a particularly important role in global innovation, as they are the firms with the highest capacity to allocate resources to innovative projects. Compared to smaller firms, such leading innovators are less financially constrained and would be able to bear the higher risk often associated with green technologies. Such technologies require longer-term investments or higher fixed costs (e.g., in energy and transport markets), and their commercial viability is often uncertain (OECD, 2011; Vysoká et al., 2021). Given that top R&D investors are typically large multinational companies, they can drive the faster adoption of technological advancements through their global operations. Furthermore, open commitments towards the net-zero goal have been coming from large corporations as well, and not only from governments (IEA, 2021). The scope and timeframe of these pledges vary significantly across companies, some not having outlined any specific way in which their targets are to be achieved. Against this background, it becomes relevant to investigate whether the urgency for new technological solutions is reflected in the recent innovation endeavours of the world's most prominent R&D investors.

I do so by combining data from two major sources. I leverage information on the world's leading 2000 R&D investing firms from the EC-JRC-OECD COR&DIP database (Amoroso et al., 2021) and link it with PATSTAT, the global patent statistics database provided by the European Patent Office (EPO). Top R&D investors are the firms with the highest absolute R&D expenditures in the world in 2018, as in Hernández et al. (2020). From all retrieved patents, I restrict attention to those filed internationally across multiple jurisdictions, as they are known to capture higher-value inventions (Harhoff, Scherer, & Vopel, 2003; Dechezleprêtre, Ménière, & Mohnen, 2017). Specifically, I focus on *triadic* patent families that include applications filed at three major patent offices: the US, European and Japanese patent offices. I obtain a sample of 1,506 top R&D investors and approximately 3.5 million corresponding patent families filed between 2012 and 2019, of which 270,787 are triadic. I identify green patents using

[1] The Paris Agreement is an international treaty adopted in 2015 under the United Nations Framework Convention on Climate Change (UNFCCC). It aims to guide nations in limiting the global temperature rise to 1.5°C above pre-industrial levels. Reaching net-zero emissions by 2050 is crucial for achieving this goal. Retrieved on 26/02/2024 from https://www.un.org/en/climatechange/net-zero-coalition.

the EPO's *Y02 tagging scheme*, a standardized classification meant to target technologies that mitigate or provide adaptation to climate change (Angelucci, Hurtado-Albir, & Volpe, 2018).[2] I build on and enhance the work of Amoroso et al. (2021) by extending the analysis period to 2012-2019, while circumventing numerous data construction challenges. This extension enables me to uncover recent trends in the green patenting activity of leading R&D investing firms, moving beyond a mere snapshot and providing a more comprehensive understanding of their contribution.

The findings indicate that their contribution to the global development of green technologies has been substantial in recent years. Overall, top R&D investors own 67% of all green patents filed between 2012 and 2019, with shares even higher than 80% in climate change mitigation technologies related to 'ICT' and 'Transportation'. However, both their relative contribution and the total number of green patents show a declining trend over this period. Specifically, since 2012, the number of green patents of top R&D investors decreased annually by approximately 5%. This decline, together with the decreasing share of green patents among all patents filed by these firms, suggests a shift towards the development of non-green inventions in recent years.

There is significant sectoral concentration among top R&D investors in terms of their green innovation efforts, with 'Transport equipment', 'Chemicals' and 'Computers & Electronics' collectively accounting for 55% of green patents. I capture sectoral specialization in the development of climate-related technologies by computing the share of green patents in all patents by sector. The results uncover the substantial variation in the level of specialization across sectors, with 'Electricity, gas & steam', 'Transport equipment', 'Chemicals' and 'Electrical equipment' standing out with a high relative share of green patents. Despite the transportation sector's large impact on global emissions, firms active in 'Transport equipment' have surprisingly been focusing relatively less on developing green technologies, as evidenced by the declining share of green patents in this sector from 2012 to 2019.

I analyze the geographical distribution of top R&D investors' green innovation, by looking at the location of their patent inventors. Japan stands out as the country where the majority of green technologies developed by top R&D investors originate from. However, Korea shows the highest specialization in climate-related technologies, with 30% of all patents invented there by the underlying firms being classified as green. Following Korea, high specialization can be observed in the case of Japan, France and Germany, with a share of green patents between 12-15%. Notably, Korea has experienced a remarkable increase in the level of specialization since 2014, indicating its growing prominence in green technology development. In contrast, Japan and the other major European economies show a slight decrease in the share of green patents invented there. Interestingly, the United States displays the most pronounced shift away from

[2] In the remainder of this paper, I use "climate change mitigation technologies", "green technologies" and "climate-related technologies" interchangeably for expositional simplicity. All patents with a *Y02* tag are referred to as "green" patents.

green technology, with the share of green patents halving over the 2012-2019 period.

I next document patterns in the development of climate-related inventions at the firm level. I show that 42% of all green patents of interest can be attributed to only 20 firms. This result highlights the significant influence of just a few firms in shaping the recent green patenting landscape, particularly when taking into account the high contribution of top R&D investors in global green patenting. On average, roughly 11% of all patents of a firm are classified as green. I provide evidence that even at the firm level, both the number of and the share of green patents have been on a decreasing trend since the adoption of the Paris Agreement in 2015. This pattern remains even when controlling for sector and firm fixed effects.

Lastly, I show that green patents of top R&D investors not only receive more citations relative to their non-green counterparts, but they are also cited by patents from a more diverse set of technological fields. This suggests that green patents exhibit, on average, higher quality and have a greater influence on the development of subsequent technologies across a broader set of domains. Furthermore, they are building on a more diverse set of knowledge sources compared to non-green patents, suggesting that they are more original. Interestingly, there is no significant difference in radicalness between green and non-green patents of top R&D investors. This concept reflects the extent to which patents draw upon knowledge from areas beyond their own technological domain. Thus, although firms draw on a wider spectrum of technologies for their green inventions, the references they rely on are not significantly more distant from their own inventions than those used for their non-green inventions. In addition, firms rely on more recent sources of knowledge in developing their green technologies. This is evidenced by the shorter average backward citation lag and the significantly higher share of recent references observed for green patents. There is no evidence that top R&D investors have significantly changed the nature of their green innovation as a response to the adoption of the Paris Agreement in 2015. Specifically, green patents filed as of 2015 do not appear to be cited more, nor broader, and they do not exhibit higher levels of originality or radicalness. Surprisingly, these recent cohorts of green patents appear to rely on relatively older knowledge sources compared to those filed before 2015.

The main contributions of this paper are as follows. First, it sheds light on recent advancements in green inventions as reflected in the global green patenting landscape. Empirical evidence on recent trends is scarce. Previous studies mostly examine earlier developments, for instance Dechezleprêtre et al. (2011) who focus on the period 1978-2005. An exception is Probst et al. (2021), who analyze global developments in green patenting between 2013 and 2017. Similar to this paper, they document a decline in the number of green inventions. I advance this literature by focusing on recent trends in patenting of high-value inventions by the most prominent R&D investors worldwide. I provide additional insights from examining the sectoral and firm-level variation in specialization in climate-related inventions. By tracking developments through 2019, I am also able to assess changes in green patenting activity following the

2015 Paris Agreement. I contribute by highlighting the disparity between the increased perceived importance of climate change mitigation technologies and the decline in green patented inventions by large R&D investors. Second, I complement the existing literature exploring differences between green and non-green technologies (e.g., Barbieri, Marzucchi, & Rizzo, 2020; Dechezleprêtre, Martin, & Mohnen, 2022). I highlight the differences between firms' patented green and non-green technologies in terms of their quality and contribution to subsequent inventions, as well as the characteristics of the knowledge base they build upon.

The remainder of this paper is organized as follows. Section 2.2 elaborates on the main data construction steps and briefly describes the sample of firms included in the analysis. Section 2.3 presents trends in the green patenting activity of the top R&D investors, while Section 2.4 focuses on the comparison between the green and non-green patents of these firms. A discussion and conclusion are provided in Section 2.5.

2.2 Data and Summary Statistics

The main objective of this paper is to analyze recent trends in the development of green inventions by the worldwide top R&D investing firms. To do so, I combine data from various sources. First, I rely on the EC-JRC-OECD COR&DIP database that includes information on the R&D activity, intellectual property portfolio as well as a number of financial indicators for the top 2000 R&D investing firms over the 2016-2018 period Amoroso et al. (2021). Second, I extend the panel dimension of the database by linking it to PATSTAT, the worldwide patent statistics database maintained by the European Patent Office. In absence of a direct correspondence between R&D investors and patent applicants, linking the two datasets was the main challenge to overcome in building a dataset with an extended time period. Due to the fact that the original source contains patents of both parents and foreign affiliates of the top R&D investors, matching firm names with patent applicant names alone was not sufficient. Therefore, in the matching process, I additionally rely on corporate ownership information from Bureau van Dijk's Orbis and Orbis IP datasets, allowing me to identify patents of the firms' foreign affiliates. Lastly, for the purpose of comparing green and non-green patents of top R&D investors presented in Section 2.4, I complement the bibliographic information from PATSTAT with patent quality indicators retrieved from the OECD[3] (see Squicciarini, Dernis, & Criscuolo, 2013). In what follows, I briefly describe the main data sources and the data construction steps as well as basic summary statistics for the sample of firms that my analysis is focused on.

[3] OECD Patent Quality Indicators, version February 2022.

2.2.1 Data Sources

COR&DIP database. In a collaborative effort, the European Commission, Joint Research Center and the OECD Directorate for Science, Technology and Innovation maintain a database covering the world's top 2000 R&D investors. The COR&DIP database (v.3) contains not only information regarding the world ranking of the large corporations based on their R&D investments, but also financial indicators, such as net sales, capital expenditures, profits and the number of employees covering the period 2015-2018. More importantly, the database includes data on the firms' patent and trademark portfolios over the 2016-2018 period. The top R&D investors included in the dataset are identified based on their R&D investments in 2018, mirroring the 2019 edition of the EU Industrial R&D Scoreboard (Hernández et al., 2020).

The *Patent Portfolio* dataset represents the main point of reference for the underlying analysis. It captures patents filed by the top R&D investors over the short time span of only three years.[4] With the aim of providing meaningful insights into the recent contribution of top R&D investors to climate change mitigation, I build on this panel and extend the observation period to include patents filed during 2012-2019. Additionally, the original dataset contains only a specific set of patent applications. The analysis of Amoroso et al. (2021) focuses on IP5 patent families, defined as families that include at least one application at one of the main five IP offices in the world and at least one other application filed in any other patent office. The *Patent portfolio* dataset in COR&DIP includes only the applications filed at the top IP offices of these families, namely the European Patent Office (EPO), the United States Patent and Trademark Office (USPTO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO) and the China National Intellectual Property Administration (CNIPA). In contrast, in my extension, I initially retrieve all patent applications filed by the top R&D investors, i.e. the complete patent families. The original dataset provides a link between company unique identifiers and PATSTAT patent application identifiers and patent publication numbers. Therefore, I use the sample of already linked patents in order to establish a baseline crosswalk between COR&DIP companies and patent applicants in PATSTAT.

PATSTAT Global. I use the PATSTAT database to retrieve the additional patents and their corresponding bibliographic data. Specifically, besides identifying all patent applications of interest, I further collect information regarding patents' inventor location, technology classes, and forward and backward citations. I rely on PATSTAT Global version Spring 2023, such that I can capture recent developments in green technologies that span up to 2019, the most recent year which should be unaffected by truncation. I extend firms' IP portfolio by collecting all patent families with the first filing year between 2012-2019. A patent family contains a bundle of patent applications that pertain to a single invention being protected in one or across

[4] The *Patent Portfolio* dataset includes patent data from PATSTAT Global version Spring 2021. Due to lags in the patenting process via the international route (PCT), the patent counts for 2018 can still be affected by truncation and therefore incomplete.

multiple jurisdictions. I construct all patent-based indicators using patent families such that I avoid counting the same invention multiple times. For brevity, in the remainder of this paper I refer to patent families as patents interchangeably. Whenever a patent is co-owned by multiple top R&D investors, I assign equal fractions to each of them. Note that the data available does not allow me to track changes in patent ownership, therefore patents are assigned to companies based solely on information contained in patent publication documents.

I identify green inventions by using the EPO's *Y02 tagging scheme*, a classification meant to target technologies that mitigate or provide adaptation to climate change. In particular, the classification developed by the EPO identifies patents related to technologies that either "help reduce the emissions of greenhouse gases or actively enhance the sinks of such gases" (Angelucci, Hurtado-Albir, & Volpe, 2018). Patents are not tagged solely based on their CPC or IPC technological areas, but an algorithm was developed such that input from patent examiners and other parties is taken into account. According to the authors, attempts to identify patents related to climate change mitigation technologies only based on their technological areas would result in either omissions or too many false positives. The Y02 tagging scheme was introduced to provide a solution to this problem. The scheme distinguishes between eight main technological fields: climate change adaptation technologies (Y02A), climate change mitigation technologies related to buildings (Y02B), capturing of greenhouse gases (Y02C), information and communication technologies (Y02D), energy generation, transmission and distribution (Y02E), production of goods (Y02P), transportation (Y02T) and waste management (Y02W). In order to assess the contribution of top R&D investors to global green technology development, I also retrieve the total number of green patents filed between 2012 and 2019 from PATSTAT.

Linking top R&D investors and PATSTAT patent applicants. With the aim of extending the patent portfolio of the top R&D investors contained in the COR&DIP database, I use the already existing sample of patents in order to establish a link between the companies of interest and patent applicants in PATSTAT. First, I match company and patent applicant names with the goal of obtaining a list of linked company × applicant pairs. I assume that once a link is established, it holds true over the extended period 2012-2019. Second, I retrieve all patents filed by the identified applicants from the latest available PATSTAT version. I assign the newly collected patents to the top R&D investors based on the compiled list of company × applicant pairs. The original database includes all patents filed by the parent firms and also those filed by their subsidiaries. This implies that information on the corporate structure of the firms is necessary to obtain a complete patent portfolio. Moreover, relying solely on the matching of firm and patent applicant names would lead to an underestimation of the total amount of patents originating from these firms. I overcome this challenge by using the corporate structure employed by Amoroso et al. (2021), which can be inferred from the patent assignment in the original COR&DIP database and patent applicant information. Whenever possible, I refine the list of company × applicant pairs by using additional data from Orbis IP and Orbis Ownership to pin down correct links between parent companies and their foreign affiliates. Appendix B.1 provides a detailed description of the matching algorithm used in order to extend the patent portfolio.

2.2.2 Baseline Sample and Summary Statistics

The baseline sample of analysis includes 1,506 top R&D investors and 3,688,908 patent families with the earliest filing year between 2012 and 2019. Of these, roughly 60% contain applications filed at a single patent office, covering only one jurisdiction. However, due to the nature of the firms under investigation and the global perspective taken in approaching the research question, a focus on international patent families is preferred for two main reasons. First, they capture a certain level of quality of the underlying technology, given that patent protection is sought outside the country of origin (Harhoff, Scherer, & Vopel, 2003). Second, this approach eliminates concerns regarding countries' different propensity to file patents (de Rassenfosse et al., 2013; Dechezleprêtre, Ménière, & Mohnen, 2017). Among international patent families, I focus on *triadic* patent families, meaning those that include applications filed at the EPO, JPO and the USPTO and that share at least one priority filing. This indicator is widely used in the innovation literature and has been shown to be less affected by different propensities to patent across countries, to improve international comparability and to reflect higher value inventions (Dernis & Khan, 2004; Guellec & Pottelsberghe de la Potterie, 2005; de Rassenfosse et al., 2013; Aghion et al., 2016). The baseline sample includes 270,787 triadic patent families, representing 7.3% of the total amount of patents retrieved.

For robustness and comparability with Amoroso et al. (2021), I also analyze IP5 patent families, albeit with a slightly modified definition. In general, IP5 families include applications filed with at least two different patent offices, one of which being an IP5 office, as described above. In the specific case of families containing an EP application, I consider this criterion to be met if the second patent office is one outside the European Patent Convention. For example, if a family includes both an EP and a German patent application, the EP application would likely cover the German market, rendering the additional German patent application redundant when considering the international dimension of the patent protection. Therefore, I do not consider such cases to be equivalent to families with patents filed at the EPO and in, say, China or the United States. Using this IP5 definition, I capture 1,418,130 patent families, representing 38.4% of the total sample. I defer results based on the sample of IP5 patent families to Appendix B.3, where I also discuss similarities and differences with the main results.

Table 2.1 presents summary statistics at the firm-year level for the 1,506 R&D investors analyzed. On average, these large corporations generate €10 billion in net sales annually and employ roughly 31,000 individuals. There is significant variation across firms in terms of their

Table 2.1: Top R&D investors: Summary statistics

	N	Mean	SD	Median	p90
Net sales, mil. €	5,896	10,435.84	24,698.10	2,698.89	25,160.67
# employees	5,559	30,962.31	58,010.22	11,000.00	79,564.00
R&D investment, mil. €	5,964	446.94	1,229.75	109.13	912.09
# patents (all)	12,048	306.18	992.15	46.00	637.50
# patents (IP5)	12,048	117.71	389.99	18.00	239.00
# patents (triadic)	12,048	22.48	70.65	2.00	50.50
# green patents	12,048	29.52	121.20	1.00	53.00
# green patents (IP5)	12,048	12.86	52.62	1.00	22.50
# green patents (triadic)	12,048	2.86	14.63	0.00	5.00
Share green patents (IP5)	11,053	0.09	0.16	0.02	0.25
Share green patents (triadic)	8,000	0.11	0.21	0.00	0.33
# trademarks	4,518	18.64	48.02	5.00	43.00
# green trademarks	4,518	1.31	6.76	0.00	3.00
Share green trademarks	3,481	0.10	0.22	0.00	0.38

Notes: This table presents annual summary statistics for the sample of 1,506 top R&D investors included in the analysis. The unit of observation is firm-year. Financial and trademark-related indicators cover the period 2016-2018, retrieved from the original COR&DIP database. Patent-based indicators are computed for the period 2012-2019. As patenting and trademarking may not occur yearly, therefore the shares of green and trademarks are defined only in years in which firms apply for at least one patent or trademark, resulting in a smaller number of observations for those indicators. p90 indicates the 90^{th} percentile of the distribution.

annual amount of investment in R&D. While the median firm spends around € 110 million a year on R&D, a firm in the 90^{th} percentile allocates nearly € 1 billion to research endeavors. Table 2.1 also confirms that these are patent-intensive firms. They file roughly 300 patents per year, out of which 118 fall under the IP5 definition and 22 are triadic. 11% of triadic patents filed by a firm in the dataset in a given year are classified as green. A similar pattern emerges when looking at firms' trademark portfolio. On average, 10% of firms' yearly trademark applications are green.

Geographically, one third of the top R&D investors are based in the US (33%), the most represented economy in the sample. Aggregating at regional level, roughly 36% of the firms are based in Asia, 17% of them being Japanese and 11% Chinese. Only one quarter of the firms are based in European countries, Germany accounting for almost 7% of the top R&D investors. Figure B.2 illustrates the geographical distribution of the firms. The most represented sector in the sample is 'Computer & Electronics', comprising 22% of the top R&D investors. 12% of the firms are active in 'Pharmaceuticals', while both 'Machinery' and 'Transport equipment' account for 8% of the sample each. Figure B.3 displays the number of R&D investors by sector.

2.3 Green Patenting Activity of Top R&D Investors

This section examines recent trends in the patenting activity of the world's leading R&D investors, with a particular focus on their inventions related to climate change mitigation and adaptation. It highlights the role of these firms in driving the global development of green technologies and analyzes the extent to which their innovation efforts have been shifted towards such technologies in recent years. Furthermore, it explores variations in green technology specialization at the sector, country and individual firm levels and documents changes in specialization since the adoption of the Paris Agreement in 2015. The main analysis focuses on firms' triadic patents filed between 2012 and 2019. Results based on the broader sample of IP5 patents are presented in Appendix B.3.

2.3.1 The Contribution of Top R&D Investors to Green Technology Development

The world's top R&D investors have contributed substantially to green technology development over the last years. This is reflected in the share of global green patents that originate from these firms. Notably, 67% of all high-value green patents filed between 2012 and 2019 are attributed to the top R&D investors under investigation.[5] Figure 2.1 illustrates that top R&D investors have been driving climate change mitigation technology development in fields related to ICT (Y02D) and transportation (Y02T), filing more than 80% of all green patents in these categories. With the exception of adaptation (Y02A) and waste management technologies (Y02W), the majority of patents in all other fields are owned by the top R&D investors with shares of more than 60%.

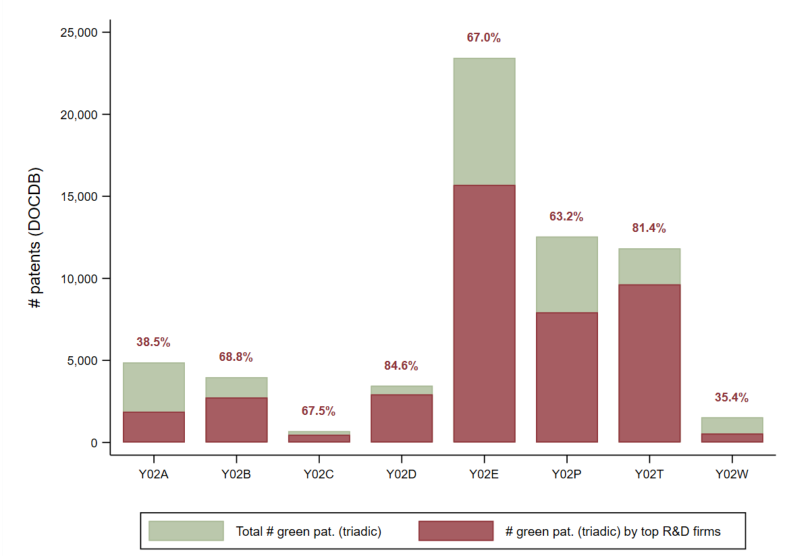
The high proportion of global green patents associated with these firms highlights the extent to which they have been shaping the green patenting landscape in recent years. This pattern underscores the importance of further analyzing their behavior in this regard. Figure 2.2 plots the evolution of the total number of green patents over time and the share that R&D investors account for. First, it reveals that, surprisingly, the total number of green patents has been declining over the last years.[6] Moreover, the share of green patents originating from top R&D investors has been decreasing over time, albeit still exceeding 60% in 2019. This pattern points to a decline in the number of green patents filed by the leading innovative firms and to

[5]As shown in Appendix B.3, this pattern is replicated when looking at the wider sample of IP5 patent families. Therefore, this finding is not driven by the patent family definition employed, meaning that it cannot be solely explained by a higher propensity of top R&D investors to seek patent protection in the three jurisdictions considered.

[6]Figure B.4 depicts the evolution of all green patents filed over time. Although the total number of patents exhibits a staggering increase, when conditioning on internationally filed and therefore higher-value patents, this increase vanishes completely. This figure alone suggests that using different definitions when compiling patent-based statistics can largely influence the results. Here, the rise is particularly driven by applications filed within a single jurisdiction, primarily in China. Figure B.5 replicates the analysis for the population of non-green patents, suggesting that the downward trend is specific to green inventions.

a mildly increasing contribution of relatively smaller firms to the development of high-quality green inventions.

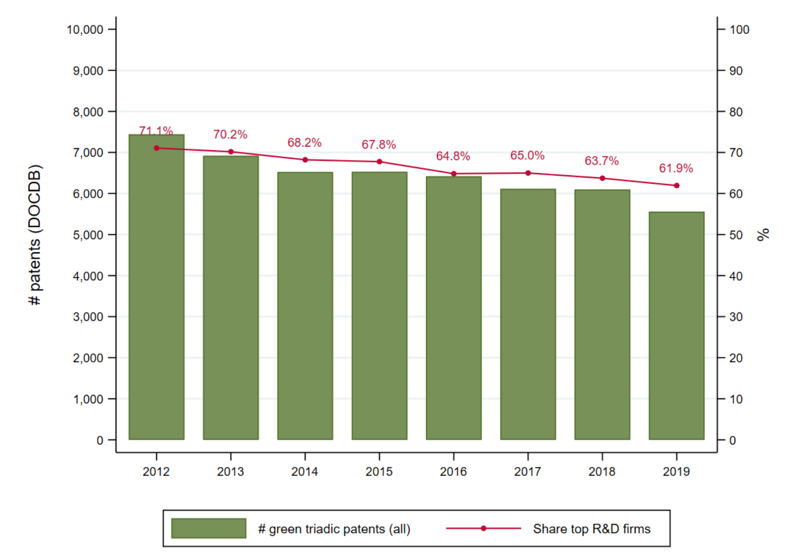
Figure 2.1: Contribution of top R&D investors to green tech development by Y02 category



Notes: This figure illustrates the proportion of green triadic patents originating from top R&D investors, by Y02 category. All green triadic patent families with earliest filing year between 2012 and 2019 are included. The red bars represent the number of green patents in each category that are filed by top R&D investors, while the labels above the bars represent the corresponding share.

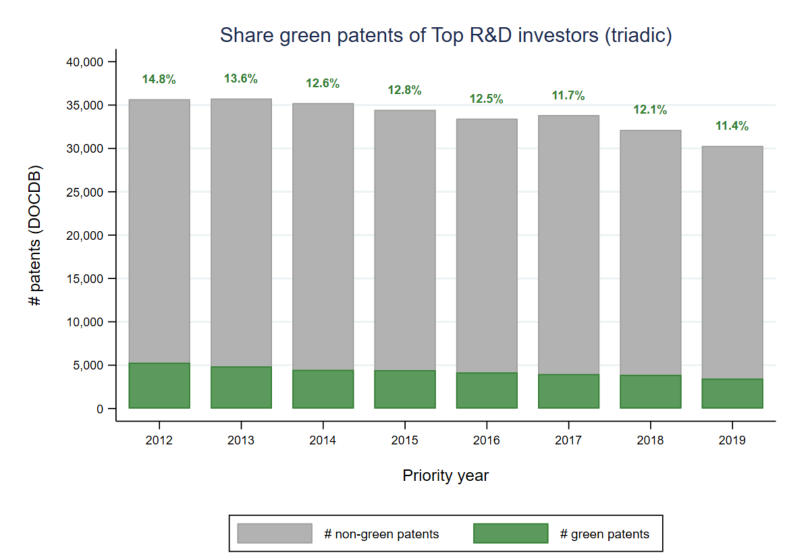
Focusing only on the sample of patents filed by top R&D investors highlights a negative trend in green technology development. Specifically, the average annual growth rate of the number of green patents filed by these firms was -5% over the period of interest. Interestingly, the number of non-green patents filed by them has also been decreasing, although at a much lower rate of only -1.7% per year. Zooming in at the technological field level, it is evident that the steepest decline in the number of green patents occurs in areas with the largest contribution from the firms of interest. Particularly, the number of green patents related to Transportation (Y02T), ICT (Y02D) and Buildings (Y02B) decreased by 8.6%, 7.5% and 7.8% per year on average. Figure 2.3 plots the evolution of the number of green and non-green patents of top R&D investors over time. The labels on top of each bar represent the share of green patents relative to all yearly filings. The share decreased from roughly 15% in 2012 to 11.4% in 2019, highlighting that top R&D investors have been focusing relatively more on developing non-green inventions in the last years.

Figure 2.2: Yearly contribution of top R&D investors to green technology development



Notes: This figure plots the total number of green triadic patent families with earliest filing year between 2012 and 2019 (left-axis) and the share that is accounted for by the top R&D investors (right-axis).

Figure 2.3: Evolution of green patenting across top R&D investors

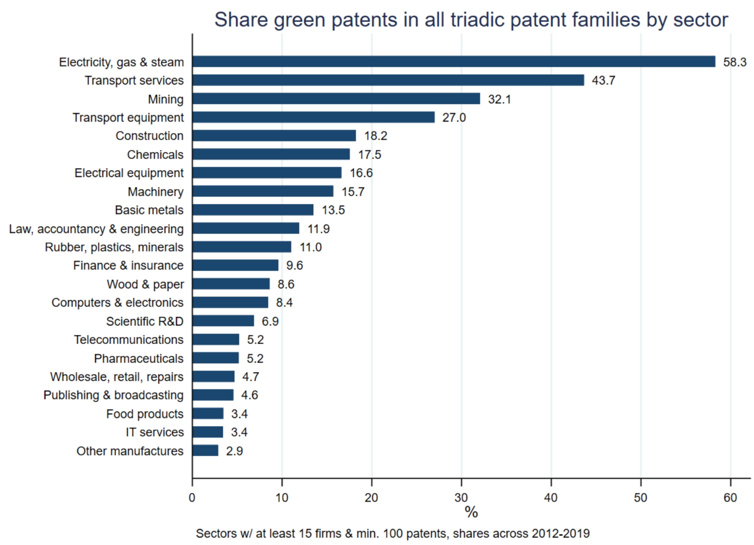


Notes: This figure plots the aggregate number of green and non-green triadic patents of top R&D investors filed over the 2012-2019 period. The label on top of each bar represents the share of green patents in all triadic patents of these firms.

2.3.2 Green Technology Specialization by Sector

I further investigate the sectors in which top R&D investors direct more of their innovation efforts toward climate change mitigation. Aggregating patent counts at the sectoral level shows that there is significant concentration in terms of green technology development. Particularly, 55% of the green patents are equally distributed among three sectors, namely 'Transport equipment', 'Chemicals' and 'Computers & Electronics', accounting for 18-19% of the patents each. The high level of concentration is interesting as it highlights the sectors where green inventions are generated on a larger scale. However, the aggregate patent count at this level is influenced by the large number of top R&D investors active in each sector as seen in Figure B.3. This points to a sector size effect rather than specialization in green technology. To address this issue, Figure 2.4 presents the share of green patents by sector, computed over the entire period of interest. The figure highlights the significant variation in green technology specialization across sectors. The relative focus on green patents appears to be the highest in the 'Electricity, gas & steam' sector, where the share is almost 60%. In contrast, green patents account for 27% of all patents in the 'Transport equipment' sector and only 8% in 'Computer & Electronics'. Smaller sectors such as 'Transport services' and 'Mining' also display a higher share of green patents of 44% and 32% respectively.

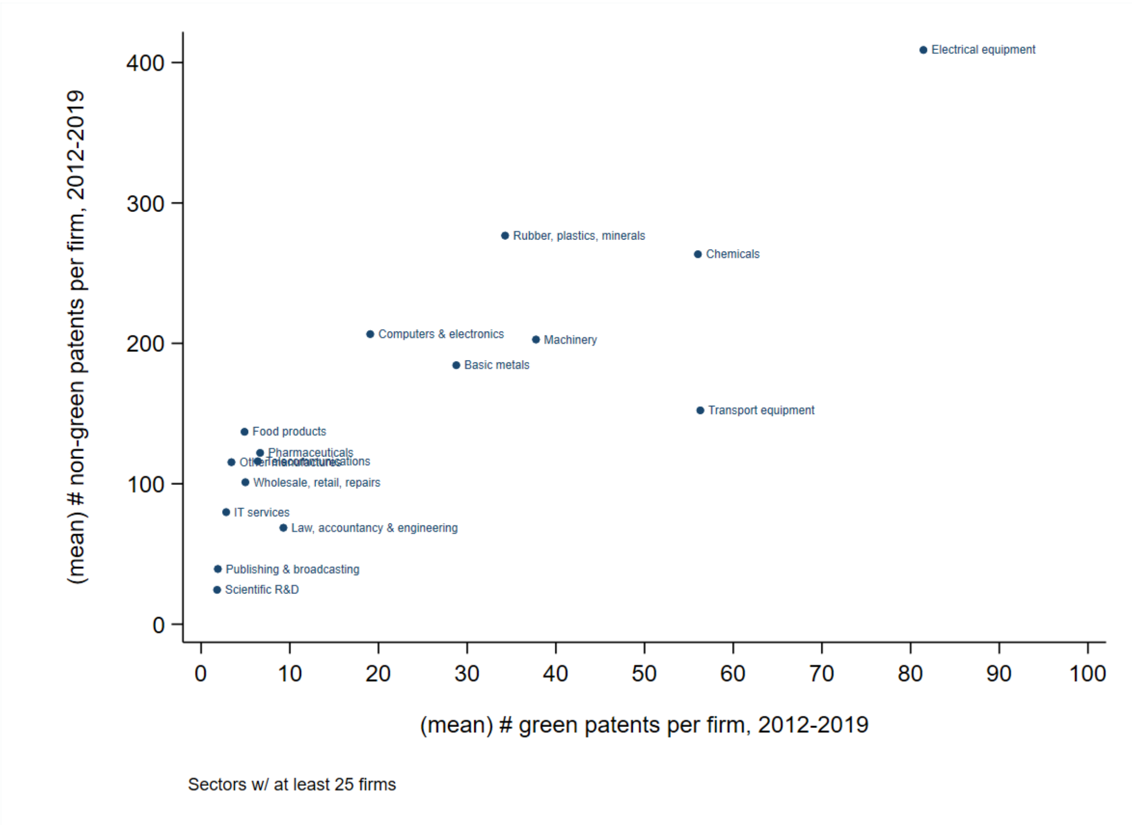
Figure 2.4: Share of green patents in all triadic patents by sector



Notes: This figure plots the share of green patents in all triadic patents filed by top R&D investors over 2012-2019, by sector. For readability, the figure includes only sectors with at least 15 firms and 100 patents. The sectoral classification is based on the aggregation of ISIC Rev.4 economic activities into 38 groups as in the COR&DIP database.

I complement these findings with an analysis of firms' behaviour with respect to both green and non-green patenting across sectors. To this end, Figure 2.5 shows the average number of

Figure 2.5: Green and non-green triadic patents per top R&D investor by sector

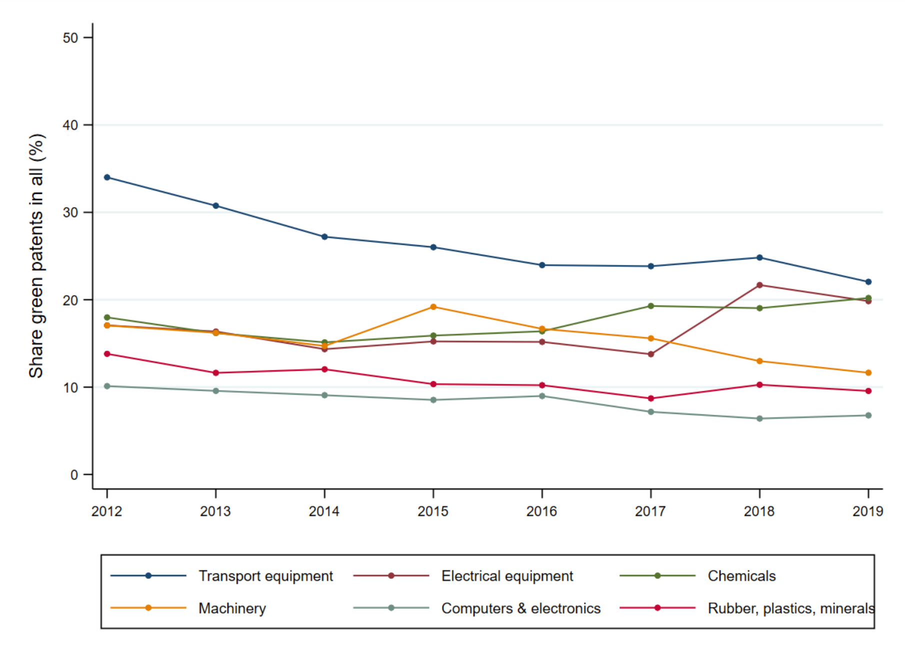


Notes: The figure plots the average number of non-green and green triadic patents per firm over the 2012-2019 period by sector. For readability, I include only sectors that have at least 25 firms present in the sample.

green (x-axis) and non-green patents (y-axis) per firm across sectors. Generally, the scatterplot identifies not only the sectors with most patent-intensive firms, but also those sectors where firms file most green patents. Thus, although it reflects the same sector-level proportions of green patents as Figure 2.4, the scatterplot also accounts for the different propensities to patent across these sectors. Notably, the sectors with the largest average number of green patents per firm are 'Electrical equipment', 'Transport equipment' and 'Chemicals'. On average, a firm in 'Electrical equipment' files 80 green and approximately 400 non-green patents between 2012 and 2019, whereas a firm in 'Transport equipment' generates around 56 green and 150 non-green patents over the same period. The bottom left corner of the scatter plot displays sectors where firms have an overall lower propensity to file patents, but also a much lower focus on developing green technologies.

The evidence presented so far suggests that certain sectors have directed significant innovation efforts towards the development of climate change mitigation technologies in recent years. However, in light of recent policy developments and the ongoing political debate on the importance of achieving climate neutrality by mid-century, it is crucial to assess the evolution

Figure 2.6: Evolution of the share of green patents in all triadic patents by sector



Notes: The figure plots the evolution of the share of green triadic patents over time at sectoral level. For readability, I include only the main sectors where top R&D investors make significant contributions to climate change mitigation.

of these efforts over the last years, namely whether they have been sustained, enhanced or rather diminished. Figure 2.6 shows how the share of green patents has evolved since 2012 for the main sectors contributing to climate change mitigation as identified above. Notably, in the case of 'Transport equipment', the share has been decreasing from 34% in 2012 to only 22% in 2019, suggesting that the firms active in this sector have been focusing relatively more on non-green inventions in the last years.[7] This is particularly surprising, given the high impact the transport sector has on global emissions. In contrast, the 'Electrical equipment' sector shows a positive development, with the share of green patents jumping to 21.7% in 2017, after averaging around 15% in the years prior to that. A steadier increase can be observed for 'Chemicals', where the share has been on an upward trend since 2015.

2.3.3 Green Technology Specialization by Inventor Country

Reaching the net-zero emissions goal requires efforts being taken globally. It is therefore important to understand the geographical distribution of green technology development and identify the countries or regions that specialize in such technologies. Probst et al. (2021) document that inventions related to climate change mitigation have been highly concentrated in

[7]Appendix Figure B.6 shows that also the absolute number of green triadic patents associated with this sector has been on a declining trend in recent years.

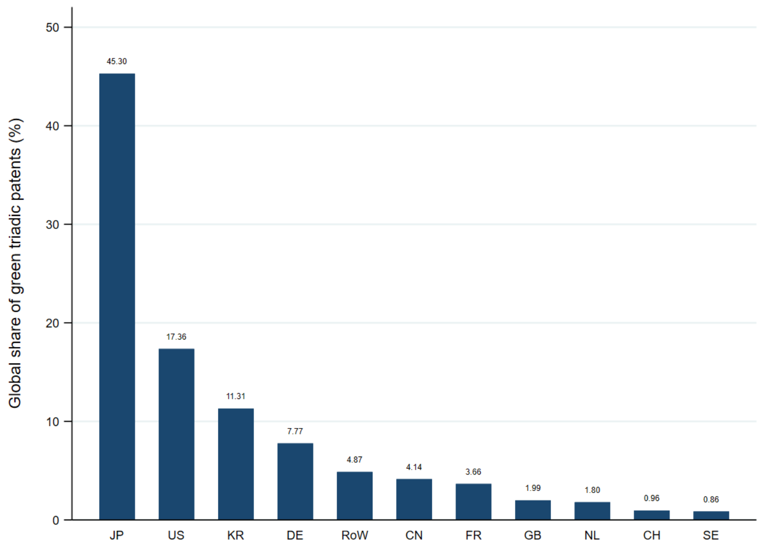
recent decades, particularly in high-income economies. In contrast to their overall perspective, I focus on the geographical dimension of top R&D investors' inventive activities and identify the locations where their green high-quality inventions are developed.

Most top R&D investors are large multinational firms with operations spanning multiple countries. Over the last decades, the increase in globalization has not only translated into the internationalization of global value chains in production, but also of innovation, with more firms offshoring R&D activities abroad (OECD, 2008). This is particularly the case for multinational firms (see e.g., Gumpert et al., 2023). Thus, assigning patents to countries based on the firms' headquarter location in order to assess the extent of green technology specialization would be misleading. Instead, one can locate the inventive activity by retrieving the location of inventors reported in patent publications. For each patent of interest, I collect information regarding inventors' addresses. In cases where inventors are located in multiple countries, I assign equal fractions to each inventor country represented in the inventor team. I compare firms' headquarter country with inventor countries in order to infer whether technology development took place domestically, or in a foreign country. 46% of green patents filed during 2012-2019 belong to Japanese firms, while only 17% originate from firms based in the United States, 11% from firms in Korea and 8% from Germany. This pattern clearly suggests that the R&D-investing Japanese firms direct significant effort towards green inventions that ultimately result in patents. An analysis of the extent of R&D offshoring in the underlying sample shows that only 5% of the green patents associated with Japanese firms have inventors located abroad. This is a very low share, especially when compared to firms based in the US (27%) and Germany (38.5%). Similarly low shares of green technology offshoring can be observed for Korean (4%) and Chinese (7%) top R&D investors.[8]

Moving on to analyze the distribution of green patents across inventor countries, I find that top R&D inventors' green inventive activity is highly concentrated geographically, with a similar pattern as above. Almost 74% of green patents originate from just three countries: Japan (45%), United States (17%) and Korea (11%). Figure 2.7 displays the share of green patents that originate in the top 10 inventor countries represented in the baseline sample. The domination of Japan in green technology development over the last years is apparent even when looking strictly at inventor location. Note however that these shares are not only influenced by the distribution of R&D investors across countries or their propensity to offshore R&D, but also by cross-country variations regarding the sectors in which these investors are active. As shown above, there is significant variation in the propensity to generally patent and to file green patents across sectors. Nevertheless, Probst et al. (2021) show a similar ranking when analyzing a wide sample of climate-related international patent families filed during 2013-2017. The similarity of results suggests that in general, the geography of green innovation among top R&D investors is representative for overall green technology development. Had

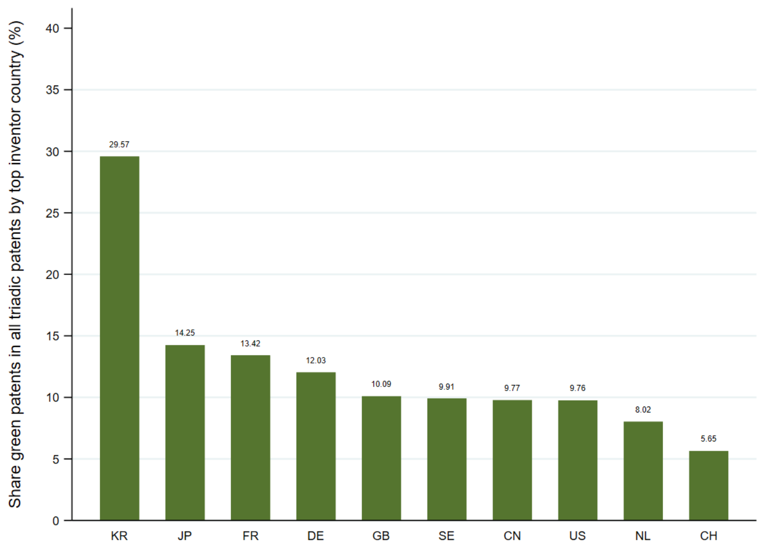
[8]Figure B.8 illustrates the extent of internationalization of R&D efforts across headquarter countries, displaying cross-country shares of green and non-green patents originating abroad.

Figure 2.7: Location of invention of green technologies



Notes: The figure includes top 10 economies where the patented green technologies of R&D investors originate from. I assign patents to countries based on the location of inventors, using fractional counts whenever there are multiple inventor countries. I compute the share of all triadic green patents that are invented in each top country over the period 2012-2019. I label the aggregated share of all other countries as RoW (rest of the world).

Figure 2.8: Share of green patents in all triadic patents by inventor country

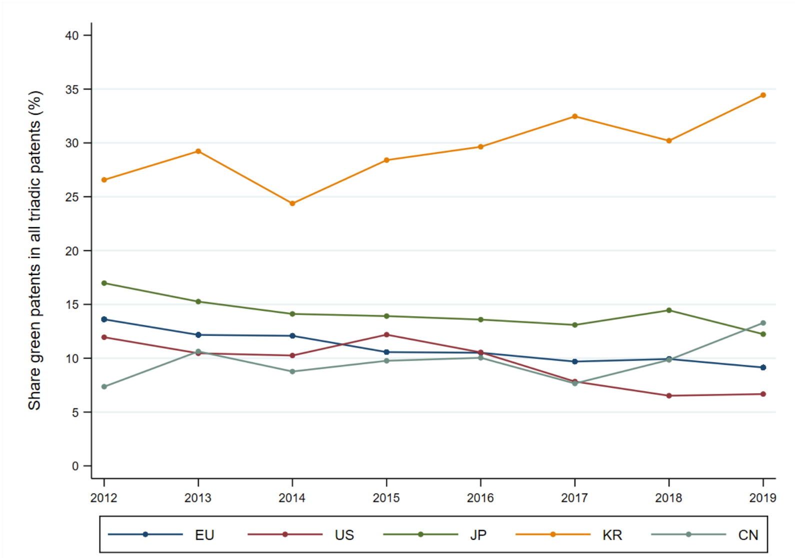


Notes: The figure includes the top 10 economies where the patented green technologies of R&D investors originate from. I assign patents to countries based on the location of inventors, using fractional counts whenever there are multiple inventor countries. For each top inventor country, I compute the share of green patents in all triadic patents invented there over the period 2012-2019.

smaller firms systematically developed green technologies in other areas, larger differences in the ranking would have been observed.

I further explore the cross-country level of specialization in climate-related inventions by computing the share of green patents invented in each country. Figure 2.8 plots the respective shares for the top inventor countries identified in the analysis above. Although most green inventions of top R&D investors originate in Japan, it seems that only 14% of all patents with inventors located in Japan are climate-related. In contrast, Korea stands out as the country with the strongest focus on green inventive activity, reflected in its share of 30%. In addition, European economies such as France, Germany and the United Kingdom exhibit a stronger focus on green technologies compared to the United States.

Figure 2.9: Evolution of the share of green patents in all triadic patents by inventor location



Notes: The figure includes top economies where the patented green technologies of R&D investors originate from. I assign patents to countries based on the location of inventors, using fractional counts whenever there are multiple inventor countries. For each top inventor country, I compute the annual share of green patents in all triadic patents invented there during 2012-2019. EU aggregates patents originating from European Union members represented in the baseline sample, including the UK.

Lastly, I investigate the extent to which countries' focus on green technologies has changed over time, as reflected in the patenting activity of top R&D investors. Intuitively, dynamics along this dimension reflect whether certain countries become more or less specialized in green technology development. I compute the share of green patents associated with each inventor country by filing year. Figure 2.9 plots the evolution of the share for selected economies: the EU (including the UK), US, Japan, Korea and China. As expected, Korea stands out as the inventor country with the highest share of green patents. Not only that, but its share has also been on an increasing trend since 2014, suggesting that over time, top R&D investors have developed relatively more high-value green technologies in Korea. In contrast, the share has been mildly decreasing for Japan and the EU region[9]. A stronger decrease can be observed for the United States where the share has almost halved over the period. China shows more volatility, the share having doubled from 2012 to 2019 when it reaches its peak at roughly 14%.

2.3.4 Green Technology Specialization by Top R&D Investor

Having examined the variation in green technology development across sectors and countries, the following analysis takes a more granular approach by focusing on the individual firm. I analyze the prevalence of green patenting among top R&D investors and identify the top firms that focus on climate-related inventions. The results so far have suggested that there is a general declining trend in the number of climate-related inventions by top R&D investors. I further investigate whether this pattern is confirmed at the firm level.

Table 2.2: Top 20 R&D investors by the number of green triadic patents

Company	Country	Sector, ISIC4	All triadic patents	Green patents	% Green	% of total green
LG Chem	KR	Chemicals	4,828	2,407	49.8	6.98
General Electric	US	Machinery	4,358	1,502	34.5	4.36
Panasonic	JP	Electrical equip.	4,751	1,244	26.2	3.61
Nissan Motor	JP	Transport equip.	2,122	985	46.4	2.86
Toyota Motor	JP	Transport equip.	2,578	976	37.9	2.83
Toyota Industries	JP	Transport equip.	2,245	850	37.8	2.46
Hitachi	JP	Electrical equip.	4,049	719	17.7	2.08
Toshiba	JP	Computers & electr.	2,949	694	23.5	2.01
Qualcomm	US	Computers & electr.	6,392	688	10.8	1.99
Mitsubishi Heavy	JP	Machinery	1,934	588	30.4	1.70
BASF	DE	Chemicals	3,082	550	17.9	1.60
Mitsubishi Electric	JP	Electrical equip.	3,668	544	14.8	1.58
Boeing	US	Transport equip.	2,273	460	20.2	1.33
LG Electronics	KR	Computers & electr.	2,618	380	14.5	1.10
Huawei Invest. & Holding Co	CN	Computers & electr.	5,069	339	6.7	0.98
Philips Lighting	NL	Electrical equip.	2,467	334	13.5	0.97
Kyocera	JP	Computers & electr.	2,361	328	13.9	0.95
Samsung Electronics	KR	Computers & electr.	2,806	323	11.5	0.94
Daikin Industries	JP	Machinery	1,589	310	19.5	0.90
Robert Bosch	DE	Machinery	1,734	280	16.2	0.81
Total (Top 20)			63,872	14,499	22.7	42.05
Total (All R&D investors)			270,787	34,483	12.7	100

Notes: This table reports the top 20 R&D investors with the highest number of green triadic patents filed between 2012 and 2019. Firms are ranked according to the total number of green triadic patents. The last column displays the share of green patents by each individual firm out of the total number generated by top R&D investors, as shown in the bottom row.

In general, it is common for these R&D-investing firms to develop green technologies, with

[9] The major European economies (Germany, France and the UK) show a generally decreasing share of green patents as shown in Figure B.9.

56% of them filing at least one higher-value green patent. Table 2.2 presents the top 20 firms ranked by their total number of green patents. These 20 companies account for 42% of all green patents filed by top R&D investors. In light of the earlier results on the overall contribution of top R&D investors to global green technology development (Figures 2.1 and 2.2), this table highlights the significant influence of just a few firms in shaping the recent green patenting landscape. The ranking is dominated by firms active in 'Computers & Electronics', 'Electrical equipment' and 'Transport equipment'. Geographically, half of the top 20 companies are based in Japan, while the US, Korea and the EU are represented by three firms each. The top five companies exhibit a strong specialization in climate-related inventions in recent years, their share of green patents ranging between 26% and 50%. LG Chem ranks first with half of its patents being classified as green. At the same time, LG Chem accounts for almost 7% of all green patents filed by top R&D investors.

The variation in the level of specialization in green technologies across firms is evident already in the ranking shown in Table 2.2. On average, 11% of a company's patents are related to green technologies. The share exceeds 30% for firms in the top decile of the distribution, while for the median R&D investor the share drops to only 4%. I further investigate how the share of green patents has developed over the last years. Figure B.10 indicates that on average, the firm-level share of green patents is also following a decreasing trend. Thus, over the course of the eight years of interest, firms' annual share of green patents decreased by roughly 2 percentage points.

Table 2.3: Green patenting of top R&D investors post Paris Agreement

	Log # green patents		Share green patents	
	(1)	(2)	(3)	(4)
Post Paris Agreement = 1	-0.067^{***}	-0.071^{***}	-0.013^{**}	-0.013^{***}
	(0.016)	(0.015)	(0.004)	(0.004)
Constant	0.755^{***}	0.763^{***}	0.115^{***}	0.115^{***}
	(0.031)	(0.010)	(0.005)	(0.002)
Sector FE	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes
N	7,999	7,923	7,999	7,923
# firms	1,293	1,217	1,293	1,217

Notes: This table reports OLS regression results investigating whether top R&D investors changed their patenting behavior after the adoption of the Paris Agreement. The dependent variables are the log number of green triadic patent (computed with offset 1) and the share of green patents in all triadic patents filed over 2012-2019. The $PostParis$ binary indicator takes the value 1 as of 2015. Standard errors clustered at the firm level are presented in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

In a regression analysis, I formally assess whether the number and the share of green patents in all filings has changed after the adoption of the 2015 Paris Agreement. Using a firm-year level dataset, I regress the number and the share of green patents on a binary indicator that takes the value 1 as of 2015. I control for time invariant sector and firm specific characteristics by adding fixed effects. The estimation results are presented in Table 2.3. The coefficient of the dummy variable $PostParisAgreement$ is negative and statistically significant across all specifications, suggesting that firms not only file fewer green patents after 2015 but also that the share of green patents is lower, on average. Thus, the decrease in the share of green patents observed in Figure B.10 holds even after controlling for sector and firm fixed effects. This illustrates that, contrary to expectations, the world's most prominent R&D investing firms have been focusing less on green technology development as reflected in their recent patenting activity.

2.4 Comparing Green and Non-green Patents of Top R&D Investors

Despite the societal need to speed up the development and diffusion of green technologies, firms with the highest innovative capacity do not seem to respond by increasing innovative activities targeting climate change mitigation. As shown in the previous section, top R&D investors have been driving green technology development, as evidenced by the high share of high value green patents they generate. However, both the number and the share of green patents in all their filings have been declining in recent years. This is surprising, particularly in light of the increasing rate with which commitments towards net-zero goals have been made ever since the adoption of the Paris Agreement in 2015.

This section provides insights into the nature of innovation pursued by top R&D investors in recent years. In particular, I compare green and non-green patents filed by these firms along a set of characteristics related to both their quality and contribution to subsequent inventions, as well as the characteristics of the knowledge base they build upon. To this end, I use a set of indicators that are based on both forward and backward patent citations to the focal patents filed by the top R&D firms during 2012-2019. In addition, I investigate whether these innovative firms have shifted their R&D efforts towards green inventions with specific characteristics in response to global sustainability commitments. Specifically, I analyze whether patenting patterns of green and non-green inventions have changed following the adoption of the Paris Agreement.

2.4.1 Patent Indicators

Forward citations. I measure patent quality using forward citations, a standard metric in the innovation literature (see e.g., Harhoff, Scherer, & Vopel, 2003). The more citations a patent receives, the more valuable it is and the more it contributes to the development of subsequent technologies. In order to account for truncation, I count forward citations over a period of three and five years since the first patent application date. I do so at the patent family level and include only USPTO citations to abstract from different citation procedures across IP offices.

Generality. The generality index captures the extent to which a focal patent is cited by patents in diverse technological fields (Trajtenberg, Henderson, & Jaffe, 1997). Intuitively, a patent scoring higher on the generality index contributes to the development of a larger number of technological fields, which implies a contribution towards a wider spectrum of later inventions. I construct the generality index following Hall, Jaffe, and Trajtenberg (2001), as: $Generality_i = 1 - \sum_j^{n_i} s_{ij}^2$, where s_{ij} is the share of forward citations received by focal patent i in technology area j, out of all n_i technology areas the forward citations belong to.

Originality. While similar to the concept of generality, the originality index focuses on a backward looking perspective and reflects the variety of knowledge sources a patent is building on (Trajtenberg, Henderson, & Jaffe, 1997; Hall, Jaffe, & Trajtenberg, 2001). It is based on the assumption that combining knowledge from various sources may lead to a more original invention. The index captures the spread in technology fields a patent draws upon by examining its backward citations, i.e. the patents referenced by the focal patent. Thus, a patent relying on a wider pool of technology fields is regarded as more original, and therefore receives a higher originality score. The OECD originality index I employ is measured as described in Squicciarini, Dernis, and Criscuolo (2013): $Originality_i = 1 - \sum_j^{n_i} s_{ij}^2$, where s_{ij} is the percentage of citations made by patent i to patent class j among all n_i patent classes comprised in the backward citations of focal patent i. The OECD Patent Quality indicators are computed for USPTO and EPO patent applications separately, taking into account the different citation patterns across IP offices. I restrict attention to the USPTO filings included in the underlying triadic patent families. Similar to Barbieri, Marzucchi, and Rizzo (2020), each patent family is assigned the maximum value of the originality index among all USPTO applications within that family.

Radicalness. The radicalness index reflects the idea that patents building on other technological fields than the one it is in should be regarded as radical (Shane, 2001). Intuitively, the more a patent references other technologies than its own, the higher it scores on the radicalness index. As described in Squicciarini, Dernis, and Criscuolo (2013), the OECD computes the index as: $Radicalness_i = \sum_j^{n_i} CT_j/n_i; IPC_{ij} \neq IPC_i$, where CT_j is the number of patent classes of cited patent j which are not present in patent i, out of n_i which represents all the IPC classes in the backward citations of patent i. Each patent family is assigned a single value of the radicalness index following the same steps as explained above.

Backward citations lag. With the aim of providing further insights into the sources of knowledge firms use in building their green and non-green technologies, I create two additional measures reflecting the age of the inventions referenced on the focal patents. First, I compute the average lag between the focal patents and all its backward citations, as the mean difference between the earliest filing date of the focal patent and the earliest filing date of the referenced patents. Thus, a patent whose references are on average 10 years old at the time of filing has an average backward citation lag of 10. In addition, I compute the share of recent citations in all backward citations of a focal patent. I consider a citation to be recent if the respective cited patent was filed at most 5 years prior to the focal patent. For example, a patent with a share of 0.2 would have 20% of its backward citations not older than 5 years old at the time of filing.

2.4.2 Methodology

I study the differences between green and non-green patents of top R&D investors by estimating variants of two main empirical specifications:

$$Y_i = \beta_0 + \beta_1 Green_i + X_i + \varepsilon_i \tag{2.1}$$

$$Y_i = \beta_0 + \beta_1 Green_i + \beta_2 PostParis_i + \beta_3 (Green \times PostParis)_i + X_i + \varepsilon_i, \tag{2.2}$$

where the dependent variable Y_i is one of the patent indicators described above, $Green_i$ is a binary indicator that takes the value 1 for green patents and X_i is a set of fixed effects I control for. Specifically, in regression 2.1, I add year, technology class (4-digit IPC class[10]) and firm fixed effects to absorb changes in patenting patterns over time, technology specific characteristics that are time invariant, and firm specific characteristics that may influence their patenting behavior irrespective of technology class or time. In a more stringent version of equation 2.1, I condition on technology class × year fixed effects, to control for differences in patent indicators across technology classes and over time within each technology class. Adding technology class dummies at such a granular level in addition to firm fixed effects ensures that I compare green and non-green patents of the same firm within identical technical fields, characterized by similar features. In regression 2.2, I investigate whether there are significant differences in the nature of innovation pursued by top R&D investors in the years after the adoption of the Paris Agreement. To this end, $PostParis_i$ is a binary indicator that takes the value 1 for patents with earliest filing year in 2015 or later. Intuitively, adding the interaction term $Green \times PostParis$ allows me to assess whether trends in green and non-green patenting of top R&D investors have diverged since 2015. This specification controls for technology class and firm fixed effects. The unit of analysis in all specifications is the patent family. The main

[10] Generally, patent filings are assigned multiple technology classes. I first identify a unique primary IPC 4-digit code for each application within the patent families of interest. Second, I identify the most frequent IPC 4-digit class at the family level. Whenever a mode is not identified, I randomly pick one of the most frequent classes.

results are based on the sample of triadic patents with earliest filing year during 2012-2019. For robustness, I provide additional estimation results using the extended sample of IP5 patent families in Appendix B.4.

2.4.3 Results

I initially take a forward-looking perspective and focus on indicators that reflect the quality of patents and their contribution to the development of subsequent inventions. Table 2.4 presents results from estimating equation 2.1 using the log number of forward citations (Columns 1-4) and the generality index (Columns 5-6) as dependent variables. I find that, on average, firms' green patents receive significantly more forward citations than their non-green patents. This implies not only that green patents are of higher quality, but also that they tend to have a greater impact on the development of subsequent technologies. In addition, Columns 5 and 6 show that green patents score higher on the generality index compared to non-green ones. This suggests that not only are green patents more cited, but they are also cited by patents belonging to a more diverse set of technological fields. Similar results are obtained when using the larger sample of IP5 patent families, as shown in Appendix Table B.3.

Table 2.4: Forward citations & Generality

	Log # citations (3y)		Log # citations (5y)		Generality	
	(1)	(2)	(3)	(4)	(5)	(6)
Green	0.049**	0.049**	0.063**	0.062**	0.024***	0.025***
	(0.018)	(0.017)	(0.020)	(0.019)	(0.003)	(0.003)
Constant	0.871***	0.872***	1.038***	1.038***	0.273***	0.272***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.000)	(0.000)
Year FE	Yes	No	Yes	No	Yes	No
Tech class FE	Yes	No	Yes	No	Yes	No
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Tech class x Year FE	No	Yes	No	Yes	No	Yes
N	255,353	254,908	255,353	254,908	217,943	217,475
# firms	1,217	1,217	1,217	1,217	1,210	1,209

Notes: This table reports OLS regression results investigating whether the green and non-green patents of top R&D investors differ in terms of quality and generality. The dependent variables are the number of forward citations received from subsequent USPTO filings within 3 and 5 years from the patent's first filing and the generality index. Citation counts are log-transformed with offset 1. A patent scoring higher on the generality index is cited by patents belonging to more diverse technological fields. I include all green triadic patents with earliest filing year during 2012-2019. Standard errors clustered at the firm level are presented in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 2.5 assesses whether green and non-green patents of top R&D investors differ across various outcomes related to the sources of knowledge they are building on. First, I focus on the originality and radicalness of patents, both of which provide insights into how patents relate

Table 2.5: Knowledge sources

	Originality		Radicalness		Avg. lag bw. citations		Share recent bw. citations	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Green	0.012^{***}	0.012^{***}	-0.005	-0.005	-0.887^{***}	-0.858^{***}	0.034^{***}	0.032^{***}
	(0.003)	(0.003)	(0.004)	(0.004)	(0.087)	(0.085)	(0.004)	(0.004)
Constant	0.777^{***}	0.777^{***}	0.368^{***}	0.367^{***}	10.252^{***}	10.240^{***}	0.425^{***}	0.426^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.011)	(0.010)	(0.000)	(0.000)
Year FE	Yes	No	Yes	No	Yes	No	Yes	No
Tech class FE	Yes	No	Yes	No	Yes	No	Yes	No
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tech class x Year FE	No	Yes	No	Yes	No	Yes	No	Yes
N	162,900	162,406	162,909	162,415	254,103	253,654	204,925	204,435
# firms	1,139	1,138	1,139	1,138	1,215	1,215	1,195	1,193

Notes: This table reports OLS regression results investigating whether the green and non-green patents of top R&D investors differ in terms of the knowledge base they draw upon. Column 1-4 focus on differences in patent's novelty, as indicated by how they score on the originality and radicalness indexes. A patent is more original if it relies on a wider pool of technology fields. A patent is more radical the more it references other technologies than its own. The dependent variables in Columns 5-8 are the patents' average lag of backward citations and the share of recent references in all backward citations. I consider a citation to be recent if the respective cited patent was filed at most 5 years prior to the focal patent. I include all green triadic patents with earliest filing year during 2012-2019. Standard errors clustered at the firm level are presented in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

to the inventions they reference. Column 1 indicates that even within firm and technology class, green patents are, on average, more original compared to non-green counterparts. This suggests that they are building on a more diverse set of technological fields. In contrast, when looking at radicalness, the coefficient of interest is close to zero and statistically insignificant. Essentially, this finding suggests that top R&D investors do not seem to rely more on knowledge from external technological domains in developing green inventions relative to non-green ones. Taken together, these results point to the fact that although firms draw on a wider spectrum of technologies for their green inventions, the references they build on in this case are not generally more distanced from the technological domain of their own inventions. Second, I analyze whether top R&D investors reference more recent technologies on their green patents relative to their non-green ones. Columns 5 and 6 show that, indeed, the average lag between green patents and their references is significantly lower compared to the average lag observed for non-green inventions. This indicates that firms rely on more state-of-the-art knowledge sources for their green inventions. A similar pattern is observed in Columns 7 and 8 where the dependent variable is the share of recent citations (with a lag of at most 5 years) in all backward cited patents. The results consistently show that a larger share of the references made on green patents are recent, suggesting rapid advancements in green technology development by top R&D investors. The robustness of the results on all indicators is confirmed when focusing on the wider sample of IP5 patent families, as shown in Table B.4.

Table 2.6: Developments after the adoption of the Paris Agreement

	(1) Log # cit. (3y)	(2) Generality	(3) Originality	(4) Radicalness	(5) Avg. lag bw. cit.	(6) Share recent bw. cit.
Green=1	0.055**	0.025***	0.012***	−0.003	−0.820***	0.043***
	(0.017)	(0.004)	(0.003)	(0.004)	(0.093)	(0.005)
Post Paris=1	−0.476***	−0.096***	−0.002	−0.026***	0.394***	0.004
	(0.016)	(0.003)	(0.002)	(0.003)	(0.055)	(0.003)
Green=1 × Post Paris=1	0.015	0.002	−0.001	−0.003	−0.130	−0.017**
	(0.027)	(0.005)	(0.004)	(0.004)	(0.107)	(0.006)
Constant	1.162***	0.326***	0.778***	0.379***	10.012***	0.423***
	(0.010)	(0.001)	(0.001)	(0.001)	(0.034)	(0.001)
Tech class FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
N	255,353	217,943	162,900	162,909	254,103	204,925
# firms	1,217	1,210	1,139	1,139	1,215	1,195

Notes: This table reports OLS regression results investigating whether top R&D investors changed the nature of their innovation after the adoption of the Paris Agreement. The $PostParis$ binary indicator takes the value 1 as of 2015. I include all green triadic patents with earliest filing year during 2012-2019. Standard errors clustered at the firm level are presented in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Lastly, I present empirical results from estimating regression 2.2 which investigates whether the nature of innovation pursued by top R&D investors changes after the adoption of the Paris Agreement. Table 2.6 reports estimation results for all indicators considered: Columns 1-2 focus on the quality and generality of patents, while Columns 3-6 look at the knowledge sourcing behavior of firms as reflected in the backward citations of their patents. For the forward-looking perspective, I concentrate on citation counts compiled over the short period of 3 years which should not be affected by truncation given the bibliographic data used. In Columns 1-2, the coefficient associated with the $PostParis$ binary indicator is negative and statistically significant, suggesting that, on average, patents filed after 2015 receive fewer citations and are cited by a less diverse set of technological domains. The main coefficient of interest, that of the interaction term $Green \times PostParis$, is positive but statistically insignificant, suggesting that after the adoption of the Paris Agreement, green patents do not exhibit different patterns in terms of their contribution to subsequent inventions.

Shifting attention to the knowledge base firms rely on, Columns 3-6 report results from estimating the same specification using the originality, radicalness, and backward citation lag measures as dependent variable. As observed in Column 3, patents of top R&D investors do not seem to score higher in originality after 2015. In addition, there is no significant difference in how original their green patents are relative to the beginning of the period. Interestingly, patents filed after the adoption of the Paris Agreement are, on average, less radical, as evidenced by the negative and statistically significant coefficient associated with the $PostParis$

dummy in column 4. At the same time, there are no diverging trends in radicalness between the green and non-green patents after the Paris agreement, as suggested by the statistically insignificant coefficient of the interaction term $Green \times PostParis$. Results in Column 5 indicate that the average lag between patents and their references is higher after 2015 across patent types, pointing to the fact that recent inventions of top R&D investors rely on older knowledge sources, on average. However, since the coefficient of the interaction term in this specification is also statistically insignificant, the average lag of backward citations of green and non-green patent is not diverging after the adoption of the Paris Agreement. Surprisingly, the pattern of backward citations on green and non-green patents does seem to deviate after 2015 when we consider the share of recent references. Thus, after 2015, green patents rely less on recent knowledge sources compared to before. To sum up, I do not find consistent evidence suggesting that top R&D investors have changed the nature of their green innovation as a response to the adoption of the Paris Agreement and increased net-zero commitments. Their recent green patents do not seem to be cited more, nor broader and they do not exhibit greater originality or radicalness. Surprisingly, they rely on older knowledge sources compared to green patents filed before 2015. Results are confirmed in the robustness analysis using the IP5 patent sample, as shown in Table B.5.

2.5 Discussion and Conclusion

This paper explores recent trends in green technology development by the world's leading R&D investing firms. The analysis highlights their significant contribution to this field, with these firms accounting for 67% of all high-value green patents filed between 2012 and 2019. Most importantly, despite global commitments to the net-zero goal following the adoption of the Paris Agreement, there has been a decline in their green patenting activity. Not only has the number of green patents decreased over time, but also their proportion relative to all patents. This pattern points to the fact that, surprisingly, firms have shifted their focus towards the development of non-green technologies in recent years. Furthermore, although green patents seem to receive more citations, be more original and rely on more recent sources of knowledge compared to firms' non-green patents, I do not find consistent evidence that the nature of firms' patents has changed following the adoption of the Paris Agreement. Specifically, the analysis shows that firms' recent green patents do not seem to be cited more, nor broader, they do not exhibit greater originality or radicalness and they surprisingly rely on older knowledge sources compared to those filed before 2015.

An overall decreasing trend in the number of green patents was also documented by Probst et al. (2021), who focus on the total number of green patents filed in at least two jurisdiction over the 2013-2017 period. They argue that this negative development is likely due to declining fossil fuel prices, low carbon prices and increased maturity of certain technologies in this space, such as solar photovoltaics. Analyzing recent patenting trends in the energy sector, Popp et al. (2021) suggests that an additional explanation for the decline in green innovation could be weakened or uncertain regulation that can hamper investments and incentives to innovate.[11] In line with the technological maturity argument, they also raise awareness about the possibility that research in certain fields has reached diminishing returns, such that advancements become more incremental in nature and less likely to be patented. Linking back to the firm-level perspective taken in the current study, further work could delve deeper into the type of innovation top R&D investors have been focusing on, whether their inventions are more likely to be efficiency-enhancing and therefore incremental, or whether they are exploring completely new technological domains. The puzzling result that their recent green patents rely on older sources of knowledge points to a change in their R&D activity that deserves further exploration.

The slowdown in green technology development by top R&D investors documented in this paper raises additional questions not only for researchers but also policymakers. A natural follow-up question is whether other types of firms react differently to the current global climate crisis. Comprehensive evidence on recent trends in market entry and the contribution of smaller firms in developing climate-related technologies is currently lacking (OECD, 2021b). Such firms are known for generating more radical and disruptive innovation, challenging existing business models and shaping new markets (Henderson, 1993; Baumol, 2005; OECD, 2011). Novel technological solutions are needed in order to reach climate objectives, therefore exploring whether smaller firms are currently picking up the pace in green technology development is particularly relevant. However, it is evident that the role of top R&D performers cannot be substituted by small firms. Beyond their overall capacity to innovate, these top performers possess global networks that are instrumental in scaling and deploying technological advancements quickly, irrespective of their nature. Any efficiency-enhancing innovation brought to a global scale would contribute significantly to emissions reductions.

Reaching the ambitious net-zero goal by mid-century requires not only the development of new green technologies but also their worldwide deployment. Evidence presented in this paper complements results by Probst et al. (2021), showing that technology development is highly concentrated in developed economies. Fostering diffusion across low- and middle-income countries is critical for curbing the damaging effects of climate change (OECD, 2011). Thus, future work can focus on uncovering patterns in the diffusion of green technologies and explore not only the geographical dimension of diffusion but also whether there is any evidence of speeding-up in recent years. Patent citation data may be used in future studies to assess how the knowledge embedded in patented technologies is diffusing over time and across borders.

[11] Berestycki et al. (2022) measure climate policy uncertainty based on newspaper coverage frequency across 12 OECD Member Countries during 1990-2018. They document that higher levels of policy uncertainty are associated with significant decreases in firms' investment.

To conclude, this paper provides valuable insights into the evolving landscape of green technology development in a decade where climate change mitigation has turned into a topic of intense public and political attention. Ultimately, the negative trends highlighted in this paper suggest that the private sector innovation mechanism is not operating at its full potential, which goes against wide-spread perceptions in policy circles and the public at large. With global policy goals becoming more ambitious, these findings raise concerns and draw attention to the need for a deeper understanding of the factors driving firms' R&D and innovation decisions in the realm of green technology.

3

Government R&D Support for Young Innovative Companies

Evidence from France

3.1 Introduction

Young innovative companies play an important role in generating radical innovations, shaping new markets and fostering job creation, having the potential to contribute to technological change and overall economic growth (Henderson, 1993; Schneider & Veugelers, 2010; Haltiwanger, Jarmin, & Miranda, 2013). Thus, policymakers around the world have been using a variety of mechanisms to support new innovative ventures, usually with the objective of stimulating their growth and R&D investments and improving their chances of obtaining venture capital finance.

The financing of such young, growth-oriented firms, also known as startups, is subject to various forms of market failure. One main reason for governments supporting firms' R&D endeav-

*This chapter is based on joint work with Laurie Ciaramella and Dietmar Harhoff. Access to some of the data used in this work was carried out in secure environments provided by CASD – Centre d'Accès Sécurisé aux Données (Ref. 10.34724/CASD).

ours is their inability to fully internalize the spillovers they create by generating new knowledge. The wedge between the private returns from investment in innovative activities and the associated social benefits leads to an underprovision of R&D (Nelson, 1959; Arrow, 1962). Furthermore, young innovative companies face significant challenges in securing access to external capital due to financial frictions (Hall & Lerner, 2010). Information asymmetries prevent potential investors from easily assessing the likelihood of success of ventures' inventive activities and the value of their investments. Financial constraints may in turn hinder startups' growth rate and R&D trajectories in the form of delayed or abandoned projects.

This study provides new evidence on the effectiveness of government R&D support schemes aimed at young innovative ventures. Specifically, our analysis focuses on a particular type of instrument that has been largely overlooked in previous studies: a scheme entailing a reduction in R&D labor cost, achieved through exemptions from social security contributions for R&D personnel. Public support programs aiming to address the aforementioned challenges have been receiving significant attention in the last decades. R&D subsidies, grants as well as loans have been shown to improve startup performance in multiple settings (e.g., Lerner, 1999; Howell, 2017; Zhao & Ziedonis, 2020; Santoleri et al., 2022). Compared to a standard R&D grant program allocated on a competitive basis, such a scheme might allow for an easier and straightforward implementation, lower bureaucratic burden as well as the provision of more long-lasting support. At the same time, as reported below, the cost of R&D labor is significant for young innovative companies, highlighting the relevance of the tool for supporting firms in the early stage of their development.

In the context of the French *Jeune Entreprise Innovante* (JEI) scheme, we examine how a reduction in R&D labor cost applied in the early stage of startups' development can influence their performance. Since 2004, the JEI scheme supports R&D efforts within the French startup environment by providing young and innovative ventures with fiscal advantages. Specifically, under certain conditions, young firms with sufficiently high R&D expenditures are temporarily granted reductions in social security contributions for their R&D-related employees. We leverage access to French administrative data and identify all young innovative companies that benefit from the JEI scheme since its introduction. Hence, we first document the uptake of the JEI status in France. In order to analyze how early stage benefits influence firm performance, we limit attention to a subset of JEI firms founded between 2006-2012 and follow them over an eight to twelve-year time period. We construct a group of similar young innovative companies that were eligible for the JEI status in their initial phase, but never benefited from the scheme. By comparing outcomes of JEI and non-JEI firms after their fourth year, we aim to shed light on how early-stage JEI benefits affect firms' survival, likelihood of acquisition, innovation output, and overall performance concerning employment and sales.

We provide evidence suggesting that JEI firms are more successful compared to similar young innovative companies not benefiting from the scheme. In particular, they are more likely to stay active by the end of their initial eight years since foundation, thus having a higher probability of survival. Additionally, JEI firms are more likely to get acquired within the first twelve years. Such a "trade sale" is an important outcome measure, since being taken over by a larger firm usually generates substantial returns for founders, inventors and the state via taxation. Our second set of results explores firms' innovation output, proxied by their patenting activity. We find that JEI firms are not only more likely to file a patent at ages five to eight compared to non-JEI firms, but they are also filing more patents. Finally, we show that during the same development stage, there are significant differences between the workforce of JEI and similar young innovative companies. We observe not only significant variations in overall employment, but also differences among employee type. Specifically, in addition to opting for more R&D-related hires, JEI firms are expanding their workforce by also recruiting more non-R&D employees. We interpret this finding as evidence of JEIs' maintained research emphasis, complemented by a focus on scaling up their operations. Although generating lower revenues initially, JEI beneficiaries seem to catch up to the non-JEI firms within the time frame we are investigating.

This paper has multiple contributions. First, we analyze a rare public R&D support instrument that, to the best of our knowledge, has been understudied with respect to diverse aspects of startup performance. Since 2004, the French government has allocated over €2 billion to support more than 10,000 young innovative firms through reductions in their social security contributions. This substantial amount highlights the extensive utilization of the scheme in the French economy. It also demonstrates how important it is for French policymakers to understand the effects of the JEI scheme in detail. Additionally, the relevance of our analysis goes beyond French borders, as other countries such as Sweden, Spain, Hungary and Turkey have introduced a similar R&D support tool since 2004.[1] The limited number of studies assessing the French JEI support scheme concentrate on labor outcomes (Hallépée & Houlou Garcia, 2012; Gautier & Wolff, 2020; Quantin, Bunel, & Lenoir, 2021). There is consensus regarding a positive employment effect of the JEI scheme. Additionally, Gautier and Wolff (2020) find that JEI beneficiaries experience an increase in hours worked and a modest increase in the average hourly wage, the latter being due to a change in the composition of the workforce rather than an increase in individual wages. More recently, Quantin, Bunel, and Lenoir (2021) conduct an *ex post* evaluation of the scheme, obtaining a small but significant effect on total and R&D-related employment and no effect on wages. In contrast, we extend the scope of the analysis by including other startup performance outcomes touching upon firm success and innovation output. Thus, we further contribute to the study of the JEI scheme by building a uniquely rich firm-level dataset that combines three main sources: various administrative datasets based on French firms' tax declarations, patent information from the French Intellectual Property Office (INPI) combined with data from PATSTAT, and firm acquisition information from Crunchbase.

[1] See OECD (2021a) for an overview of R&D tax relief measures applied by selected OECD and non-OECD member states.

Second, our paper contributes to the broader literature on the effectiveness of public R&D support programs targeting young firms. A small yet growing strand of research has investigated the effect of direct governmental support in the form of R&D subsidies specifically targeting startups. Howell (2017) and Santoleri et al. (2022) find that early-stage grants, such as those from the US Small Business Innovation Research (SBIR) program and the European SME Instrument, positively impact firms firms' patenting, venture capital attainment, survival and commercial success. Santoleri et al. (2022) further show that grants reduce firms' financial frictions through funding, rather than a certification effect. Other evidence suggesting that such programs benefit young firms has been found for Italy (Bronzini & Piselli, 2016), Israel (Conti, 2018), and Sweden (Söderblom et al., 2015). Additionally, several governments have introduced debt-based instruments such as subsidized loans in their effort to support small firms. Zhao and Ziedonis (2020) examine the impact of an R&D loan program for technology startups in Michigan, finding that loan recipients experienced increased survival rates and improved access to venture capital, particularly for very young firms. Meanwhile, Hottenrott and Richstein (2020) focus on knowledge-intensive startups in Germany, comparing the effects of grants and subsidized loans. They discover that both types of programs stimulate firm growth, but only grants lead to increased R&D investments.

The remainder of this paper is structured as follows: Section 3.2 provides institutional details on the French JEI scheme, together with overall descriptive findings on the uptake of the JEI status since its introduction. Section 3.3 elaborates on our empirical approach, introducing the data, construction of baseline sample and outcome variables. Results are presented in Section 3.4. We provide a discussion and conclude in Section 3.5.

3.2 Institutional Background

In 2004, the French government introduced the *Jeune Entreprise Innovante* (JEI) scheme with the aim to incentivize R&D and innovation activities of French young firms (Moutaabbid, 2016). The JEI status grants firms a number of fiscal benefits, primarily exempting them from paying social security contributions for staff involved in carrying out R&D projects. The scheme is a horizontal type of policy tool, as it applies to all young innovative companies, irrespective of industry, technology or growth potential. In order to be eligible for the status, firms must meet a number of criteria, including a specific R&D expenditure threshold.

Specifically, in order to qualify as a young innovative company subject to the JEI scheme, firms must meet five eligibility criteria.[2] First, they need to be a small or medium enterprise (SME), as defined by the European Union: they should have less than 250 employees, a turnover not

[2]The JEI status has been established by Article 13 of the Finance law for 2004 n° 2003-1311 from December 30, 2003. The eligibility criteria for the JEI status are defined by Article 44 sexies-0 A of the General Fiscal Code.

Table 3.1: The evolution of the JEI benefits

A. Social Security Exemption Scheme

	2004-2010	2011	2012-2013	2014-2020
- Age 1-4	100%	100%	100%	100%
- Age 5	100%	75%	80%	100%
- Age 6	100%	50%	70%	100%
- Age 7	100%	30%	60%	100%
- Age 8	100%	10%	50%	100%
Ceilings				
- by employee	none	4.5 × minimum wage	4.5 × minimum wage	4.5 × minimum wage
- by establishment	none	3 × annual social security ceiling	5 × annual social security ceiling	5 × annual social security ceiling

B. Other Fiscal Benefits

B1. Exemption or reduction of corporate tax / income tax

	2004-2010	2011	2012-2013	2014-2020
- year 1	100%	100%	100%	100%
- year 2	100%	100%	50%	50%
- year 3	100%	100%	None	None
- year 4	50%	50%	None	None
- year 5	50%	50%	None	None

B2. Exemption from taxes on capital gains from the sale of JEI shares

	2004-2010	2011	2012-2013	2014-2020
- 1-3 years ownership	0% of IR + PS	0% of IR + PS	0% of IR + PS	50% of IR + PS
- 4-7 years ownership	0% of IR + PS	0% of IR + PS	0% of IR + PS	35% of IR + PS
- ≥ 8 years ownership	0% of IR + PS	0% of IR + PS	0% of IR + PS	15% of IR + PS

Notes: This table presents the evolution of the JEI benefits since its introduction in 2004. Year in Panel B refers to the beneficiary financial year or tax period. In Panel C, IR refers to the personal income tax, which is progressive and therefore varies across individuals. PS refers to social deductions, which are fixed.

exceeding €50 million or a balance sheet total not higher than €43 million.[3] Firms must be less than 8 years old and carry out a new activity, i.e. they should not be the result of restructuring operations, part of a merger, expansion, etc. Additionally, firms must be independent, such that the majority of their capital is not under the ownership of a different entity, with few exceptions such as other young innovative companies or research organizations.[4] Lastly, firms'

[3]Prior to 2008, the SME criteria included the same employee count limit, a turnover of up to €40 or a balance sheet total of at most €27 million.

[4]According to the scheme, a firm is considered independent if at least 50% of its capital is held by one of the following: individuals, other young innovative companies, scientific associations or foundations of public interest, investment structures such as venture capital firms and private equity funds, and public research and teaching institutions.

annual R&D expenditures must represent at least 15% of their annual tax deductible expenses in order for them to be eligible for the JEI status each year. Firms lose their JEI status if they do not meet one of the conditions in a given year. However, they can regain their status if conditions are met in subsequent years.[5]

Table 3.1 summarizes the fiscal advantages obtained by JEI beneficiaries over time. Most importantly, young innovative companies are exempted from paying employers' social security contributions for the high-skilled employees that are contributing to R&D projects.[6] This includes engineers, technicians, R&D project managers, lawyers dealing with industrial protection and technology agreements, and employees in charge of pre-competitive testing. Fiscal benefits include a full exemption or reduction in corporate tax in the first beneficiary period(s), as well as exemptions from tax on capital gains from the sale of shares.[7] Innovative startups are known to experience losses in the early years of their existence, suggesting that among all types of benefits, it is the reduction in social security contributions for R&D related staff that is most likely to meaningfully impact JEI beneficiaries. In 2011, a policy reform introduced a gradual exit from the scheme, such that firms closer to the age of eight were granted a decreased rate of exemption relative to younger firms. Furthermore, the policy change enforced a cap on the amount of exemptions that can be granted per employee and per establishment. Thus, since 2011, only the share of an employee's wage that is below 4.5 times the minimum wage is subject to the exemption. At the establishment level, the amount of exemptions granted should not exceed five times the annual social security ceiling, valued at approximated €40,500 in 2019. The change in exemption rate was revoked after only a few years, a full exemption being applied ever since 2014.

All young firms that meet the eligibility criteria may use the JEI status and benefit from the scheme. In contrast to standard R&D subsidies or grants, the status is not assigned on a competitive basis, meaning that firms are not required to go through a formal application procedure followed by an evaluation. Instead, firms may self-declare as JEI on their tax declaration. Doing so allows them to directly apply reductions to their social security contribution statements filed with the tax authorities. Although firms have the option to enquire whether they are eligible for the status, this is not a prerequisite for being granted the exemptions. However, beneficiaries must be able to justify their JEI eligibility in case of an *ex post* investigation.[8] Being granted the JEI status does not prohibit firms from being eligible to the general French R&D tax credit (CIR) or other support mechanisms. In fact, most JEI participants benefit from

[5]Note that during 2004-2007 the loss of the JEI status was irreversible. Only after 2007 firms could regain the status if conditions were met in the future.

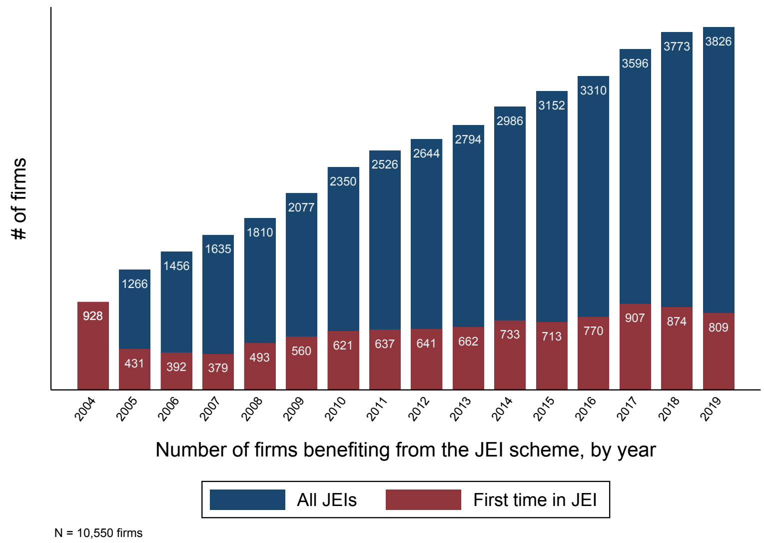
[6]Article 131 of the Finance law for 2004 n° 2003-1311 from December 30, 2003 defines the social security exemptions granted by the JEI scheme.

[7]Some municipalities also grant JEI beneficiaries exemptions from local property tax.

[8]Additionally, in order to receive the benefits they need to have been carrying out their reporting and payment obligations with Urssaf, the agency responsible of collecting all social security contributions in France. No other checks are conducted by tax authorities or Urssaf before firms' claiming the JEI status and reporting reduced social security contributions.

both support schemes, as shown in Figure C.6.

Figure 3.1: Firms benefiting from the JEI scheme



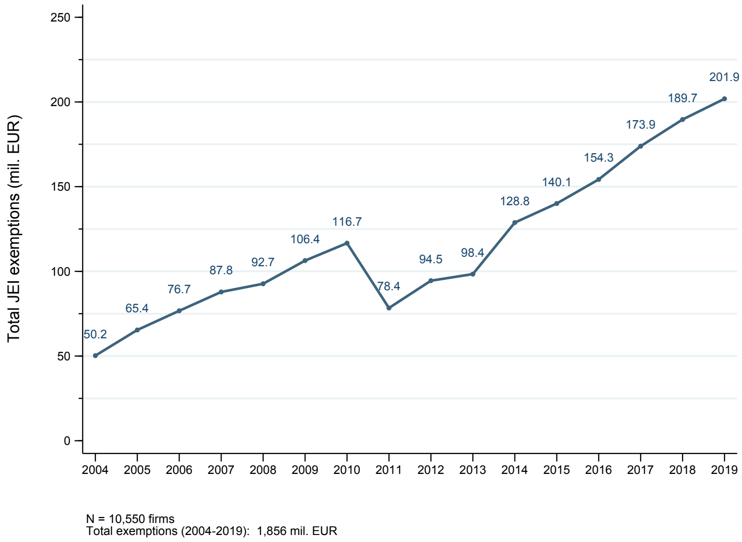
Notes: We plot the number of firms with the JEI status by year (entire bar) and distinguish those firms that enter the JEI scheme for the first time (red segment of the bar). From the total number of firms comprised in the JEI dataset, we restrict our attention to those that receive benefits at ages one to eight, in line with the eligibility criteria. We additionally remove firms for which we cannot establish the date of foundation, such that the age at entry and exit into the scheme cannot be identified.

During the 2004-2019 period, more than 10,000 firms have benefited from the JEI scheme, as illustrated in Figure 3.1. In 2019, the last year we record, there were approximately 3,800 beneficiaries, 800 of which being new entrants.[9] Since its introduction, the number of firms joining the scheme each year has been steadily increasing. This pattern could be attributed to firms gradually learning about the scheme or to the overall development of the French startup ecosystem. Figure 3.2 plots the total amount of exemptions that the French government has granted the 10,550 beneficiaries we track since 2004. As expected, the annual amount of JEI-related exemptions has risen consistently, from € 50 million in 2004 to more than € 200 million in 2019, marking a more than fourfold increase. Since the introduction of the status, the total amount of exemptions granted to young innovative companies exceeds € 2 billion, underlining the extensive use of the scheme as well as the relevance of understanding its effects.

JEI beneficiaries tend to enter the scheme in the early stage of their development. Appendix Figure C.3 suggests that more than 50% of the firms claim the JEI status within the first two years since their establishment. On average, a JEI firm maintains its status for approximately 4.7 years. Panel A in Appendix Figure C.2 shows the distribution of JEI beneficiaries by the number of years of granted benefits. The figure suggests that, even when considering only the subset of firms observed for the entire eight-year period, there is significant variation in firms'

[9]Figure C.1 presents the distribution of JEI beneficiaries by founding year.

Figure 3.2: Total granted exemptions



Notes: We plot the annual amount of JEI social security exemptions granted to the 10,550 beneficiaries we observe over the 2004-2019 period.

duration of participation in the JEI scheme. Panel B further shows that the average participation length varies only modestly across cohorts. In general, firms may lose the status prior to the end of their eighth year since foundation for multiple reasons. These reasons may include no longer meeting the R&D criterion, ceasing operations, exceeding the SME threshold, or losing independence. Appendix Figure C.4 illustrates the average amount of exemptions a firms in each cohort received. Thus, a JEI beneficiary, observed over a full eight-year period, received an average exemption of roughly €200,000 in social security contributions. As shown in Appendix Table C.3, a JEI beneficiary in its first year of existence would be granted an average exemption of €11,000, while a firm of age eight, an average exemption of approximately €69,000. In a given fiscal year, the exemption granted by the JEI scheme would amount to roughly 16% of a firms' total cost related to wages.[10]

3.3 Empirical Strategy

This paper analyzes how the benefits provided by the JEI scheme impact firm performance. To do so, we build a sample of firms that benefit from the scheme (JEI firms) and firms that were eligible for the JEI status but never benefited from it (non-JEI firms). In constructing the baseline sample, we limit our attention to the first four years of a firm's existence. Thus, we consider only JEI firms that have entered the scheme at the latest at age four. Similarly,

[10]The mean value is computed based on a subset of firms where wage information is available. We retrieve total wage expenses and social security charges from firms' income statement information in FARE (2009-2019).

we identify eligible non-JEI firms by checking their eligibility within their first four years of existence. In doing so, we are able to provide evidence on the immediate, early impact of a reduction in R&D labor cost during the first years of innovative firms' development. We compare JEI and non-JEI firm outcomes after the age of four and provide insights into how such early-stage benefits can influence firms' survival, acquisition probability, innovation activity, and early performance.

3.3.1 Data

This subsection elaborates on the main data sources used in building our baseline sample. For the purpose of our analysis, we undertake a major data construction effort and build a novel dataset on French young companies by combining multiple high-quality administrative firm-level datasets provided by the French Secure Data Access Center (CASD). We rely on information regarding the JEI status, firms' balance sheet, ownership structure, employment, as well as beneficiaries of the French R&D tax credit. We complement these data with comprehensive patent-level information from the French Intellectual Property Office (INPI) and PATSTAT, the worldwide patent database maintained by the European Patent Office. Additionally, we retrieve firm acquisition data from the online platform Crunchbase.

JEI. The French Central Agency for Social Security Organisms (ACOSS) assembles a panel with yearly information on young firms that declare a reduction in social security contributions based on the JEI status. We observe firms that benefited from the JEI scheme over the period 2004-2019. The dataset contains establishment-level information regarding the number of employees, amount of social security contributions paid, amount of social security exemptions received, amount of other type of exemptions received. The panel dimension of the JEI dataset allows us to identify the length of firms' JEI status. For our analysis, we aggregate all variables at the firm level, using the SIREN number (the French firm unique identifier).

FICUS/FARE. We obtain balance sheet information for the population of French firms from the FICUS/FARE datasets which comprise data from firms' tax declarations. We retrieve firms' founding date, closure date (if available), number of employees, turnover, total balance sheet, total expenses, economic activity (5-digit NAF classification [11]) and legal form. FICUS/FARE are instrumental in building our baseline sample, as they allow us to establish firm age, firms' SME eligibility as well as R&D eligibility for the JEI status. In addition, we use this data to build a number of outcome variables as explained below.

Enquête LIFI & LIFI. Firms can only be eligible for the JEI status if they are independent, i.e. they are not controlled by a different entity. In order to assess the independence criterion, it

[11] *Nomenclature d'activités française.* The most recent version of NAF is consistent with the European Union NACE Rev.2 classification.

is necessary to determine whether the firms of interest have financial links with other entities, either French or foreign firms. Therefore, we build a panel dataset that includes all financial links between our focal firms and other firms in France or abroad. Furthermore, we obtain information on whether there is a group head that controls the focal firms each year. We construct these panels using the Enquête LIFI and LIFI data sources, which contain the structure of corporate groups active in France. Not only does these data allow us to check firms' independence, but also provides us with a way to track changes in ownership and firm acquisitions, one of our main outcome variables of interest. Moreover, we are able to identify and remove French subsidiaries of foreign multinational firms that are outside the scope of the underlying analysis.

GECIR. One of the main limitations in retrieving the population of French young firms that would have been eligible for the JEI status is imposed by the availability of firm R&D expenditure data. We overcome this challenge by focusing on firms that additionally benefit from the French R&D tax credit (CIR). Such firms are required to declare their R&D expenditures on their tax declaration. The GECIR dataset, produced by the French Ministry of Finance, contains the R&D related information reported on firms' tax declaration since 2008. We combine R&D expenditure information from GECIR with total expenditures data from FARE and calculate the R&D ratio required by the JEI R&D eligibility criterion. We consider the criterion to be met for any firm-year combination where the R&D ratio is higher than 15%. Due to data availability, we can only compute the R&D ratio for the period 2009-2019. We validate this measure by calculating the ratio for the subsample of JEI firms where both variables are non-missing, and find that more than 84% have a ratio above 15% in the years in which they receive the benefits.

Other sources. Finally, we use additional sources for building some of our main outcomes of interest. We retrieve patent-level information by combining data from INPI and PATSTAT over the period 2000-2017. In our analysis, we focus on patent families that contain French patent applications. Unlike PATSTAT, INPI patent data allows us to directly link patent applicants and our firms of interest using the SIREN number.[12] In a second step, we collect all other patent applications within the same patent families from PATSTAT. We rely on this two-step approach in order to circumvent the absence of a cross-walk between PATSTAT patent applicants and firm SIREN numbers. We distinguish patents families that include a European Patent application and refer to them as EP patents. Moreover, we obtain firm acquisition data from the online venture capital platform Crunchbase. We initially extract all acquisition events that involve a French acquiree and match them with our baseline sample using firm names as explained in Appendix C.1. Lastly, we use employer-employee administrative data from DADS (Déclaration

[12]Since we are analyzing small French firms, it is likely that their initial inventions would predominantly target the domestic market. Consequently, we anticipate their first patent applications to be filed at the INPI, the patent office closest to their location.

Annuelle de Données Sociales) in order to generate a proxy for the R&D personnel eligible for exemptions. We rely on the *DADS Salariés* tables that include the primary job of each employee. We categorize employees as R&D related using the detailed nomenclature of occupations and socio-professional categories (INSEE, 2003) available in the dataset. We consider engineers, technicians and project managers in tech-related fields. Appendix Table C.2 contains the full list of occupations included in our R&D personnel measure. For robustness, we generate an additional measure using the tech-worker definition of Harrigan, Reshef, and Toubal (2023).

3.3.2 Sample Construction and Summary Statistics

We construct a sample of firms that includes those benefiting from the JEI scheme and firms eligible for the scheme that never used the JEI status. Our focus is on firms founded between 2006 and 2012 for two reasons. First, we begin with firms founded in 2006 to ensure that all eligibility criteria can be evaluated given the available data. Second, we include firms founded up to 2012 to allow for a full eight-year observation period for all firms. From the population of firms founded during this period, we gradually remove firms based on a sequence of steps described below. We generally condition on firms receiving R&D tax credits at least once within the first four years since their establishment. This rule is essential for calculating the R&D ratio whose evaluation is necessary based on the JEI eligibility criteria. This step further ensures that JEI and non-JEI firms are comparable with respect to their engagement in R&D activities. Simultaneously, by focusing solely on firms benefiting from the R&D tax credit in their early stages, we address the potential endogeneity concern. Additionally, we only consider JEI and non-JEI firms that are registered as companies, as indicated by their legal form.[13]

JEI firms. Besides the above conditions generally applied to all firms, we take the following additional steps in building the baseline sample of JEI firms. First, we only consider firms that enter the scheme the latest at age four. This allows us to analyze the growth trajectory of innovative firms that face a reduction in R&D labor costs in the very early stages of their existence. From these, we remove firms that appear to be controlled by other entities during the years in which they receive JEI benefits.[14] The remaining sample contains 2,677 JEI firms.

[13]We restrict our sample to firms with the following type of legal form and corresponding code in the French classification of *catégories juridiques*: limited liability companies (54), public limited company with a board of directors (55), public limited company with a management board (56) and simplified joint-stock company (57).

[14]There are two main reasons for which such a case may occur. First, the independence criterion requires that at least 50% of a JEI firm's capital be held by either a natural person or a certain type of entity, as explained in Footnote 4. It could be that some of the entities holding the majority of the firms' capital while receiving JEI benefits are of this nature, which would not have endangered the eligibility of their JEI status. Alternatively, a second possibility is that such firms were in fact controlled by another company outside this list of exceptions and should not have received the benefits for the years in question. With the data available, we are unable to distinguish between the two scenarios in a systematic way. We therefore remove all JEI firms that are majority-owned while being part of the scheme. This approach further ensures comparability with the identified non-JEI firms.

Non-JEI firms. We identify a set of firms that were eligible for the JEI status in the first years of existence, but never entered the scheme. After applying the general conditions mentioned above, we assess each eligibility criterion in turn. First, we evaluate whether the non-JEI firms meet the SME criteria every year within the first four years since foundation.[15] At the same time, we remove firms that never report any employees, as they would not be able to benefit from the scheme. Second, in line with the independence criterion, we remove firms that are majority-owned by another entity in any year within the same time frame. All remaining non-JEI firms benefit from the R&D tax credit at least once by the end of their fourth year of existence. This implies that for those respective years, they also report their R&D expenditures. We compute the R&D ratio by diving their R&D expenditures by the total expenses they report in their income statement in the same year. We keep only non-JEI firms for which the 15% threshold is met at least once within the initial four-year period. Given the validity test mentioned above, we are confident that this measure allows us to reliably identify firms that were eligible for the JEI status up to the age of four.

We are primarily interested in generating a group of comparable firms to those that benefit from the JEI scheme, i.e. young and innovative startups that engage in R&D activities. Imposing the same main restrictions on both JEI and non-JEI firms guarantees their comparability and eliminates potential sources of endogeneity. We provide a systematic, clear way of identifying eligible firms in the first years of existence, while circumventing various data availability problems.

Final sample. The final sample contains 4,433 firms, out of which 1,756 are non-JEI firms. Tables 3.2A and 3.2B present an overview of the baseline sample with respect to firm founding year and sector. The majority of JEI firms in the sample are active in *ICT* (41.5 %), while 19% of them report R&D as their primary activity. We group such firms into one category defined *Engineering and other technical & scientific activities*. Roughly 12% of the JEI firms are active in *Life Sciences*, which represents biotech startups or firms that conduct R&D in other natural science related fields. In comparison, the subsample of non-JEI firms contains a larger share of *Manufacturing* firms and a lower share of *Life Science* related startups.[16] We include sector controls in all regression specifications to account for specific sectoral characteristics that may influence both the growth trajectory of startups but also their likelihood of taking up the JEI status.

Summary Statistics. Table 3.3 reports summary statistics for the sample of JEI and non-JEI firms over the first eight years of existence. During this period, 75% of the firms in our sample remain active and only 2% of them are acquired. 17% of the firms seek patent protection

[15]This restriction removes a negligible amount of firms, as it is uncommon for young firms to exceed the SME criteria within their first four years.

[16]The *Other* group is also over-represented among the non-JEI firms, this being due to a larger share of firms reporting *Head office activities* as being their primary activity.

Table 3.2A: Baseline sample composition by founding year

Founding year	JEI	Non-JEI	Total
2006	261	159	420
2007	300	246	546
2008	349	261	610
2009	423	330	753
2010	471	272	743
2011	440	275	715
2012	433	213	646
Total	2,677	1,756	4,433

Table 3.2B: Baseline sample composition by sector

	JEI		Non-JEI	
	N	%	N	%
ICT	1,111	41.50	537	30.58
Life Science	320	11.95	88	5.01
Engineering and other technical & scientific activities	512	19.13	274	15.60
Manufacturing	224	8.37	243	13.84
Other	510	19.05	614	34.97
Total	2,677	100.00	1,756	100.00

Notes: We group firms in the baseline sample into five main sectors based on their first available economic activity code (NAF rév.2 classification). Appendix Table C.1 elaborates on the activity codes corresponding to each sector.

within the first four years of existence. Notably, firms tend to file their first patent quite early, with an average age at first filing of only three years. Moreover, by the end of the initial eight-year period, a patenting firm in our sample builds a portfolio that contains, on average, three patents. The average JEI firm in the sample receives approximately €230,000 worth of exemptions in social security contributions over the period it benefits from the JEI status. Annually, firms report an average of six employees and sales of roughly €0.5 million. Based on our measure of R&D personnel, an average of three employees are engaged in R&D-related activities. Patenting activity has a sporadic nature, especially among young firms. This pattern is also evident within our sample where the average annual filing rate does not exceed 0.08 patents.

3.3.3 Empirical Specification and Outcome Variables

We examine the sample of JEI beneficiaries and comparable non-JEI firms that were eligible for the scheme and evaluate firms' survival, innovation outcomes, employment and sales outcomes within the initial eight-year period since foundation. Our focus extends to a twelve-year period

Table 3.3: Summary statistics

	N	Mean	SD
Firm level			
Located in Paris (0/1)	4,433	0.23	0.42
Closed within first 8 years (0/1)	4,433	0.23	0.42
Being acquired within first 8 years (0/1)	4,433	0.02	0.15
Age at acquisition	97	6.70	1.42
Filing a patent within first 4 years (0/1)	4,433	0.17	0.37
Age at first patent filing	930	3.02	1.72
Patent stock at age 8	930	2.71	4.13
Total JEI exemptions (thousand €)	2,677	232.88	283.85
Firm-year level (8-year window)			
Sales (thousand €)	31,059	525.57	2,081.53
# employees	28,490	5.82	12.19
# employees (R&D)	25,551	3.27	10.84
# patents	31,082	0.08	0.49

Notes: We report summary statistics for the sample of JEI and non-JEI firms, observed for the first eight years of their life. We include firms that are founded between 2006 and 2012. Yearly statistics are computed conditional on firm survival.

when analyzing firms' probability of acquisition, as explained below. We estimate variants of the following baseline regression equation at the firm level:

$$Y_i = \beta_0 + \beta_1 JEI_i + \varphi_i + \delta_f + \lambda_s + \varepsilon_i, \tag{3.1}$$

where Y_i is one of the outcome variables described below and JEI_i is a binary indicator that takes the value 1 for firms that receive JEI benefits and value 0 for the non-JEI firms. We control for firm location using φ_i, an indicator of whether the startup is located in Paris or anywhere else in France. Firms located in Paris may benefit from being part of a larger entrepreneurial ecosystem which may, in turn, impact both their performance and their likelihood of taking up the JEI status. For instance, such firms may have access to a wider talent pool for skilled employees or researchers, may benefit from larger networking opportunities or a better developed startup support infrastructure. Furthermore, we condition on founding year fixed effects, δ_f, to absorb characteristics associated with each cohort of firms established in the same year. Sector fixed effects λ_s account for sector specific characteristics that may influence startups' survival, growth trajectory or propensity to patent. We aggregate firms' primary economic activities (NAF classification) into five main sectors, as described in Appendix Table C.1.[17]

First, we assess whether firms that benefit from the JEI scheme in the first years are more likely to survive until age eight, relative to firms that were eligible for JEI but never benefited from

[17] Firms' primary address and main activity can change over time. Whenever available, we take into account the information from the founding year. Otherwise, we rely on the first available observation with non-missing information for the two variables.

the scheme. The outcome variable *Closed* is a binary indicator that takes the value one for firms that closed down before the age of eight. The FICUS/FARE panels contain information regarding firm closure date and activity status (active, inactive, closed). Therefore, we consider a firm to be closed in any of the following cases: (1) a closure date is included in the data (FICUS), (2) the firm switches to an inactive or closed status (FARE), (3) the firm drops out of the data, suggesting that there are no other tax declarations filed. In the latter case, we consider the closure year to be that of the last occurrence in the panel, i.e. the year of the last tax declaration. A closed status may also include a positive outcome in the case of a merger or acquisition. Thus, we assign a value of zero to all identified acquired startups, treating them as active throughout the entire analysis period, as their closure is not related to a failure.

Second, for evaluating firms' probability of being acquired within the analysis period, we combine ownership information from LIFI and acquisition data from Crunchbase. As mentioned above, we retrieve all financial links of firms in our baseline sample that appear in LIFI. We identify acquisitions by tracking instances where there is a transition from a lower ownership percentage to 100%, which subsequently stands as the final recorded value in the LIFI panel. We ensure that we are not simply capturing internal restructuring events by conditioning on the owner not being a holding company or a head office. We consider the acquisition year the one in which the change to 100% ownership is documented. We identify additional acquisitions using information from Crunchbase. In absence of a direct link between the two sources, we match firms in our baseline sample with those that appear to have been acquired in Crunchbase based on their names.[18] As shown in Table 3.3, we observe 97 firms being acquired within the first eight years since foundation. Acknowledging that this time frame may be too restrictive for capturing acquisition events, our main results focus on an extended twelve-year period. Hence, our second performance measure, *Acquired*, is a binary indicator that takes the value one if the firm has been acquired within the extended period.

Since the JEI scheme aims to support the growth of young innovative startups and boost R&D activities within this space, evaluating firms' innovation outcomes is of particular importance in order to get a general understanding of the policy's effectiveness. For this purpose, we examine firms' patenting activity as a proxy for their innovation efforts. For emerging startups, alternative measures such as new products, services or processes are widely unavailable, rendering patents a valuable source of information. They not only represent the outcome of an invention process but also offer insights into the nature of the patented technology. Specifically, we analyze firms' patenting outcomes at ages five to eight both at the extensive and intensive margin. Thus, we generate a binary indicator that takes the value one if the firm filed a patent

[18] We obtain a similarity score between the full strings using the default matching technique employed by the *matchit* `Stata` command, which is based on a vectoral decomposition algorithm (Raffo, 2020). By cross-checking firm addresses, we noticed that even minor name differences leading to high similarity scores were generally causing false positive errors. Therefore, we match firms in both sources based on similarity scores of value one, strings being exactly the same. We identify few additional cases with lower similarity scores by manually inspecting the data.

in a given year and an annual patent count measure. As mentioned above, we group all patent applications into patent families, such that we count each invention only once. In addition, we identify patent families that include a European patent application filed at the European Patent Office, which we label EP patents for simplicity. A firm's decision to seek patent protection internationally signals its intention to enter additional markets, which, in turn, may capture the value of the underlying invention.[19] We therefore construct this measure as an indicator for higher-value inventions. We additionally capture the quality of inventions by building an alternative measure that weighs each patent by the number of forward citations it receives within five years since the first filing date.[20] Prior literature has shown that citation-weighted patent counts are a good indicator for innovation quality (see e.g., Harhoff, Scherer, & Vopel, 2003; Moser, Ohmstedt, & Rhode, 2018).

Lastly, we use log sales and the log number of employees as indicators of overall firm size. We primarily assess firms' performance along these dimensions at ages five to eight. Similar to Kerr, Lerner, and Schoar (2014), we account for survival bias by coding both employment and sales with zero values in the years following firm closure in the case of firms that have not survived.[21] We further distinguish between R&D and non-R&D workers based on information regarding employees' occupations and not on firm self-reported values. While acknowledging the endogeneity problem R&D employment is subject to in our setting, this analysis provides valuable insights into how JEI beneficiaries react to a decrease in R&D labor cost during their early stages. In particular, we are interested in understanding whether this reduction translates into an increase in R&D employment or rather in an increase in complementary staff, which could be indicative of an expansion of their operations.

3.4 Results

3.4.1 Firm Survival

We first provide results from a survival analysis as reported in Table 3.4. Columns 1 and 2 show results from estimating equation 3.1 with a binary indicator for firm closure as the dependent variable. We find that JEI beneficiaries are significantly less likely to close by the end of their eighth year compared to other young innovative companies eligible for the scheme. Specifically, a JEI beneficiary has a roughly 3 percentage points lower probability to close compared to a non-JEI firm. Results are robust to using a probit specification, as shown in Columns 3 and 4. Furthermore, results from a Cox proportional hazards model are shown in Columns 5 and 6. The negative and statistically significant coefficient confirms that JEI firms

[19]Harhoff, Scherer, and Vopel (2003) find that the size of the patent family is a good indicator for a patent's value.
[20]We count all citations received from subsequent EP applications within the 5-year window.
[21]We capture few post-acquisition years in the case of successful firms that get acquired within the eight-year window. We discard these observations from our analysis.

are slower to close compared to the non-JEI firms. Based on the estimate in Column 6, a JEI firm would be 1.14 times less likely to close within its first eight years relative to a non-JEI firm, holding everything else constant.

Table 3.4: JEI participation and firm survival

Variable	(1) OLS *Closed*	(2) OLS *Closed*	(3) Probit *Closed*	(4) Probit *Closed*	(5) Cox hazard *Years-to-close*	(6) Cox hazard *Years-to-close*
JEI	-0.030^{*}	-0.025^{+}	-0.097^{*}	-0.081^{+}	-0.151^{*}	-0.130^{*}
	(0.013)	(0.013)	(0.043)	(0.044)	(0.062)	(0.064)
from Paris (0/1)	−0.018	−0.018	−0.061	−0.064	−0.088	−0.088
	(0.015)	(0.015)	(0.050)	(0.052)	(0.075)	(0.077)
Constant	0.203^{***}	0.216^{***}	-0.838^{***}	-0.798^{***}		
	(0.021)	(0.023)	(0.076)	(0.082)		
Founding year FE	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	No	Yes	No	Yes	No	Yes
N	4,433	4,433	4,433	4,433	4,433	4,433
# JEI	2,677	2,677	2,677	2,677	2,677	2,677

Notes: The table examines the relationship between firm survival and JEI participation. The outcome variable *Closed* is a binary indicator that takes the value one for firms that closed down before the age of eight. We determine firm closure based on information regarding their official closure date (when available), activity status and their pattern of filing tax declarations. Identified acquired startups are assigned a value of zero and therefore considered alive. In Columns 5-6, we report results from estimating a Cox proportional hazards model. The negative and statistically significant coefficient indicates that JEI beneficiaries are slower to close compared to non-JEI firms. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.4.2 Firm Acquisition

In Table 3.5, we analyze the relationship between firms' likelihood of acquisition within their initial 12 years and JEI participation. The evidence suggests that JEI firms are 3.4% more likely to be acquired within the period of interest relative to non-JEI firms. The coefficient estimates are highly statistically significant at the 1% level across specifications. Moreover, Column 2 shows that firms active in *ICT* are significantly more likely to be acquired compared to other types of firms. Similar results are obtained when estimating a probit model, as shown in Appendix Table C.4.

For completeness, Appendix Table C.5 replicates the analysis using only the 2006-2011 cohorts in order to address any potential concerns regarding truncation when considering the extended analysis period.[22] The estimates align closely with the main results. Additionally, Appendix Table C.6 reports results from estimating equation 3.1 with the probability of being acquired

[22]Only the 2012 cohort of firms is partly affected by truncation given that the LIFI data covers only years up to 2019. However, the information extracted from Crunchbase allows us to track acquisitions for the last cohort as well, given that the data covers all years up to 2022.

within the limited eight-year window as the dependent variable. Also here, JEI firms are significantly more likely to be acquired relative to non-JEI firms. Furthermore, in Appendix Table C.7 we explore whether there is a relationship between the timing of acquisition and being a JEI beneficiary. The results indicate that there is no statistically significant difference in the age at acquisition between JEI and non-JEI firms. Therefore, although benefiting from the JEI status is associated with a higher likelihood of acquisition, it does not appear that JEI beneficiaries are acquired faster.

Table 3.5: JEI participation and firm acquisition

Variable	(1) OLS *Acquired*	(2) OLS *Acquired*	(3) OLS *Acquired*
JEI	0.034***	0.032***	0.034***
	(0.006)	(0.006)	(0.006)
from Paris (0/1)	0.023**	0.020*	0.018*
	(0.009)	(0.009)	(0.009)
ICT (0/1)		0.024**	
		(0.007)	
Life Science (0/1)		−0.015	
		(0.010)	
Constant	0.024***	0.018***	0.025***
	(0.004)	(0.005)	(0.005)
Founding year FE	Yes	Yes	Yes
Sector FE	No	No	Yes
N	4,433	4,433	4,433
# JEI	2,677	2,677	2,677

Notes: The table investigates the relationship between firms' probability to be acquired and JEI participation. The outcome variable is a binary indicator that takes the value 1 if a firms has been acquired within its first 12 years since establishment. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.4.3 Innovation Output

We build measures of firms' innovation output based on their patenting activity. Specifically, we examine firms' probability to patent and the number of patents they file at different moments in their life-cycle. Table 3.6 reports linear probability model estimates for two outcome variables measured at age five and at age eight: the probability to file a patent (Column 1 and 2) and the probability to file an EP patent (Column 3 and 4).[23] We find that JEI firms exhibit a higher likelihood to file patents at both ages. The results hold when focusing on EP patents, suggesting that JEI firms are more likely to seek patent protection internationally. This result

[23] Appendix Table C.8 reports regression results for outcomes at each age within the five-eight year-window. Estimates are similar across ages.

speaks to their intention to extend their market and reflects a higher value of their underlying inventions.

Table 3.6: JEI participation and the probability of patenting

	Any patent (0/1)		Any EP patent (0/1)	
	(1) Age = 5	(2) Age = 8	(3) Age = 5	(4) Age = 8
JEI	0.038***	0.023***	0.024***	0.024***
	(0.006)	(0.005)	(0.004)	(0.004)
from Paris (0/1)	−0.020***	−0.006	−0.008+	−0.005
	(0.006)	(0.005)	(0.005)	(0.004)
Constant	0.027***	0.011***	0.012***	0.002
	(0.004)	(0.003)	(0.003)	(0.002)
Founding year FE	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes
N	4,425	3,337	4,425	3,337
# JEI	2,670	1,958	2,670	1,958

Notes: This table reports regression results from estimating a linear probability model investigating the relationship between firms' likelihood of patenting and JEI participation. The outcome variables are binary indicators that take the value 1 if the firm files a patent in a given year (Columns 1-2) and an EP patent (Columns 3-4). EP patents labels patent families that include an EP application. The year is determined by the first filing date within each patent family. The outcome variables are coded with 0 values for closed firms to account for survival bias. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Next, we focus on assessing whether the number of patents differs across JEI firms and comparable young firms. Results shown in Table 3.7 suggest that on average, JEI beneficiaries file around 3% more patents at ages five and eight. Similar results are obtained when we restrict attention to EP patents. In Columns 5-6, we run additional specifications using citation-weighted patent counts as dependent variable. The results show that, on average, the patents of JEI firms receive more citations, which, in turn, suggests that they are of higher value. The positive and highly statistically significant coefficient is maintained across specifications and at all ages five to eight as shown in the complete Appendix Table C.9.

3.4.4 Employment and Sales

Lastly, we turn to an analysis of performance indicators related to firm size, such as employment and total sales. Columns 1-4 in Table 3.8 estimate equation 3.1 using the log number of employees at ages five to eight as dependent variable. In year five since foundation, a JEI beneficiary has roughly 20% more employees relative to a comparable non-JEI firm. The magnitude of the difference between the two groups is similar across all ages, as shown in subsequent Columns 2-4. Panel A of Appendix Figure C.7 shows that on average, both JEI and non-JEI firms are on a growth trajectory in terms of their number of employees. Taken together, these

Table 3.7: JEI participation and patent output

	Log # patents		Log # EP patents		Log # cit. weighted patents	
	(1) Age = 5	(2) Age = 8	(3) Age = 5	(4) Age = 8	(5) Age = 5	(6) Age = 8
JEI	0.034***	0.026***	0.020***	0.022***	0.020***	0.020***
	(0.006)	(0.005)	(0.004)	(0.004)	(0.006)	(0.005)
from Paris (0/1)	−0.012*	−0.008	−0.006+	−0.004	−0.010+	−0.006
	(0.006)	(0.005)	(0.004)	(0.004)	(0.006)	(0.004)
Constant	0.022***	0.009***	0.009***	0.002	0.013***	0.001
	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
Founding year FE	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes
N	4,425	3,337	4,425	3,337	4,425	3,337
# JEI	2,670	1,958	2,670	1,958	2,670	1,958

Notes: This table reports OLS regression results from investigating the relationship between firms' innovation output as reflected by their patenting activity and JEI participation. The outcome variables are the log number of patents, the log number of EP patents, and the log number of citation-weighted patents, respectively. All dependent variables are computed as $log(1 + \#\text{patents})$. EP patents labels patent families that include an EP application. Patents are weighted by the number of forward citations received from subsequent EP applications within 5 years since their first filing. The outcome variables are coded with 0 values for closed firms to account for survival bias. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

results suggest that JEI firms hire systematically more employees relative to other young innovative companies. Additionally, Appendix Table C.14 indicates the employment gap between JEI and non-JEI firms is maintained across ages nine to twelve.

With the reduction in R&D labor cost instrumented by the exemptions from paying social security contributions, JEI firms may either hire more employees or increase wages for existing employees. Prior work evaluating the impact of the JEI scheme on labor outcomes similarly finds an increase in employment, but either no effect or a mild increase in wages (Gautier & Wolff, 2020; Quantin, Bunel, & Lenoir, 2021). There are different hiring strategies that firms can take that explain an increase in employment. For instance, the additional resources may allow firms to further strengthen their focus on R&D activities and thus hire more R&D-related employees. On the other hand, firms may also enter a scaling phase where complementary, non-R&D employees would be required. We distinguish between R&D and non-R&D personnel using information regarding employees' occupations and analyze employment differences by type between the two groups of firms. Results presented in Table 3.9 suggest that JEI firms' higher number of overall employees is explained by both a higher number of non-R&D and

Table 3.8: JEI participation and firm size indicators

	Log # employees				Log Sales			
	(1) Age = 5	(2) Age = 6	(3) Age = 7	(4) Age = 8	(5) Age = 5	(6) Age = 6	(7) Age = 7	(8) Age = 8
JEI	0.181***	0.168***	0.185***	0.186***	−0.409**	0.008	0.089	0.143
	(0.030)	(0.034)	(0.036)	(0.038)	(0.143)	(0.162)	(0.179)	(0.192)
from Paris (0/1)	0.098**	0.110**	0.103*	0.092*	0.000	0.093	0.132	0.091
	(0.035)	(0.039)	(0.042)	(0.045)	(0.171)	(0.190)	(0.209)	(0.226)
Constant	1.349***	1.296***	1.203***	1.119***	3.727***	3.021***	2.394***	1.836***
	(0.025)	(0.027)	(0.029)	(0.030)	(0.112)	(0.130)	(0.143)	(0.153)
Founding year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4,149	4,153	4,226	4,248	4,397	4,394	4,390	4,360
# JEI	2,539	2,544	2,567	2,571	2,647	2,649	2,646	2,627

Notes: This table examines the relationship between JEI participation and firm size indicators measured at ages 5-8. Dependent variables are the log number of employees and log sales computed with offset 1. They are coded with 0 values for closed firms to account for survival bias. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

R&D employees.[24] Indeed, higher coefficients are observed for R&D compared to non-R&D personnel. We interpret these findings as evidence of the fact that innovative young firms facing a reduction in R&D labor costs in their early stages are able to sustain a focus on R&D activities, by employing a significantly larger number of R&D employees relative to comparable firms. Because of the interrelation between the number of R&D workers and JEI status by the end of age eight, we complement our results with an analysis of the same outcomes in later years, at ages nine to twelve . Appendix Table C.11 shows that throughout the additional period, JEI beneficiaries still employ a significantly higher number of R&D workers, albeit with a decreasing difference. In contrast, over the age of eight, the number of non-R&D employees of JEI and non-JEI firms does not differ significantly anymore.

Our final results focus on the relationship between JEI participation and firm size measured by the log of sales, as shown in Columns 5-8 of Table 3.8. We observe that at age five, JEI firms have significantly lower sales relative to other young innovative companies. However, in the following years the difference between the two groups declines as evidenced by the lower point estimates which also turn statistically insignificant. Additional results presented in Appendix Table C.12 provide evidence consistent with the idea that JEI beneficiaries catch up to the comparable young firms in terms of revenue generation. The same pattern can be visually observed in Panel B of Appendix Figure C.7. This finding acts as a valuable supplementary

[24]We alleviate concerns regarding JEI firms' potential strategic relabelling of employees as R&D personnel during the eligibility period by constructing an objective measure and not relying on self-reported counts. We additionally showcase the robustness of our results employing the tech-worker definition of Harrigan, Reshef, and Toubal (2023) instead, as shown in Appendix Table C.10.

Table 3.9: JEI participation and type of employment

	Log # non-R&D employees				Log # R&D employees			
	(1) Age = 5	(2) Age = 6	(3) Age = 7	(4) Age = 8	(5) Age = 5	(6) Age = 6	(7) Age = 7	(8) Age = 8
JEI	0.071*	0.100**	0.111**	0.083*	0.240***	0.240***	0.264***	0.272***
	(0.032)	(0.035)	(0.037)	(0.039)	(0.029)	(0.032)	(0.034)	(0.036)
from Paris (0/1)	0.142***	0.177***	0.155***	0.179***	0.026	0.014	0.075^{+}	0.029
	(0.037)	(0.041)	(0.044)	(0.047)	(0.034)	(0.039)	(0.042)	(0.044)
Constant	1.285***	1.211***	1.131***	1.066***	0.710***	0.746***	0.729***	0.721***
	(0.026)	(0.028)	(0.030)	(0.032)	(0.024)	(0.026)	(0.027)	(0.028)
Founding year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	3,839	3,735	3,709	3,685	3,839	3,735	3,709	3,685
# JEI	2,386	2,313	2,293	2,272	2,386	2,313	2,293	2,272

Notes: This table explores the relationship between JEI participation and employment, distinguishing between R&D and non-R&D staff. The dependent variables are the log number of non-R&D workers (Columns 1-4) and the log number of R&D-workers (Columns 5-8) evaluated at ages 5-8. They are coded with 0 values for closed firms to account for survival bias and computed with offset 1. We identify R&D employees based on information regarding employees' occupations and not on firm self-reported values. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

result highlighting JEI beneficiaries' growth trajectory.

3.5 Discussion and Conclusion

This paper analyzes the French *Jeune Entreprise Innovante* program aimed at supporting innovative startups through exemptions from social security contributions for their R&D employees. The JEI scheme is a rare and unconventional government R&D support tool which has been utilized by more than 10,000 startups in France since its inception in 2004. Young innovative startups as supported by the JEI scheme typically experience financing constraints and high cost of R&D labor on which they rely strongly, especially in their early stages of development. We document the uptake of the JEI support since its introduction and highlight its extensive utilization, as evidenced by the significant amount of exemptions granted.

Leveraging access to uniquely rich French administrative data complemented with patent and firm acquisition information from external sources, we study the early impact of the reduction in social security contributions for R&D labor on firm performance. We find that startups receiving JEI benefits in the initial phase of their development are significantly more likely to survive, to be acquired, and to patent relative to comparable young innovative firms that never use the JEI status although being eligible. Additionally, our analysis reveals that compared to non-JEI startups, JEI beneficiaries employ more workers of both R&D and non-R&D type. Moreover, while these firms are typically smaller (in terms of revenues) than non-JEI firms of the same age, they catch up in terms of average revenues in the years following their entry into the scheme. We interpret these results as evidence consistent with the idea that JEI benefits received in the initial stage of development support firms on their growth trajectory.

The ideal scenario for causally identifying the effect of JEI participation on firm outcomes would entail the JEI status being randomly assigned to firms. However, the policy is set-up in such a way that firms self-select into the scheme, as they have to declare their JEI eligibility status when filing their tax declarations. This may raise concerns regarding the existence of a potential selection bias affecting the validity of our results. In other words, if firms that take up JEI are systematically different from those that do not along characteristics that correlate with firm outcomes, the results of our analysis would be biased. Imagine a case of positive selection, where innovative and high-potential startups are more likely to make use of the JEI status. Even in the absence of JEI benefits, they are naturally more likely to survive or get acquired due to their innovative nature. In such a case, we would likely overestimate the effect of the scheme on firm performance.

We attempt to control for the potential selection bias by identifying a sample of firms comparable to JEI beneficiaries and not using the scheme. Evaluating firms' eligibility for the JEI status is empirically challenging given the five criteria that they must fulfill. However, the richness of our data enables us to confidently identify a set of similar young innovative companies eligible for the scheme in their initial stages, but not benefiting from it. To this end, we are able to condition on firms' early engagement in R&D activities which are measured with high precision in the French administrative data. This is generally difficult to achieve in other settings due to the widespread unavailability of R&D expenditure data for such young firms. Our approach allows us shed light on the performance of young companies that make use of the scheme. Despite our careful construction of the comparison group, we cannot claim that JEI beneficiaries are more likely to be successful solely due to the JEI benefits they receive and therefore provide a causal interpretation of our results. While we acknowledge this limitation, the persistent positive results we obtain are consistent with the idea that the scheme aids firms in their development and they cannot be solely attributed to a selection bias.

Our study provides insights into startups' utilization of the JEI scheme as well as its effectiveness in supporting firms in the early stages of their development. However, many questions regarding the policy tool remain unanswered. Our current setting does not allow us to provide evidence on potential mechanisms through which JEI benefits might positively influence firms' performance. Startups may leverage their JEI status to signal their innovative capability to potential investors or acquirers. At the same time, the status might allow them to attract better-qualified researchers. Alternatively, the financial resources saved through the reduction in social security contributions may be directed toward additional investments, such as the further development of their technology or the pursuit of new and exploratory ideas. In an extension of the current analysis, we aim to retrieve data on venture capital financing rounds and investigate whether JEI beneficiaries have better access to external financing. One particular challenge in this setting revolves around the finding that JEI beneficiaries are more likely to patent relative to the other young companies. Patents have been shown to increase firms' access to venture capital financing (see e.g., Haeussler, Harhoff, & Mueller, 2014; Farre-Mensa, Hegde, & Ljungqvist, 2020). Therefore, we would need to disentangle the effect of patenting from the effect of the JEI participation on access to financing.

Over the last years, French policymakers made changes to the JEI support scheme and gradually extended its application. The relevance of our overall positive findings is highlighted by their continued investment and further promotion of the JEI scheme. Specifically, in the Finance bill for 2024-2028, the government has committed to broaden its scope and cover also startups with an R&D expenditure threshold below 15%, but above 5%, introducing the new status of Young Innovative Growth Company (*Jeune Entreprise Innovante de Croissance*). Furthermore, they introduced the status of Young Disruptive Innovation Company (*Jeune Entreprise d'Innovation de Rupture*) for startups with at least 30% of their expenses going towards R&D activities.[25] This further extension speaks to the French government's continued commitment toward supporting the financing of young startups. These changes provide future avenues for research exploring impact heterogeneity and the mechanisms through which different types of firms benefit from such an R&D government support tool.

[25] Besides defining two new groups of firms, the amendment also introduces new types of support, in addition to exemptions on social security payments. For details see: https://www.frenchtechbordeaux.com/comprendre-le-statut-de-jeune-entreprise-innovante-jei/.

A

Appendix to Chapter 1

Multinational Firms and Global Innovation

A.1 Theoretical Appendix

Proof of Proposition 1. Firm profits $\pi(\varphi)$ are supermodular in productivity and innovation quality.

Intensive margin:

$$\frac{\partial^2\pi(\varphi)}{\partial q_j^A \partial\varphi} = R\left(\frac{\alpha P}{w_s}\right)^{\sigma-1} \frac{(\sigma-1)^2}{\sigma}\left(1+\sum_j q_j^B(\varphi)\right)\left(1+\sum_j q_j^A(\varphi)\right)^{\sigma-2} \varphi^{\sigma-2} > 0, \quad \text{and}$$

$$\frac{\partial^2\pi(\varphi)}{\partial q_j^B \partial\varphi} = R\left(\frac{\alpha P}{w_s}\right)^{\sigma-1} \frac{(\sigma-1)}{\sigma}\left(1+\sum_j q_j^A(\varphi)\right)^{\sigma-1} \varphi^{\sigma-2} > 0, \quad \text{and further}$$

$$\frac{\partial^2\pi(\varphi)}{\partial q_j^A \partial q_j^B} = R\left(\frac{\alpha P}{w_s}\right)^{\sigma-1} (\sigma-1)\left(1+\sum_j q_j^A(\varphi)\right)^{\sigma-2} \frac{\varphi^{\sigma-1}}{\sigma} > 0. \qquad \square$$

Extensive margin: As profits are increasing and supermodular in innovation quality and productivity, more productive firms are more likely to amortize the fixed costs of innovation for every innovation type and location.

Proof of Proposition 2. Follows from Proposition 1, the ranking of fixed costs of innovation, and the assumption that fixed costs have to be paid for each country.

Proof of Proposition 3. Firm profits $\pi(\varphi)$ are submodular in inventor wages and innovation quality.

Intensive margin:

$$\frac{\partial^2\pi(\varphi)}{\partial q_j^A \partial r_j} = -\sum_j \mathbf{1}(q_j^A > 0)(q_j^A(\varphi))^{\beta-1} < 0, \quad \text{and}$$

$$\frac{\partial^2\pi(\varphi)}{\partial q_j^B \partial r_j} = -\sum_j \mathbf{1}(q_j^B > 0)(q_j^B(\varphi))^{\beta-1} < 0, \quad \text{and further}$$

$$\frac{\partial^2\pi(\varphi)}{\partial q_j^A \partial q_j^B} = R\left(\frac{\alpha P}{w_s}\right)^{\sigma-1} (\sigma-1)\left(1+\sum_j q_j^A(\varphi)\right)^{\sigma-2} \frac{\varphi^{\sigma-1}}{\sigma} > 0. \qquad \square$$

Extensive margin: Result follows from intensive margin result along with profits being increasing in innovation qualities.

Proof of Proposition 4. Follows from ranking of fixed costs: Fixed costs of applied innovation are strictly lower when co-located with production. Fixed costs of basic innovation are independent of a firm's organizational structure, so applied innovation is *ceteris paribus* more likely to be co-located with production than basic innovation.

A.2 Data Construction

A.2.1 Microdatabase Direct Investment (MiDi)

This paper uses foreign direct investment administrative data on German multinational firms from the Microdatabase Direct Investment (MiDi) maintained by the Deutsche Bundesbank.[1] This database contains annual German outward and inward FDI information for the period 1999-2016. Since we are interested in the global network of affiliates of German multinational firms, we limit our analysis to firms reporting outward direct investments. MiDi contains information at the individual investment relationship level. Both direct and indirect investment relationships between a German parent company and its foreign subsidiaries are included.

Based on the German Foreign Trade and Payments (*Aussenwirtschaftsverordnung*) decree, German companies are required to report information regarding their foreign direct investments to the Deutsche Bundesbank if they:

- own directly at least 10 % of the shares (or voting rights) in a foreign company that has a balance sheet total above EUR 3 mil.
- own a combined controlling share of more than 50% in a foreign company with a balance sheet total above EUR 3 mil (either indirectly or through a combination of direct and indirect shares).

These reporting rules have been in place since 2007, after two main historical changes in 2002 and 2007. For our analysis, we take into account firms that were not affected by changes in the reporting requirements over time, i.e. firms that meet all reporting requirements during 1999-2016. Following this strategy, we implicitly remove all firms that voluntarily report to the Deutsche Bundesbank, without being required to. Additionally, we remove all public or private households that fall under the reporting requirements. Given that firms are legally bound to report information regarding their foreign operations, MiDi contains highly reliable, "close to complete" data (Drees, Schild, & Walter, 2018).

In this paper, we are primarily interested in the location of German MNCs' foreign affiliates such that we construct an annual mapping of their global operations. In addition, we use both parent and affiliate turnover information in order to construct our main productivity proxy, global sales. For this computation, we weight each affiliate's turnover by the parent's

[1] This paper uses the 2018 version of the MiDi database. DOI: 10.12757/Bbk.MiDi.9916.04.05. See Drees, Schild, and Walter (2018) for detailed information on the database.

total participation in the firm. As a robustness check, we compute an alternative productivity proxy, parent sales per employee.

Our main outcomes of interest relate to firms' patenting activity. Therefore, the construction of our baseline sample of German MNCs requires a linkage between MiDi and patent data obtained from PATSTAT Global. In absence of a direct link between MiDi parent firms and patent applicants in PATSTAT, we rely on information from Bureau van Dijk's Orbis dataset. The Deutsche Bundesbank Research and Data Center has developed a mapping from MiDi parent firms to Orbis firm identifiers (BvD ID) using supervised machine learning see Schild, Schultz, and Wieser, 2017. Through a crosswalk from Orbis (version 2016) to PATSTAT, we are able to link the two databases of interest. We retrieve the firms' primary industry of activity in the 2-digit NACE 2.0 classification from Orbis as well. Therefore, our baseline sample of MNCs comprises parent firms with at least one foreign affiliate active in MiDi that is also captured in Orbis, such that we are able to assess whether they had filed any patents in the period of interest.

A.2.2 PATSTAT

In order to analyze firms' patenting activity, we retrieve all patents filed during 1999-2016 by the MNCs in our baseline sample from PATSTAT Global (version autumn 2018). In our analysis, our focus remains on the patents that are filed by the parent firms in our baseline sample, abstracting from patents originating from firms' affiliates alone. The reason for this is twofold: first, the data available does not allow us to link firms' foreign affiliates in MiDi with patent applicants in PATSTAT. Second, affiliate innovation strategies could be influenced by local market characteristics which could lead to systematic differences in the patents that are filed by affiliates relative to the parent firms. Different intellectual property strategies of affiliates and parents may also introduce systematic differences in the type or quality of the patents filed. We further restrict our sample to patents that have only a unique MNC owner. Therefore, we remove co-inventions across multiple MNCs that would involve strategic decisions that go beyond our paper. However, note that our sample would still include patents that have other applicants outside of the sample of MNCs that we observe.

We group all patent filings originating from our baseline firms into patent families. A patent family is a collection of patent applications that are filed across multiple jurisdictions but that essentially cover a single invention or technology. Throughout the analysis, we avoid multiple counting of the same invention by using DOCDB patent families instead of single applications. For each patent family, we determine the year of the first patent filing within the family, as the closest point in time to the development of the invention. Additionally, for each patent family, we identify the main technological area among the 34 areas proposed by Schmoch (2008). We choose the technology area that is most common across all filings in the patent family.

Whenever a mode cannot be identified, we select the main technological area of the first filing within the family. We drop patent families that do not contain any application that represents a patented invention, i.e. we remove families that only contain utility models or design patents. By EP patents we refer to patent families that contain an EP application (European patent application filed at the European Patent Office). We measure patent quality by summing up all EP patent citations each focal patent family received within a 5-year window. We count forward citations originating from EP applications in order to maintain comparability, as citation patterns vary systematically across patent offices.

We obtain inventor information from the latest publication document of each patent application retrieved from PATSTAT. Since inventor information can be incomplete across applications within the same family, we develop an algorithm such that we harmonize information at the DOCDB family level. Specifically, we prioritize information from applications filed at the European Patent Office, at the United States Patent and Trademark Office, German Patent and Trademark Office (DPMA) and the World Intellectual Property Organization, as they would be the most likely to contain complete information. We separated patent families that contain at least one of the applications above and those that do not. For each group in turn we take the following cleaning steps: (1) we count the number of inventors for each application and compute the number of inventors where country information is missing, (2) we identify the application with the lowest number of inventors with country information missing and is also the earliest within the family. For each patent family, we retrieve the inventor location information from this identified application. We remove patents for which we cannot identify the inventor location.

Once we combine data on affiliate location in MiDi with inventor countries in PATSTAT, we are able to distinguish between:

- **domestic patents**: all inventors are located in Germany
- **offshore patents**: at least one inventor is located abroad
  - **offshore *co-located* patents**: at least one of the foreign inventor countries match with an affiliate country (affiliates active in the same year as the patent filing year)
  - **offshore *not co-located* patents**: none of the foreign inventor countries match any of the affiliate countries (affiliates active in the same year as the patent filing year)

A.2.3 Patent Type: Science-based and Non-science-based Patents

We distinguish between two different patent types as proxies for firms' applied and basic R&D activities. We do so by using patents' distance to science following Ahmadpoor and Jones (2017). Specifically, we assume that a patent with a short distance to a scientific article would result from basic R&D activities. We define these as science-based patents. Alternatively, patents that are more distant from fundamental science are associated with applied R&D activities and labeled as non-science-based patents.

We construct the distance to science measure for EP patent applications included in our baseline sample. We restrict our attention to applications filed only in one patent office to ensure comparability, given that citing patterns differ across offices. We retrieve backward citations for all applications of interest from PATSTAT Global. Additionally, we link all focal patents and their corresponding patent citations with the *Reliance on Science* open-access dataset provided by Marx and Fuegi, 2020. This includes patent front-page citations to scientific articles retrieved from Microsoft Academic Graph and PubMed.

Patents that directly cite a scientific article receive a distance to science score of 1. For the remaining patents, we consequently check whether their cited patents in turn cite a scientific article. We repeat this step until we are able to establish how many degrees distant the focal patents are from scientific articles. Therefore, our measure produces a score of $\{1, 2, 3, ...\}$, with lower scores indicating a more closer connection to fundamental science. We label patents receiving a score of $\{1, 2\}$ as *science-based* and patents receiving a score of at least 3 as *non-science-based*.

A.2.4 Revealed Comparative Advantage in Innovation

We propose a measure of countries' revealed comparative advantage (RCA) in innovation that captures countries' capacity to enable innovation that results in patent filings. We define country c's revealed comparative advantage in technology area a and year t as:

$$RCA_{act} = \frac{\sum PAT_{act}}{\sum PAT_{ct}} \times 100,$$

where $\sum PAT_{act}$ represents the total number of patents in technology area a in year t that originate in country c. We scale by the total number of patents invented in country c in the same year in order to account for country size. The measure allows us to identify the technology areas where countries have most expertise.

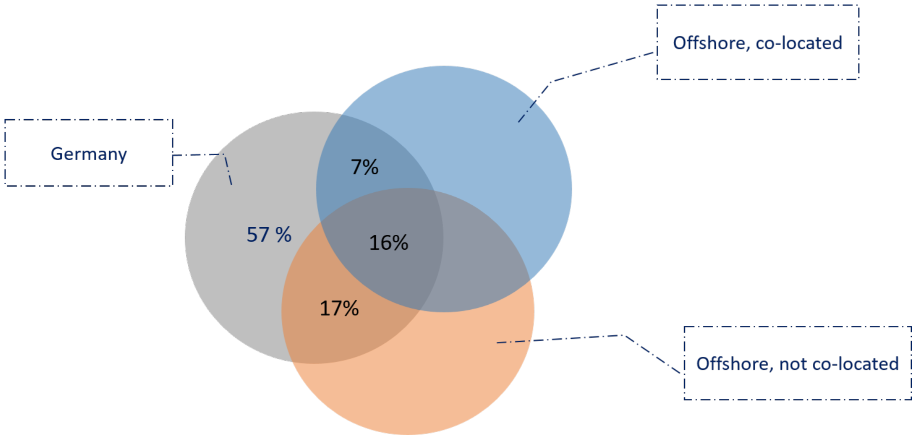
We retrieve the full set of international patent families available in PATSTAT Global (version spring 2021)[2] that contain at least one application filed at three of the top five leading patent authorities in the world. These include the European Patent Office (EPO), the United States

[2]This is the only part of the analysis that relies on a different PATSTAT version. We do so in order to capture the complete 1999-2016 period. Due to lags in the patenting process via the international route (PCT), where it can take up to 32 months from the first patent application to subsequent filings, our original PATSTAT Global (v. autumn 2018) would have included truncated data for the last years of interest. See Dechezleprêtre, Ménière, and Mohnen (2017) for a discussion on international patent families.

Patent and Trademark Office (USPTO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO) and the China National Intellectual Property Administration (CNIPA). Hence, we consider patent families that include at least one application at the EPO, one at the USPTO and an additional one at either JPO, KIPO or CNIPA. Using this definition allows us to generate a comparable measure across countries that is not affected by countries' different patent filing propensities. Additionally, only higher-quality inventions are patented in multiple jurisdictions (Harhoff, Scherer, & Vopel, 2003). Firms would only seek protection in a larger geographical region and in turn, incur the higher patent filing costs that come with that decision for higher quality inventions. We exclude all patent families that have German applicants in constructing our measure. Ideally, we would have excluded only the patents filed by the German MNCs included in our analysis so that we ensure that our measure does not capture patenting decisions of the firms we are interested in. However, due to confidentiality rules at the Research Data and Service Center of the Deutsche Bundesbank, we have no way to link the MNCs and PATSTAT outside of the research center. Therefore, we took the more conservative approach of removing all patents that have a German applicant among those we select for constructing the RCA measure. We retrieve inventor information for all patents of interest and follow the same cleaning steps as mentioned in Appendix A.2.2. We assign each patent to inventor countries using fractional counts as explained above.

A.3 Additional Descriptive Findings

Figure A.1: Location of Global MNC Innovation: All Patents



Notes: This Venn diagram summarizes the global organization of German MNC patent activity in 1999-2016. Each segment indicates the share of firms that file patents with inventors residing at home in Germany, offshore in a country with an MNC affiliate, and/or offshore in a country with no MNC affiliate. N = 2,374 MNCs. *Data sources:* Research Data and Service Center of the Deutsche Bundesbank (MiDi, 1999-2016) and PATSTAT, own calculations.

Table A.1: Top Foreign Innovation Hubs for German MNCs

	Overall, 1999-2016		2000		2015	
	Country	% offshore patents	Country	% offshore patents	Country	% offshore patents
1	US	19.2 %	US	33.2 %	US	16.6 %
2	FR	8.0 %	AT	9.1 %	AT	7.6 %
3	AT	6.9 %	FR	7.2 %	FR	6.0 %
4	CH	5.2 %	CH	5.1 %	IT	5.0 %
5	IT	4.0 %	JP	3.5 %	CN	4.9 %

Notes: This table lists the top-5 foreign countries where German MNCs invent patents. Countries are ranked by their share of all offshore MNC patents. Fractional counts are used for patents with multiple inventor countries. *Data sources:* Research Data and Service Center of the Deutsche Bundesbank (MiDi, 1999-2016) and PATSTAT, own calculations.

Figure A.2A: MNC Size and Innovation Intensity: All Patents

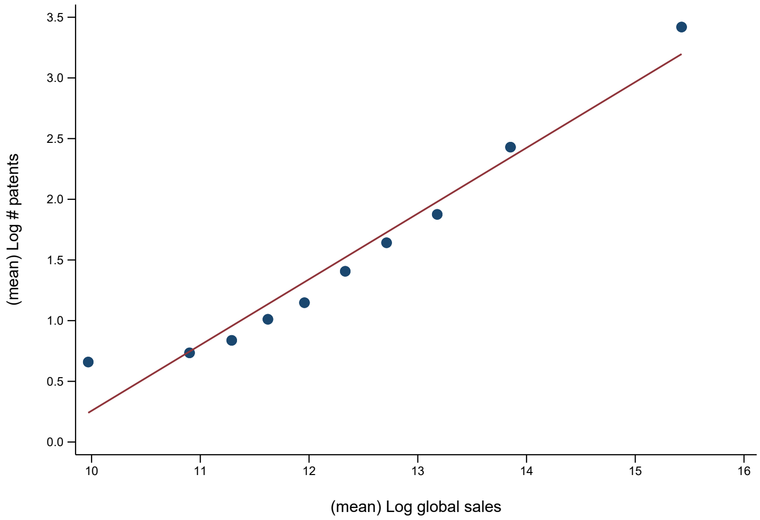
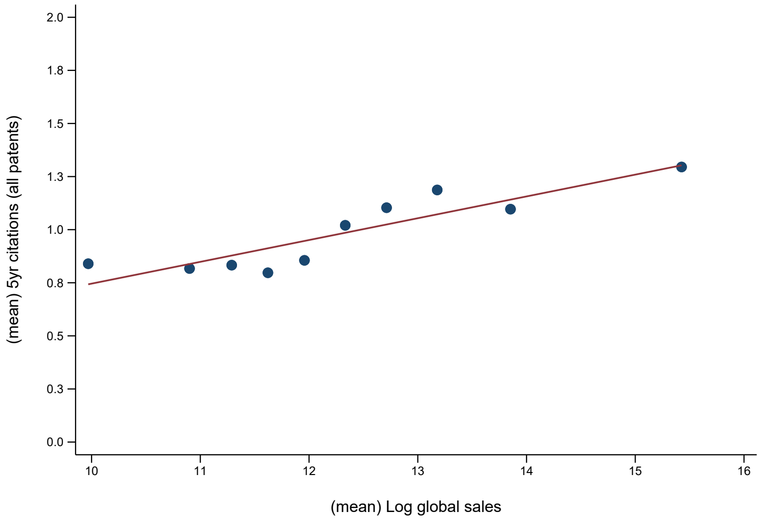
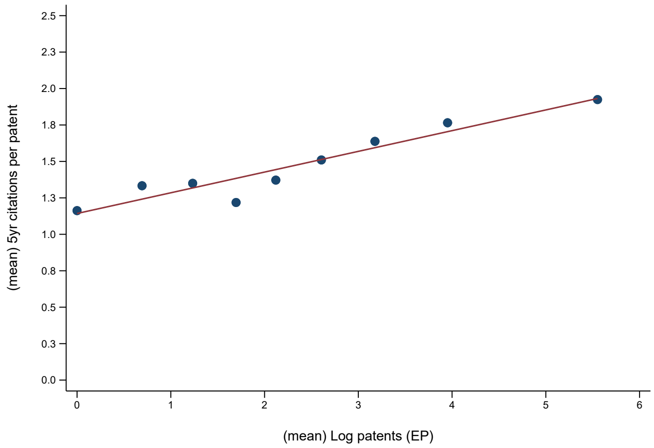


Figure A.2B: MNC Size and Innovation Quality: All Patents



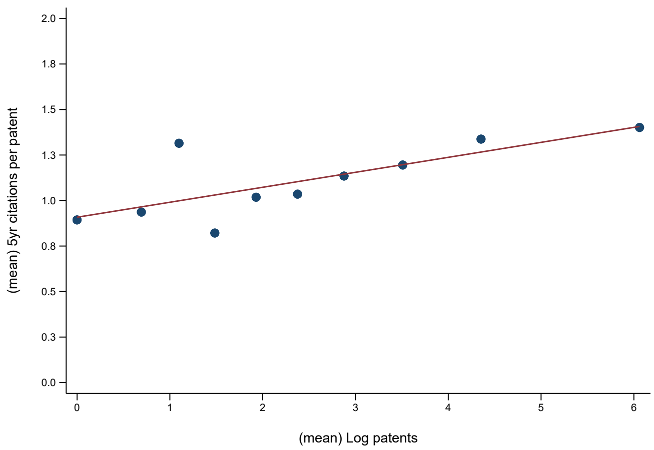
Notes: These binscatters plot the log average annual number of all patents per firm in 1999-2016 and the average number of 5-year forward citations per patent per firm in 1999-2011, by firm size bin. German MNCs are assigned to ten bins each year according to their annual global sales. Year fixed effects are absorbed. *Data sources:* Research Data and Service Center of the Deutsche Bundesbank, MiDi, 1999-2016, combined with PATSTAT and World Bank National Accounts data, own calculations.

Figure A.3: Innovation Intensity and Quality Across MNCs: EP patents



Notes: This binscatter plots the average number of 5-year forward citations per EP patent per firm in 1999-2011, by firm patent intensity bin. German MNCs are assigned to ten bins each year according to their annual number of EP patents. Year fixed effects are absorbed. *Data sources:* Research Data and Service Center of the Deutsche Bundesbank, MiDi, 1999-2016, combined with PATSTAT, own calculations.

Figure A.4: Innovation Intensity and Quality Across MNCs: All Patents



Notes: This binscatter plots the average number of 5-year forward citations per patent per firm in 1999-2011, by firm patent intensity bin. German MNCs are assigned to ten bins each year according to their annual number of patents. Year fixed effects are absorbed. *Data sources:* Research Data and Service Center of the Deutsche Bundesbank, MiDi, 1999-2016, combined with PATSTAT, own calculations.

A.4 Robustness Checks

Table A.2: Innovation Intensity (All patents)

Dependent variable	(1) any patent (0/1)	(2) log # patents	(3) log # citation weighted patents	(4) avg log # citations
Log global sales	0.040***	0.567***	0.553***	0.022***
	(0.002)	(0.029)	(0.031)	(0.004)
# MNC-years	68,999	11,837	7,329	11,837
Sector FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: This table reports empirical results from estimating equation (1.1) to evaluate Proposition 1 *More productive firms are more likely to innovate and to innovate more intensively*. Column 1 examines the extensive margin in the full panel of German multinational firms. The outcome variable is an indicator for any patenting activity within each firm - year combination taking into account the full patent sample. Columns 2-4 evaluate the intensive margin in the panel of patenting multinational firms. Outcome variables are the log number of patents, log number of citation weighted patents and the average log number of citations across patents by firm-year, respectively. Standard errors are clustered at firm level. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Data sources:* Research Data and Service Center of the Deutsche Bundesbank (MiDi, 1999-2016) and PATSTAT, own calculations.

Table A.3: Innovation Intensity (Robustness)

Dependent variable	(1) any patent (0/1)	(2) log # patents	(3) log # citation weighted patents	(4) avg log # citations
Panel A1. All patents				
Log domestic sales/employees	-0.021^{***}	0.192^{***}	0.241^{***}	0.029^{***}
	(0.004)	(0.047)	(0.058)	(0.008)
# MNC-years	40,680	11,202	6,944	11,202
Panel A2. EP patents				
Log domestic sales/employees	-0.013^{***}	0.173^{***}	0.228^{***}	0.029^{**}
	(0.003)	(0.043)	(0.058)	(0.010)
# MNC-years	40,680	9,047	5,858	9,047
Panel B. EP Basic (science-based)				
Log domestic sales/employees	0.054^{***}	0.201^{**}	0.243^{*}	0.022
	(0.012)	(0.065)	(0.095)	(0.015)
# MNC-years	8,555	3,796	2,423	3,796
Panel C. EP Applied (non-science-based)				
Log domestic sales/employees	-0.002	0.148^{***}	0.184^{***}	0.031^{**}
	(0.007)	(0.040)	(0.051)	(0.010)
# MNC-years	8,555	7,820	4,968	7,820
Sector FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: This table reports empirical results from estimating equation (1.1) to evaluate Proposition 1 *More productive firms are more likely to innovate and to innovate more intensively*. In this robustness check, we proxy firm productivity using domestic sales/employees. Column 1 examines the extensive margin in the full panel of German multinational firms. The outcome variable is an indicator for any patenting activity within each firm - year combination. Columns 2-4 evaluate the intensive margin in the panel of patenting multinational firms. Outcome variables are the log number of patents, log number of citation weighted patents and the average log number of citations across patents by firm-year, respectively. Standard errors are clustered at firm level. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Data sources:* Research Data and Service Center of the Deutsche Bundesbank (MiDi, 1999-2016) and PATSTAT, own calculations.

Table A.4: Innovation Offshoring (Robustness)

	EP patents		Basic EP patents		Applied EP patents	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Dependent variable: Any offshore patent (0/1)						
Log (domestic sales/employees)	0.054***		0.067***		0.049***	
	(0.012)		(0.018)		(0.013)	
# MNC-years	9,047		3,796		7,820	
Panel B. Dependent variable: Share offshore patents						
Log (domestic sales/employees)	0.025***		0.037***		0.018**	
	(0.007)		(0.011)		(0.007)	
# MNC-years	9,047		3,796		7,820	
Panel C. Dependent variable: # foreign inventor countries						
Log (domestic sales/employees)	0.309	0.132	0.445+	0.317	0.176	0.005
	(0.199)	(0.214)	(0.228)	(0.253)	(0.144)	(0.124)
# affiliate countries		0.149***		0.115**		0.131***
		(0.040)		(0.037)		(0.036)
# MNC-years	2,746	2,746	1,251	1,251	2,175	2,175
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table provides empirical results in line with Proposition 2 *More productive MNCs are more likely to offshore innovation and to innovate in more countries*. Panel A examines the extensive margin in the sample of patenting MNCs: we analyze whether a firm generates any patents with at least one foreign inventor, either overall or separately within each innovation type. For brevity, we refer to patents with at least one foreign inventor as *offshore patents*. Panel B examines the intensive margin by studying the percent share of patents with a foreign inventor in the sample of patenting MNCs, conditioning on MNCs having at least one basic or applied patent in columns 3 and 5, respectively. Lastly, Panel C studies the number of foreign inventor countries in the panel of patenting MNCs, conditioning on MNCs having basic or applied patents in columns 3-4 and 5-6, respectively. At the firm x year level, we count the number of distinct foreign inventor countries related to all patents, basic patents and applied patents for each corresponding column. In columns 2, 4 and 6, we additionally control for the number of affiliate countries the MNCs are present in. Standard errors clustered at the firm level are presented in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Data sources:* Research Data and Service Center of the Deutsche Bundesbank (MiDi, 1999-2016) and PATSTAT, own calculations.

Table A.5: Mixed Innovation Offshoring (Robustness)

Base level: Any offshore not co-located patent	(1)	(2)	(3)	(4)
Any offshore co-located patent				
Log (domestic sales/employees)	0.251^{*}	0.242^{*}	0.050	0.122
	(0.098)	(0.096)	(0.089)	(0.090)
# affiliate countries			0.127^{***}	0.135^{***}
			(0.016)	(0.018)
Both co-located and not co-located offshore patents				
Log (domestic sales/employees)	0.421^{**}	0.365^{*}	0.239	0.247
	(0.140)	(0.149)	(0.155)	(0.164)
# affiliate countries			0.127^{***}	0.137^{***}
			(0.016)	(0.018)
Year FE	Yes	Yes	Yes	Yes
Sector FE	No	Yes	No	Yes
# MNC-years	2,756	2,750	2,756	2,750

Notes: This table presents estimation results from a multinomial logit regression aimed to assess Corollary 1 *More productive MNCs are more likely to innovate both in locations with and in locations without a production affiliate.* The outcome variable takes the value 1 if the firms has any patent with inventors located in countries without affiliates in a given year, value 2 if there is any offshore patent co-located with production and value 3 if both strategies are employed. The three categories are mutually exclusive. We regress the outcome variable on a proxy for productivity in the form of domestic sales per employees. We further add the number of affiliate countries as a control in later specifications. We condition on year and sector fixed effects. Standard errors clustered at the firm level are presented in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Data sources:* Research Data and Service Center of the Deutsche Bundesbank (MiDi, 1999-2016) and PATSTAT, own calculations.

B

Appendix to Chapter 2

Top R&D Investors and the Net-Zero Transition

B.1 Data Construction

This paper uncovers patterns in green patenting of the Top 2000 Global R&D Investors, as identified by Hernández et al. (2020). The EC-JRC-OECD COR&DIP database v.3 contains information on the patents owned by these firms and filed between 2016 and 2018 (for details see Amoroso et al., 2021). Starting from the COR&DIP *Patent Portfolio* dataset, I retrieve additional patents applied for by the companies of interest and extend the analysis period to 2012-2019. This appendix explains in detail the data construction process.

B.1.1 Data Sources

COR&DIP v.3 database. The original *Patent Portfolio* dataset included in the COR&DIP v.3 database contains 1,210,977 patent applications owned by 1,541 top R&D investors. These are applications filed at the top 5 IP patent offices in the world: European Patent Office (EPO), United States Patent and Trademark Office (USPTO), Japan Patent Office (JPO), Korean Intellectual Property Office (KIPO), Chinese National Intellectual Property Administration (CNIPA). Note that the dataset consists only of *IP5 patent families* which are defined as families of patent applications that include at least one IP5 application and an application filed in one other patent office worldwide.

The IP portfolio of each firm includes patents filed by the corporate group. This means that the dataset also includes patents filed by the subsidiaries of the top R&D investors. In order to extend the panel dimension retroactively, it is necessary to link patent applicants with the R&D investors and their subsidiaries. The original dataset can be linked to the EPO patent database (PATSTAT Global) via the unique patent application ID or the respective patent number. However, the dataset does not contain the direct match between the top R&D investors and patent applicants. Therefore, the information provided is not sufficient to collect additional patent applications filed by the same set of firms before or after the 2016-2018 time period. Linking COR&DIP firm names and patent applicants is necessary for extending the dataset.

PATSTAT Global. The patent data used in this analysis is retrieved from PATSTAT Global, the EPO's worldwide patent database. I start by using information from PATSTAT Global version Spring 2021 in order to ensure data comparability with the COR&DIP database (Amoroso et al., 2021). First, I focus on information regarding patent applicants such that I can establish a link between COR&DIP firms and PATSTAT applicant IDs. In the current analysis, patent assignment is based on the applicant information retrieved from one publication document (see details in section B.1.2). Therefore, I abstract from patent transfers as I am not able to track changes in patent ownership with the data available.

In a second step, I collect additional patents and corresponding bibliographic information from PATSTAT Global version Spring 2023. I count each invention only once by using DOCDB patent families as the main unit of observation. I focus on patent families with the earliest filing year between 2012 and 2019. For the underlying PATSTAT version, 2019 is the last year where data is complete given that it can take up to 32 months from the first patent application and subsequent filings (see Dechezleprêtre, Ménière, and Mohnen (2017) for a discussion on international patent families).

My main analysis focuses on firms' triadic patent families, i.e. families that include applications filed at the EPO, USPTO and JPO. These are known to capture the highest-value inventions and allow for cross-country comparisons (Dernis & Khan, 2004; Guellec & Pottelsberghe de la Potterie, 2005). In addition, for robustness and comparability of results with Amoroso et al. (2021), I generate an indicator that distinguishes IP5 patent families, albeit with a slightly modified definition. In general, IP5 families include applications filed with at least two different patent offices, one of which being an IP5 office, as described above. In the specific case of families containing an EP application, I consider this criterion to be met if the second patent office is one outside the European Patent Convention. For example, if a family includes both an EP and a German patent application, the EP application would likely cover the German market, rendering the additional German patent application redundant when considering the international dimension of the patent protection. Therefore, I do not consider such cases to be equivalent to families with patents filed at the EPO and in, say, China or the United States. I identify green patents using the EPO's *Y02 tagging scheme*, a standardized classification meant to target technologies that mitigate or provide adaptation to climate change (Angelucci, Hurtado-Albir, & Volpe, 2018).

Besides firm-level patent counts, I additionally construct various patent-level indicators based on the number of forward citations, backward citations, and type of backward citations that are the basis for the comparison between green and non-green patents presented in Section 2.4. I count forward citations from subsequent USPTO applications at the family level within the first 5 years since the first filing data. I restrict attention to citations coming from a single patent office to account for the fact that citation patterns vary systematically across patent offices (Michel & Bettels, 2001). Additional patent-level indicators based on bibliographic data are retrieved from the OECD Patent Quality Indicators (version February 2022), i.e. the originality and radicalness indexes. The OECD computes these indicators for USPTO and EPO patent applications separately. Given our context, I restrict attention to the USPTO filings. Similar to Barbieri, Marzucchi, and Rizzo (2020), I assign each focal patent family the maximum value of the indicators among all USPTO applications within that family.[1]

Orbis IP. The IP portfolio of COR&DIP firms contains patents of both parent firms and their subsidiaries. In order to capture such ownership linkages, I complement PATSTAT information with data from Bureau can Dijk's Orbis Intellectual Property database (v. August 2020). This

[1]From all USPTO filings within an family, I remove utility models, design patents and all derived filings such as divisionals or continuations.

databases includes both information regarding patents' direct owner (equivalent to patent applicants in the case of patents that have never been transferred) but also on patents' indirect owners, i.e. higher-order owners such as the global ultimate owner of the patent applicant. In cases where patents are filed by subsidiaries of COR&DIP firms that do not have a common name, the string matching algorithm I employ in later steps will not establish a link between the two entities. In such cases, I check whether the name of the COR&DIP firms match with the name of any of the higher-order owners identified in Orbis IP.

I link patents in COR&DIP with Orbis IP's *Patent Matchings* table using patents' publication number, with a match rate of 60%. Roughly 40% of the matched applications have information regarding higher-order owners. I harmonize owner names by removing special characters, acronyms related to firms' legal form and common words.

Orbis Ownership. The approach mentioned above has one limitation. In the case of patents with multiple applicants, the *Patent Matchings* dataset in Orbis IP does not allow us to identify which higher-order owner is connected to which direct owner. This means that I would not be able to link COR&DIP firms with their right subsidiary in such cases. Therefore, for the subsample of patents with multiple applicants and potential indirect ownership linkages with COR&DIP firms, I complement the data in Orbis IP with firms' complete ownership structure in Bureau van Dijk's Orbis Historical Ownership database (v. August 2020). This source contains yearly snapshots of firms' ownership structure. Given the time dimension explored in this study, I combine ownership information from the 2017 and 2020 vintages.

From the list of patents that match with Orbis IP, I select those that have multiple applicants. I then match the applicant names in PATSTAT with the firm names in Orbis IP using string matching techniques.[2] I obtain a list of paired applicant person IDs and Orbis unique identifiers (BvD IDs). For each BvD ID paired, I retrieve the full list of shareholders in Orbis Ownership 2017 and 2020 vintages. The dataset contains linkages between parents and their subsidiaries, with information on the direct and indirect ownership percentage. There are different types of links included, not just those that reflect majority ownership. For the , I keep global ultimate owners, domestic ultimate owners and all other links where either the direct ownership percentage or the indirect ownership percentage is larger than 50%. I harmonize all shareholder names as explained above. Within the matching algorithm, I am then able to compare COR&DIP firm names with that of each shareholder and correctly identify parent × subsidiary pairs. Prior to string matching with COR&DIP firm names, I harmonize the names of all identified shareholders.

[2] I use the same matching procedure as explained in detail in section B.1.

B.1.2 Linking Top R&D Investors and PATSTAT Patent Applicants

The main goal from linking firm names and patent applicants is to retrieve all PATSTAT applicant person IDs that are associated with the top R&D investors of interest. This would include both the parent firms but also subsidiaries of these firms. Using the patents in the COR&DIP database that have already been associated with the R&D investors, I identify the firm × person ID links of interest. Starting from all firm × person ID combinations at the patent level, the main challenge of this exercise becomes removing false positive links that occur from patents that have either been assigned to multiple owners in COR&DIP or that have multiple applicants in PATSTAT. In what follows I elaborate on the main steps in identifying the correct links.

I retrieve all patent applicants for the applications included in the *Patent Portfolio* dataset. I use the PATSTAT version Spring 2021 similarly to Amoroso et al. (2021) in order to ensure comparability. I retrieve applicant information from the latest publication document with the number included in the original dataset. Before applying string matching techniques on firm and applicant names, I perform the following cleaning steps:

- I remove applicants that are not companies. These include inventors that are also patent applicants, private individuals, universities and research institutes or organizations, governmental institutions, non-profit organizations and hospitals.
- I harmonize firm and applicant names by removing special characters, common words (e.g., enterprise, corporation, group, etc.) and endings related to the legal form of the firm (e.g., LLC, Ltd., Inc., SARL, GmbH, etc.) taking into account different acronyms used across countries.

Keeping in mind the overarching goal of obtaining a list of patent applicant person IDs that are associated with each COR&DIP firm, I employ a two-step matching algorithm, as follows.

> **Step 1:** I separate patents that are assigned to only one COR&DIP firm and have only one remaining patent applicant after the cleaning steps above. This is a subsample of patents for which the firm × person ID link should be correct, since it is unique.
>
> **Step 2:** I analyse the combinations for the remaining patents that are either assigned to multiple firms or have multiple patent applicants. I ease the matching process by removing the already established links in Step 1. I then proceed with investigating the combinations that are left.

Within both Step 1 and 2, I first identify firm × applicant pairs where the strings match directly. I then parse each remaining firm name and harmonized patent applicant name (*psn_name* in PATSTAT) into single words. I remove a series of most common words such as *technologies*, *industries*, *pharmaceuticals*, etc. For each pairwise combination of words, I obtain a similarity score that ranges from 0 to 1, where a score of 1 is given to an exact match.[3] At the pair level, I compute the maximum similarity score obtained. This way, I can easily identify names that have at least parts of the string in common. I allow for potential typos that can decrease the score and consider that a pair is matched whenever the similarity score is > 0.6 in Step 1. For Step 2, I only consider a match once the maximum score is 1.[4]

There are several reasons for which some firm × applicant combinations remain unmatched. Either the patent has been co-filed by multiple entities and not all of them are top R&D investors, or the applicant is a subsidiary of a top R&D investor and does not have a common name, or simply the patent was not filed by top R&D investors and the link established in COR&DIP cannot be explained with the information available. Additional correct firm × applicant combinations can be identified whenever the second scenario occurs. For this, I rely on information from either Orbis IP or Orbis Ownership. In the case of patents with only one applicant, I retrieve higher order owners from Orbis IP and compare their names with the names of the firm. In the case of patents with multiple applicants, I retrieve shareholder information from Orbis Ownership and compare the name of each shareholder to the name of the R&D investor. This additional step is necessary in order to ensure that the correct applicant is linked to the firm of interest. Note that in Orbis IP, whenever there are multiple applicants, all potential higher-level owners are included without being able to infer which shareholder is connected to which applicant. Once all additional information is retrieved, the same string matching algorithm is applied as above. I only consider cases where the maximum similarity score is 1 at the firm x shareholder/owner pair. Figure B.1 summarizes the main steps of the matching algorithm.

Since I use different sources of data and different matching steps, it can be that the same firm name and harmonized patent applicant name appear both as matched and as unmatched. Once a pair is matched at least once, I consider it matched across all person IDs linked to the same name in the dataset. Note that applicant names are not fully disambiguated in PATSTAT. As mentioned above, for the purpose of this exercise, I use one form of disambiguation included in the database, namely the person's *psn_name*. At the end of the algoritm, I obtain 60,907 company × person_ID matched pairs.

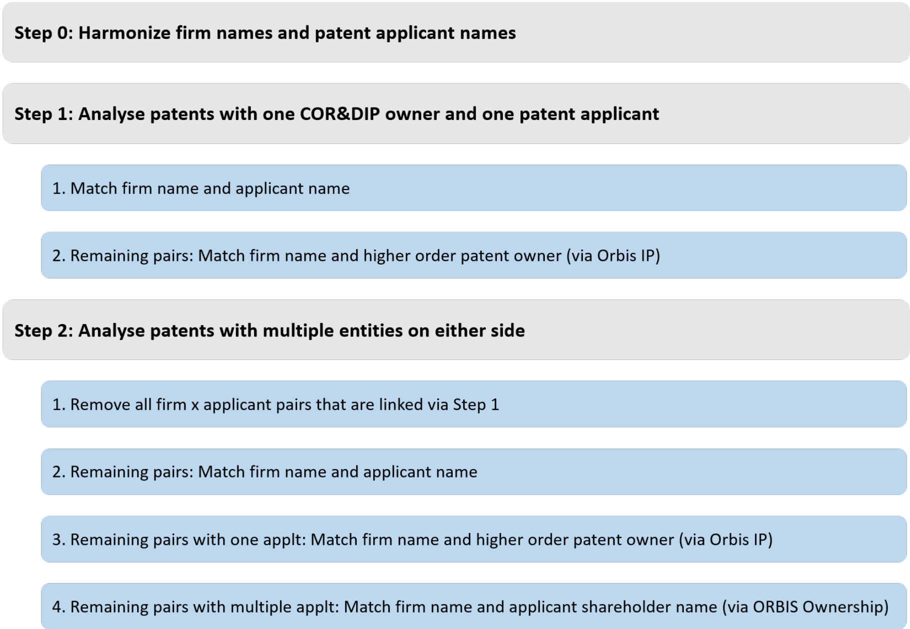
B.1.3 Retrieving Patents of Top R&D Investors

In order to capture patents of the top R&D investors from more recent years as well, I rely on PATSTAT version Spring 2023. First, I use the list of applicant person IDs identified through

[3]The similarity score is obtained using the default matching technique employed by the *matchit* `Stata` command which is based on a vectoral decomposition algorithm Raffo (2020).

[4]In this case, even a similarity score > 0.7 included a significant amount of false positives. I manually corrected scores for cases with obvious typos.

Figure B.1: Linking COR&DIP firms and PATSTAT patent applicants

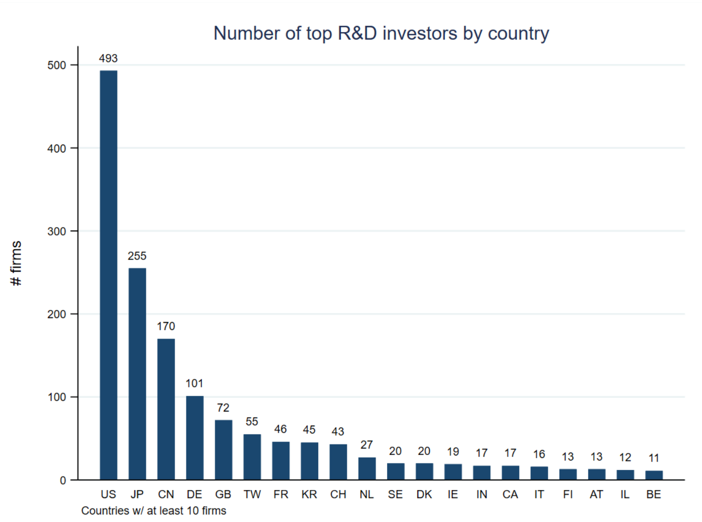


the matching algorithm to obtain the disambiguated psn IDs, which are generated for each PATSTAT version. Second, I retrieve all person IDs that are associated with the obtained disambiguated names. This allows me to capture applications filed by the same entities but that received different person IDs outside the 2016-2018 time-frame. For direct ownership links, I take into account the full list of person IDs obtained. However, for indirect links such as subsidiaries identified via Orbis IP and Orbis Ownership, I only include the original person IDs identified. I retrieve a total number of 404,489 person IDs linked to COR&DIP companies, whereas the starting number of IDs was only 59,271.[5] Finally, I extract all patents that are associated with the person IDs of interest based on the information recorded in the most recent publication (PATSTAT table *tls207*). During 2012-2019, the identified person IDs have filed 8.5 million patent applications that correspond to 3.7 million DOCDB patent families, out of which 270,787 are triadic patent families and 1,418,130 are IP5 patent families.

[5]Note that due to the COR&DIP algorithm that assigned patents to firms, there are person IDs that are linked to multiple top R&D investors. For those person IDs that are always linked to multiple investors, I maintain the same rule for patent assignment. However, in cases where person IDs are either linked to one or more companies, I prioritize the firm that the person ID is associated with most times.

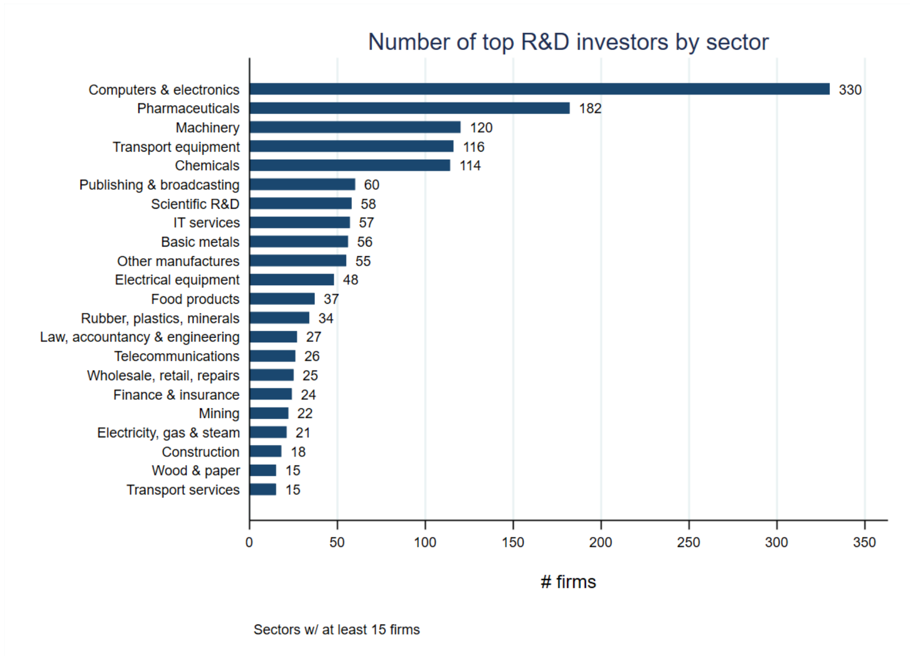
B.2 Additional Figures

Figure B.2: Geographical distribution of top R&D investors



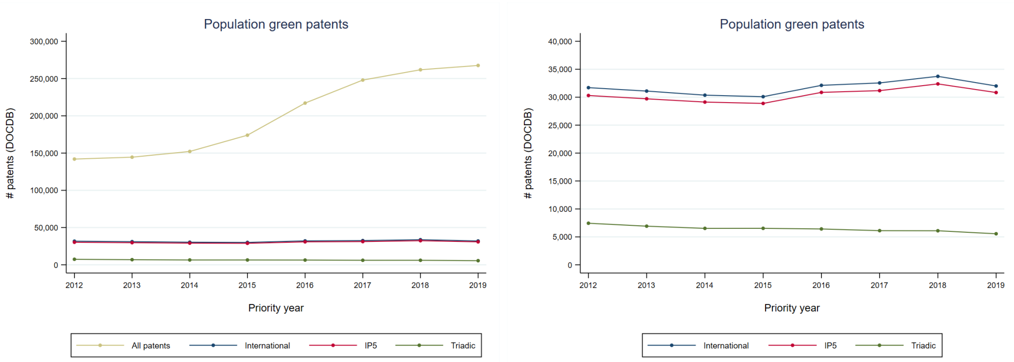
Notes: The figure includes only countries with at least 10 firms out of the total of 1,506 top R&D investors included in the baseline sample.

Figure B.3: Sectoral distribution of top R&D investors



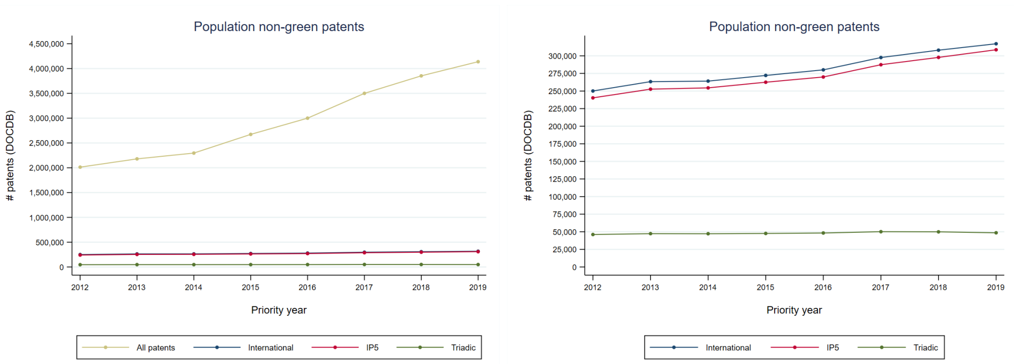
Notes: The figure includes only sectors (ISIC rev. 4) with at least 15 firms out of the total of 1,506 top R&D investors included in the baseline sample.

Figure B.4: Global development of green inventions



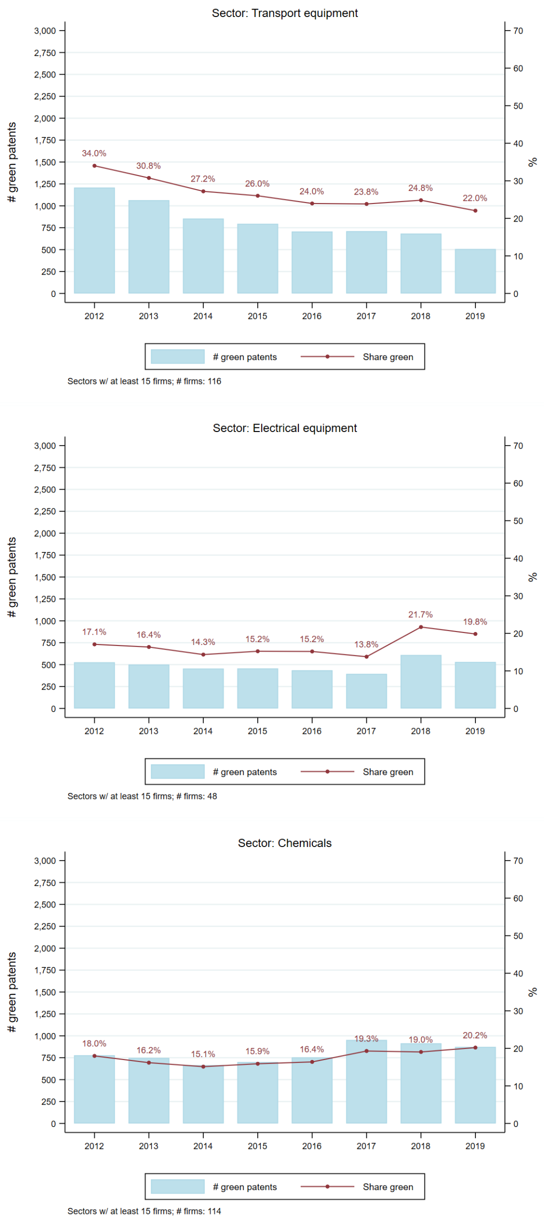
Notes: This figure illustrates the global development of green patenting over the period 2012-2019. All green patent families are included. International patent families include applications filed in at least two jurisdictions. Triadic patents families contain applications filed at the EPO, JPO and USPTO. IP5 patent families contains at least one application at the one of the main IP5 offices and one other application in a different jurisdiction. For readability, the panel on the right focuses only on patents filed in at least two jurisdictions.

Figure B.5: Global development of non-green inventions



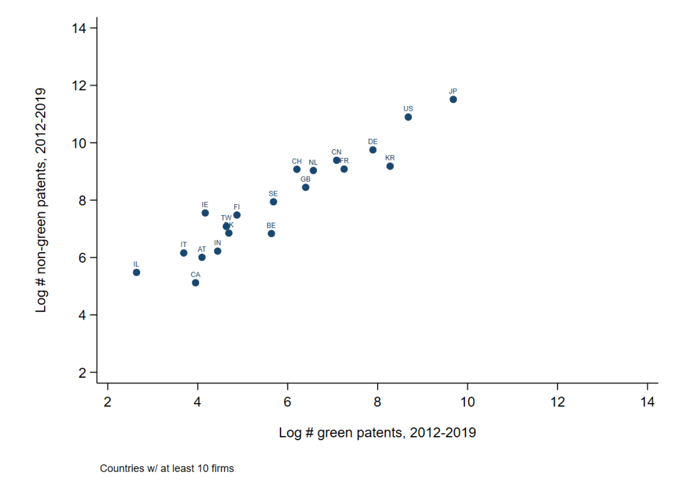
Notes: This figure illustrates the global development of non-green patenting over the period 2012-2019. All green patent families are included. International patent families include applications filed in at least two jurisdictions. Triadic patents families contain applications filed at the EPO, JPO and USPTO. IP5 patent families contains at least one application at the one of the main IP5 offices and one other application in a different jurisdiction. For readability, the panel on the right focuses only on patents filed in at least two jurisdictions.

Figure B.6: Evolution of green triadic patenting of top R&D investors in key sectors



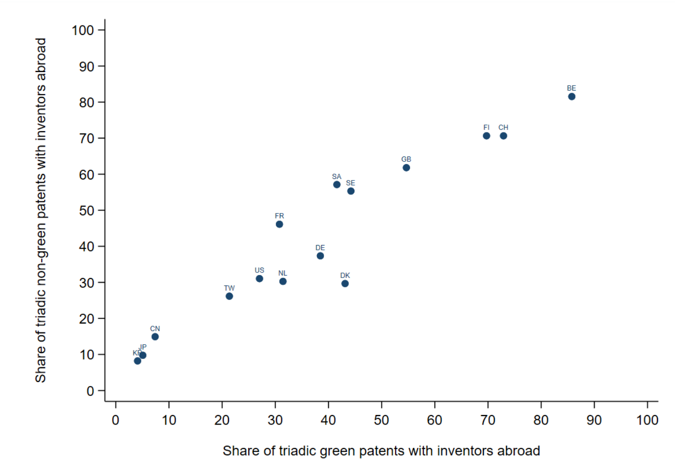
Notes: This figure illustrates the evolution of green patenting by top R&D investors in key sectors over the period 2012-2019. I include 'Transport equipment', 'Electrical equipment' and 'Chemicals', as sectors with high contributions towards climate change mitigation. I plot the absolute number of green triadic patents filed by top R&D investors each year (left axis) and the share of green patents relative to all filings (right-axis).

Figure B.7: Green and non-green patents of top R&D investors by headquarter country



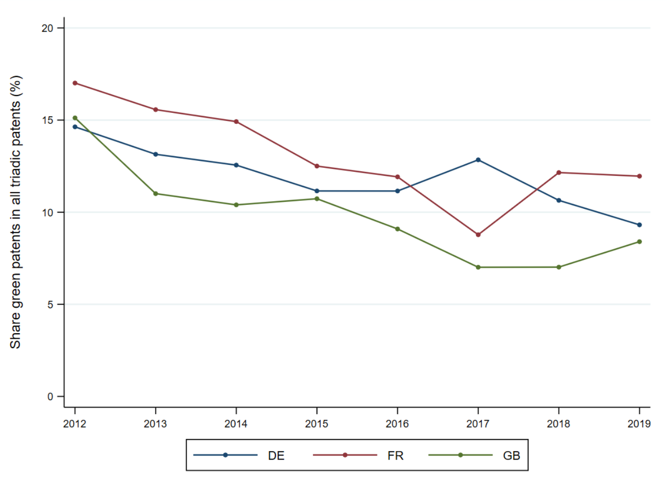
Notes: This figure plots the log number of green and non-green triadic patents over the period 2012-2019 attributed to each headquarter country. I include only countries with at least 10 firms. In the case of patents with multiple owners, I assign equal shares to each HQ country.

Figure B.8: Internationalization of R&D activities of top R&D investors



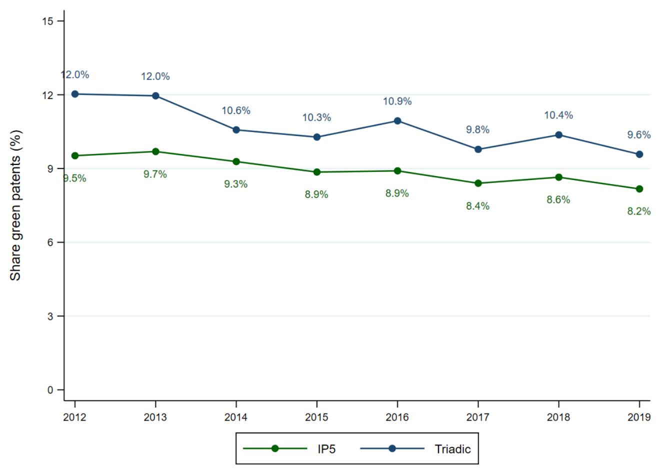
Notes: The figure illustrates the pattern of internationalization of R&D activities of top R&D investors by headquarter (HQ) country. For each country, I plot the share of green and non-green triadic patents that have inventors located abroad. The figure includes the top 15 economies with the highest number of green triadic patents filed by leading R&D investors during 2012-2019. I assign patents to countries based on the location of the firms that file them. In the case of patents with multiple owners, I assign equal shares to each HQ country. I distinguish between patents that have inventors located domestically (in the HQ country) and inventors located abroad.

Figure B.9: Evolution of the share of green patents by top European inventor country



Notes: The figure plots the annual share of green patents in all triadic patents of leading R&D investors invented in European countries during 2012-2019. I include the European countries with the highest number of green triadic patent. I assign patents to countries based on the location of inventors, using fractional counts whenever there are multiple inventor countries.

Figure B.10: Evolution of the average firm-level share of green patents



Notes: The figure plots the annual average share of green patents in all filed patents across top R&D investors. Both triadic (blue) and IP5 (green) patent family are considered.

B.3 Robustness: Green Patenting Activity of Top R&D Investors (IP5)

This Appendix presents recent trends in green patenting activity of top R&D investors based on the wider sample of IP5 patent families. This analysis complements the main results in Section 2.3 which are based on triadic patents families. It ensures comparability and extends the earlier work of Amoroso et al. (2021).

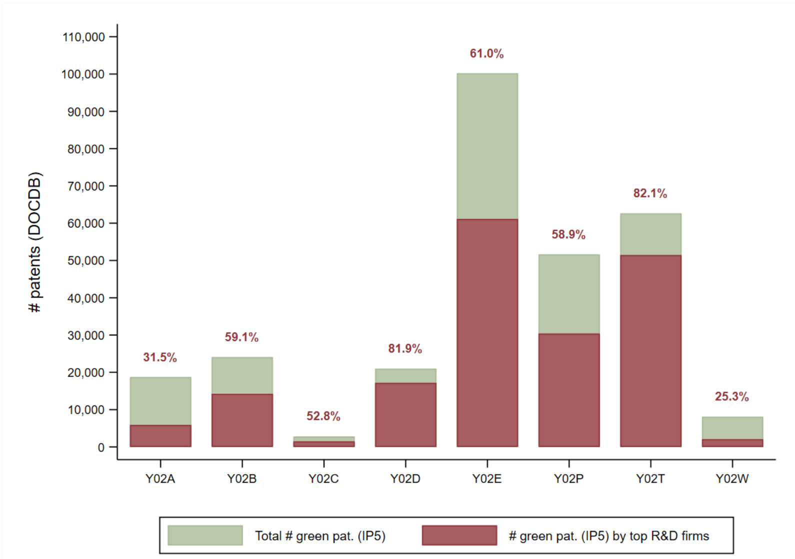
The patterns documented in Section 2.3 generally hold for the broader sample of IP5 patents. First, top R&D investors appear similarly relevant in terms of their contribution to global green technology development, resulting from the fact that they hold 64% of all green IP5 patent filed between 2012 and 2019 (Figure B.11). Second, the trend of a (relative) decline in green patenting holds also for this sample. On the one hand, we observe a similar decline in the contribution of top R&D investors, as shown in Figure B.12. On the other hand, the average annual growth rate of green IP5 patent families for top R&D investors was also negative at -0.8%, leading to a declining share of green patents (Figure B.13). The declining share in green patents after the adoption of the Paris Agreement in 2015 is confirmed in a regression analysis, controlling for sector and firm fixed effects as shown in Table B.2.

The strong sectoral concentration of green technology development among top R&D investors also holds for IP5 patents, with more than 50% of them being attributed to 'Transport equipment' (29%) and 'Computer & Electronics' (25%). Figure B.14 highlights the similar sectors with the highest specialization in green technology development: 'Electricity, gas & steam', 'Transport equipment' and 'Construction'. Differences to the main results are largely due to the sectors with general lower propensity to patent, reflected also in Figure B.15. The focus of firms active in 'Transport equipment' on developing green technologies is even more evident when looking at IP5 patent families compared to the main results.

While the relative distribution of green patents of top R&D investors across inventor countries is similar when looking at the IP5 sample (Figure B.17), the share of green patents within countries shows differences compared to the main results based on triadic patent families. As shown in Figure B.18, focusing on IP5 patent families reveals that most of the top inventor countries have a share of green patents between 10-15%, the only exception being Denmark with a share of 29%. Korea ranks second with a share of 14.5%, closely followed by Germany, France and the United Kingdom. Japan's relative focus on green technologies appears to be even lower, with a share of only 11%. China also shows a stronger upward trend in its share of green patents when IP5 patents are considered, as shown in Figure B.19. In general, using the IP5 patent families illustrates a cross-country pattern closer to Probst et al. (2021). This further strengthens the idea that the activity of top R&D investors is representative for the overall green technology development worldwide. Moreover, discrepancies in patterns between the main results and this analysis underscore the varied perspectives different patent-based metrics offer on countries' specialization in green technology. This emphasizes the need for caution when drawing conclusions for policy-making.

Results at the firm level suggests that, generally it is common for top R&D investors to develop technologies resulting in green patent applications. Over the span of 2012-2019, approximately 83% of these companies filed at least one green patent application, with 78% of them holding at least one green IP5 patent. Table B.1 presents the top 20 firms with the highest number of green IP5 patents, with Samsung Electronics being ranked first and followed by Panasonic and General Electric. similar to the main results, this top is also dominated by firms active in 'Transport Equipment' and 'Computer & Electronics' or firms originating in Japan, despite differences in the actual ranking.

Figure B.11: Contribution of top R&D investors to green tech development by Y02 category (IP5)

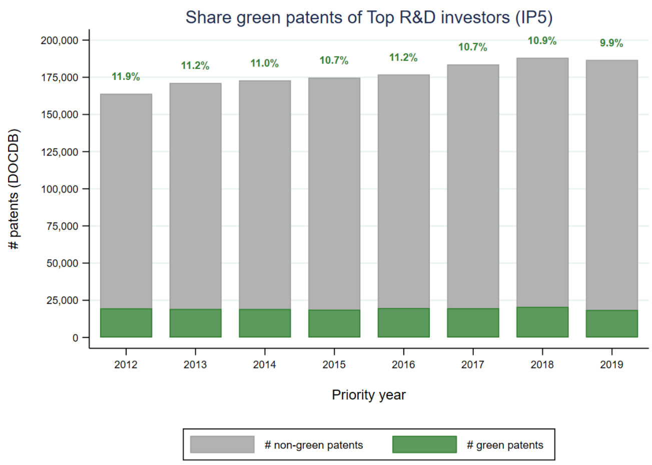


Notes: This figure illustrates the proportion of green IP5 patents originating from top R&D investors, by Y02 category. All green IP5 patent families with earliest filing year between 2012 and 2019 are included in the analysis. The red bars represent the number of green patents in each category that are filed by top R&D investors, while the labels above the bars represent the corresponding share.

Figure B.12: Yearly contribution of top R&D investors to green technology development (IP5)

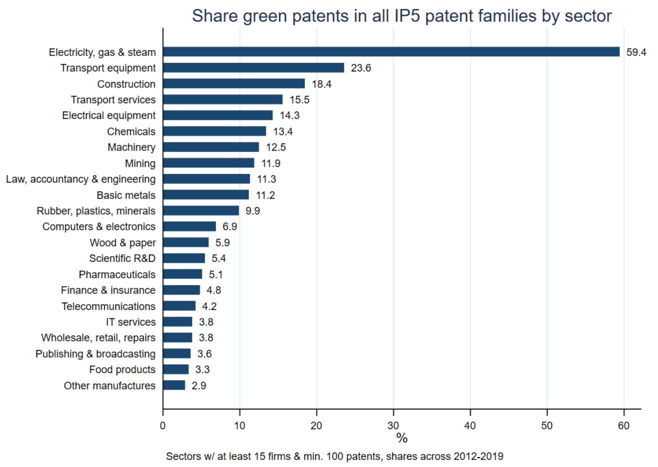
Notes: This figure plots the total number of green IP5 patent families with earliest filing year between 2012 and 2019 (left-axis) and the share that is accounted for by the top R&D investors (right-axis).

Figure B.13: Evolution of green patenting across top R&D investors (IP5)



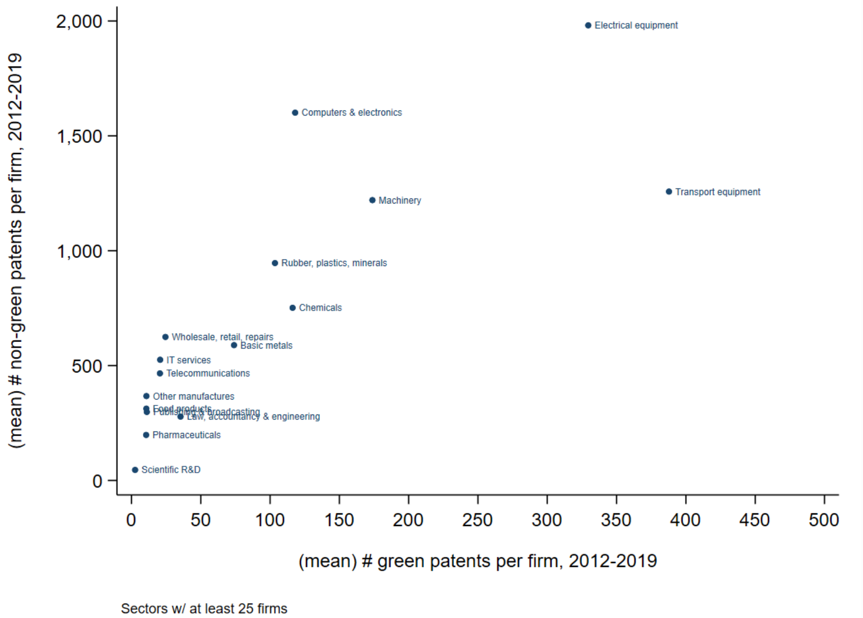
Notes: This figure plots the aggregate number of green and non-green IP5 patents of top R&D investors filed over the 2012-2019 period. The label on top of each bar represents the share of green patents in all IP5 patents of these firms.

Figure B.14: Share of green patents in all IP5 patents by sector



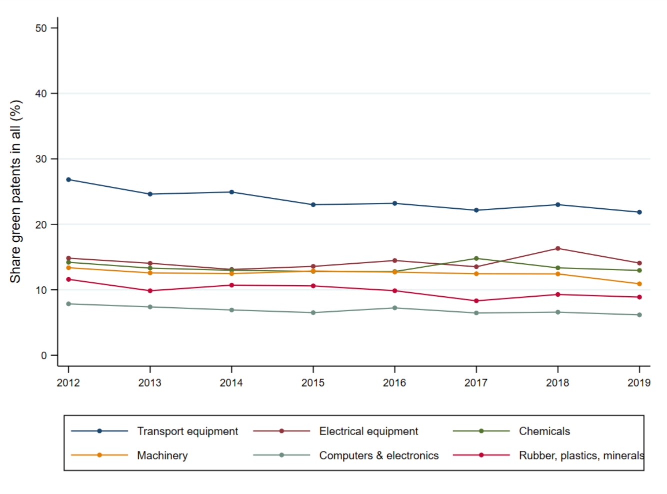
Notes: This figure plots the share of green patents in all IP5 patents filed by top R&D investors over 2012-2019, by sector. For readability, the figure includes only sectors with at least 15 firms and 100 patents. The sectoral classification is based on the aggregation of ISIC Rev.4 economic activities into 38 groups as in the COR&DIP database.

Figure B.15: Green and non-green IP5 patents per top R&D investor by sector



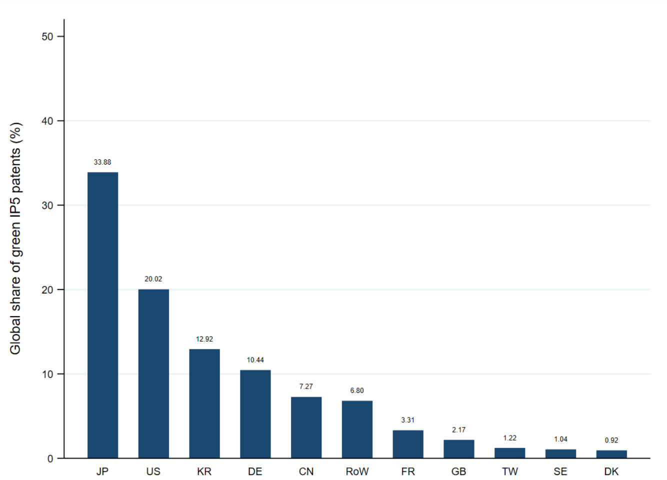
Notes: The figure plots the average number of non-green and green IP5 patents per firm over the 2012-2019 period by sector. For readability, I include only sectors that have at least 25 firms present in the sample.

Figure B.16: Evolution of the share of green patents in all IP5 patents by sector



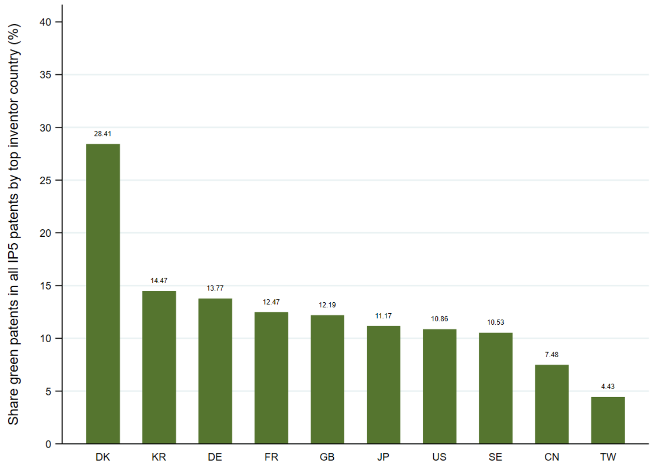
Notes: The figure plots the evolution of the share of green IP5 patents over time at sectoral level. For readability, I include only the main sectors where top R&D investors make significant contributions to climate change mitigation.

Figure B.17: Location of invention of green technologies (IP5)



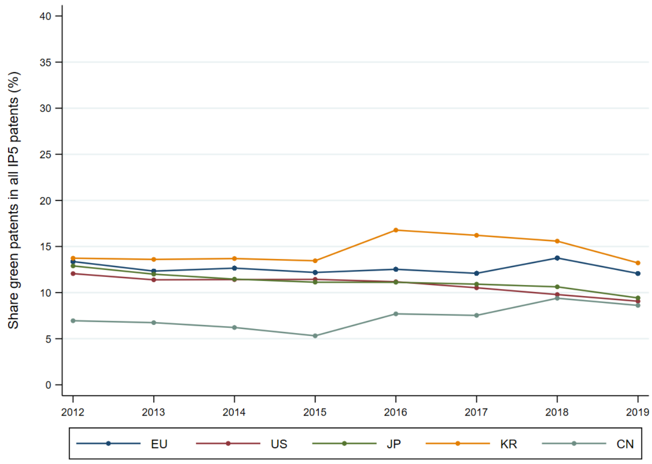
Notes: The figure includes top 10 economies where the patented green technologies of R&D investors originate from. I assign patents to countries based on the location of inventors, using fractional counts whenever there are multiple inventor countries. I compute the share of all IP5 green patents that are invented in each top country over the period 2012-2019. I label the aggregated share of all other countries as RoW (rest of the world).

Figure B.18: Share of green patents in all IP5 patents by inventor country



Notes: The figure includes top 10 economies where the patented green technologies of R&D investors originate from. I assign patents to countries based on the location of inventors, using fractional counts whenever there are multiple inventor countries. For each top inventor country, I compute the share of green patents in all IP5 patents invented there over the period 2012-2019.

Figure B.19: Evolution of the share of green patents in all IP5 patents by inventor location



Notes: The figure includes top economies where the patented green technologies of R&D investors originate from. I assign patents to countries based on the location of inventors, using fractional counts whenever there are multiple inventor countries. For each top inventor country, I compute the annual share of green patents in all IP5 patents invented there during 2012-2019. EU aggregates patents originating from European Union members represented in the baseline sample and the UK.

Table B.1: Top 20 R&D investors by the number of green IP5 patents

Company	**Country**	**Sector, ISIC4**	**All IP5 patents**	**Green patents**	**% Green**	**% of total green**
Samsung Electronics	KR	Computers & electr.	62,594	4,860	7.8	3.14
Panasonic	JP	Electrical equip.	21,916	4,513	20.6	2.91
General Electric	US	Machinery	15,728	4,442	28.2	2.87
LG Chem	KR	Chemicals	13,998	4,410	31.5	2.85
Toyota Motor	JP	Transport equip.	11,464	4,056	35.4	2.62
Ford Motor	US	Transport equip.	13,768	3,943	28.6	2.55
United Technologies	US	Transport equip.	14,856	3,940	26.5	2.54
Toyota Industries	JP	Transport equip.	9,607	3,668	38.2	2.37
Robert Bosch	DE	Machinery	18,107	3,277	18.1	2.12
Hitachi	JP	Electrical equip.	22,873	2,974	13.0	1.92
Siemens	DE	Machinery	15,673	2,804	17.9	1.81
Hyundai Motor	KR	Transport equip.	9,089	2,764	30.4	1.78
Volkswagen	DE	Transport equip.	3,432	2,456	71.5	1.59
Samsung SDI	KR	Computers & electr.	12,255	2,439	19.9	1.57
Honda Motor	JP	Transport equip.	8,255	2,385	28.9	1.54
General Motors	US	Transport equip.	8,842	2,221	25.1	1.43
Mitsubishi Electric	JP	Electrical equip.	20,251	2,166	10.7	1.40
Huawei Invest. & Holding Co	CN	Computers & electr.	35,460	2,090	5.9	1.35
Toshiba	JP	Computers & electr.	19,546	2,023	10.4	1.31
LG Electronics	KR	Computers & electr.	21,815	2,005	9.2	1.29
Total (Top 20)			359,527	63,436	17.6	40.95
Total (All R&D investors)			1,418,130	154,903	10.9	100.00

Notes: This table reports the top 20 R&D investors with the highest number of green IP5 patents filed between 2012 and 2019. Firms are ranked according to the total number of green IP5 patents. The last column displays the share of green patents by each individual firm out of the total number generated by top R&D investors, as shown in the bottom row.

Table B.2: Green patenting of top R&D investors post Paris Agreement (IP5)

	Log # green patents		Share green patents	
	(1)	(2)	(3)	(4)
Post Paris Agreement = 1	−0.012	0.020	-0.009^{***}	-0.009^{***}
	(0.014)	(0.014)	(0.003)	(0.002)
Constant	1.148^{***}	1.129^{***}	0.095^{***}	0.095^{***}
	(0.034)	(0.009)	(0.004)	(0.001)
Sector FE	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes
N	11,053	11,040	11,053	11,040
# firms	1,504	1,491	1,504	1,491

Notes: This table reports OLS regression results investigating whether top R&D investors changed their patenting behavior after the adoption of the Paris Agreement. The dependent variables are the log number of green IP5 patent (computed with offset 1) and the share of green patents in all IP5 patents filed over 2012-2019. The $PostParis$ binary indicator takes the value 1 as of 2015. Standard errors clustered at the firm level are presented in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

B.4 Robustness: Comparison Green and Non-green Patents (IP5)

Table B.3: Forward citations & Generality (IP5)

	Log # citations (3y)		Log # citations (5y)		Generality	
	(1)	(2)	(3)	(4)	(5)	(6)
Green	0.073***	0.072***	0.085***	0.084***	0.025***	0.026***
	(0.009)	(0.009)	(0.010)	(0.010)	(0.003)	(0.003)
Constant	0.719***	0.719***	0.866***	0.866***	0.247***	0.247***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Year FE	Yes	No	Yes	No	Yes	No
Tech class FE	Yes	No	Yes	No	Yes	No
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Tech class x Year FE	No	Yes	No	Yes	No	Yes
N	1,334,600	1,334,366	1,334,600	1,334,366	1,072,551	1,072,270
# firms	1,487	1,487	1,487	1,487	1,483	1,483

Notes: This table reports OLS regression results investigating whether the green and non-green patents of top R&D investors differ in terms of quality and generality. The dependent variables are the number of forward citations received from subsequent USPTO filings within 3 and 5 years from the patent's first filing and the generality index. Citation counts are log-transformed with offset 1. A patent scoring higher on the generality index is cited by patents belonging to more diverse technological fields. I include all green IP5 patents with earliest filing year during 2012-2019. Standard errors clustered at the firm level are presented in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.4: Knowledge sources (IP5)

	Originality		Radicalness		Avg. lag bw. citations		Share recent bw. citations	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Green	0.019***	0.019***	−0.002	−0.002	−0.877***	−0.861***	0.032***	0.031***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.073)	(0.071)	(0.003)	(0.003)
Constant	0.759***	0.759***	0.373***	0.373***	9.380***	9.376***	0.450***	0.450***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.008)	(0.008)	(0.000)	(0.000)
Year FE	Yes	No	Yes	No	Yes	No	Yes	No
Tech class FE	Yes	No	Yes	No	Yes	No	Yes	No
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tech class x Year FE	No	Yes	No	Yes	No	Yes	No	Yes
N	741,275	740,921	741,308	740,954	1,067,563	1,067,301	918,863	918,548
# firms	1,436	1,436	1,436	1,436	1,471	1,471	1,463	1,463

Notes: This table reports OLS regression results investigating whether the green and non-green patents of top R&D investors differ in terms of the knowledge base they draw upon. Column 1-4 focus on differences in patent's novelty, as indicated by how they score on the originality and radicalness indexes. A patent is more original if it relies on a wider pool of technology fields. A patent is more radical the more it references other technologies than its own. The dependent variables in Columns 5-8 are the patents' average lag of backward citations and the share of recent references in all backward citations. I consider a citation to be recent if the respective cited patent was filed at most 5 years prior to the focal patent. I include all IP5 triadic patents with earliest filing year during 2012-2019. Standard errors clustered at the firm level are presented in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.5: Developments after the adoption of the Paris Agreement (IP5)

	(1) Log # cit. (3y)	(2) Generality	(3) Originality	(4) Radicalness	(5) Avg. lag bw. cit.	(6) Share recent bw. cit.
Green=1	0.091***	0.023***	0.017***	−0.002	−0.843***	0.043***
	(0.013)	(0.003)	(0.003)	(0.003)	(0.073)	(0.004)
Post Paris=1	−0.341***	−0.088***	0.001	−0.027***	0.261***	0.004
	(0.012)	(0.002)	(0.001)	(0.002)	(0.042)	(0.002)
Green=1 × Post Paris=1	−0.020	0.005	0.004+	0.000	−0.057	−0.020***
	(0.012)	(0.003)	(0.002)	(0.004)	(0.063)	(0.005)
Constant	0.938***	0.299***	0.758***	0.386***	9.213***	0.447***
	(0.008)	(0.001)	(0.001)	(0.001)	(0.029)	(0.002)
Tech class FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
N	1,334,600	1,072,551	741,275	741,308	1,067,563	918,863
# firms	1,487	1,483	1,436	1,436	1,471	1,463

Notes: This table reports OLS regression results investigating whether top R&D investors changed the nature of their innovation after the adoption of the Paris Agreement. The $PostParis$ binary indicator takes the value 1 as of 2015. I include all green IP5 patents with earliest filing year during 2012-2019. Standard errors clustered at the firm level are presented in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

C

Appendix to Chapter 3

Government R&D Support for Young Innovative Companies

C.1 Data Construction

This appendix elaborates on the data sources used and the main steps taken in building the sample of analysis.

JEI. The French Central Agency for Social Security Organisms (ACOSS) assembles a panel with yearly information on young firms that declare a reduction in social security contributions based on the JEI status. We observe firms that benefited from the JEI scheme over the period 2004-2019. The dataset contains establishment-level information regarding the number of employees, amount of social security contributions paid, amount of social security exemptions received, amount of other type of exemptions received. For our analysis, we aggregate all variables at the SIREN level.

During 2004-2019, a total of 11,519 firms report the JEI (or JEU) status. From those, we remove firms for which the timing of their treatment seems implausible with respect to the age eligibility criterion. In other words, we remove firms that are either too old to have been eligible for the scheme in the first year they appear in the JEI dataset and those firms that are too old in the year they exit the scheme, i.e. firms at age nine or higher. This implies that we additionally remove firms for which the foundation date is missing in FARE/FICUS, as their age at entry and exit into the scheme cannot be identified. We compute the age of firm i in year t as $age_{it} = year_t - founding\_year_i + 1$. The age is therefore determined by calendar year. For instance, a firm founded in Oct 2010 would have age one in 2010 and age two in 2011. This is in line with how the French authorities assess the age of the firm, as confirmed in the legal case *SAS Open Pricer vs. URSSAF* in October 2018: "The status applies until the last day of the seventh calendar year following that of the creation of the company." We obtain a sample of 10,550 that is the baseline for our descriptive analysis in Section C.2.

FICUS/FARE. We obtain balance sheet information for the population of French firms from the FICUS/FARE datasets which comprise data from firms' tax declarations. We retrieve firms' founding date, address, closure date (if available), number of employees, turnover, total balance sheet, total expenses, economic activity (5-digit NAF classification [1]) and legal form. We rely on this source for both sample construction, but also for building outcome variables of interest. Specifically, these panels allow us to assess firms' age, as well as the SME and R&D eligibility criteria for the JEI scheme. Additionally, we determine firm closure based on the firm closure date and activity status (active, inactive, closed) available in FICUS/FARE. We consider a firm to be closed in any of the following cases: (1) a closure date is included in the data (FICUS), (2) the firm switches to an inactive or closed status (FARE), (3) the firm drops out of the data, suggesting that there are no other tax declarations filed. In the latter case, we

[1] *Nomenclature d'activités française.*

consider the closure year to be that of the last occurrence in the panel, i.e. the year of the last tax declaration. We group economic activities into 5 main sectors as described in Table C.1.

Enquête LIFI & LIFI. In order to be able to assess the independence criterion, we need to identify whether the firms of interest have financial links to other entities - French or foreign firms. We therefore build a panel dataset that contains all financial links between our focal firms and other firms in France or abroad. We rely on the Enquête LIFI and LIFI data sources that contain the structure of corporate groups active in France. This allows us to identify French subsidiaries of foreign multinational firms that are beyond the scope of our analysis. Additionally, the financial links give us information on firms' financing events and whether they become controlled by a different entity, which would imply a loss of eligibility for the JEI status. Furthermore, we identify acquisitions using information regarding firms' financial links. We track instances where there is a transition from a lower ownership percentage to 100%, which subsequently stands as the final recorded value in the LIFI panel. We ensure that we are not simply capturing internal restructuring events by conditioning on the owner not being a holding company (NAF 6420Z) or head office (NAF 7010Z). We consider the acquisition year the one in which the change to 100% ownership is documented.

DADS Salariés. We use employer-employee data from DADS (Déclaration Annuelle de Données Sociales) in order to generate a proxy for the R&D personnel for which exemptions could be granted. We rely on the *DADS Salariés* tables that include the primary job of each employee. We categorize employees as R&D related using the detailed nomenclature of occupations and socio-professional categories (PCS) available in the dataset. We consider engineers, technicians and project managers in tech-related fields. Table C.2 contains the full list of occupations included in our R&D personnel measure. For robustness, we generate an additional measure based on the list of occupations considered by Harrigan, Reshef, and Toubal (2023) when identifying tech workers.

GECIR. French firms that receive R&D tax credit are required to declare their R&D expenditures on their tax declaration. The GECIR dataset produced by the French Ministry of Finance contains information reported on the 2069 tax declaration related to the R&D tax credit. We retrieve declaring firms' R&D expenditures over the 2008-2019 period. This allows us to compute the R&D ratio necessary for assessing JEI eligibility.

INPI and PATSTAT. We link firms to patent applications filed at the French Patent Office (INPI) up to 2017 using firms' SIREN code. In our analysis, we focus on patent families that contain at least one application filed at the INPI. In a second step, we collect all other patent applications within the same identified DOCDB patent families from PATSTAT Global (v. spring 2023). We rely on this two-step approach in order to circumvent the absence of a cross-walk between

PATSTAT patent applicants and firm SIREN numbers. Furthermore, based on bibliographic information we compute the number of forward citation each patent family receives from subsequent applications at the European Patent Office within 5 years since its first filing.

Crunchbase. We identify additional firm acquisitions using the online venture capital platform Crunchbase (extracted in April 2023). In absence of a direct link between the sources, we match firms in our baseline sample with those that appear to have been acquired in Crunchbase based on their names. First, I harmonize firm names by removing special characters, common words (e.g., enterprise, corporation, group, etc.) and endings related to the legal form of the firm (e.g., SA, SARL, etc.). We obtain a similarity score between the full strings using the default matching technique employed by the *matchit* `Stata` command, which is based on a vectoral decomposition algorithm Raffo (2020). By cross-checking firm addresses, we noticed that even minor name differences that lead to high similarity scores were generally causing false positive errors. Therefore, we match firms in both sources based on similarity scores of value one, strings being exactly the same. We identify few additional cases with lower similarity scores by manually inspecting the data. Whenever acquisitions are observed in both Crunchbase and LIFI but the timing differs, we consider the year of acquisition to be the earliest between the two sources.

Table C.1: Sectoral classification based on the NAF rév. 2 (2008)

Sector	NAF rév. 2 code
ICT	58-63
Life sciences	721
Engineering and other technical & scientific activities	7112B, 7120B, 7490B
Manufacturing	10-33
Other	all remaining codes

Notes: We group firms into five main sectors, based on the French economic activity classification (NAF rév. 2 (2008). We distinguish between sectors represented in the startup space that we are interested in analysing. We include *Engineering and other technical and scientific activities* as a separate sector given the prevalent occurrence of the selected activity codes among the R&D related startups that we analyse.

Table C.2: Classification R&D employees

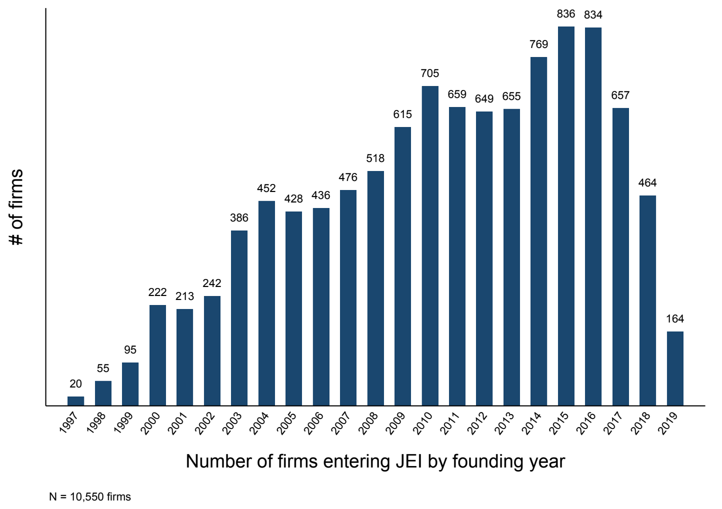
PCS	Description (EN)
342e	Public research scientists
380a	Technical managers in large companies
383a	Design, research and development engineers and managers in electricity and electronics
383b	Manufacturing engineers and managers in electrical and electronic equipment
384a	Mechanical and metalworking design, research and development engineers and managers
384b	Mechanical and metalworking manufacturing engineers and managers
385a	Design, research and development engineers and managers in the processing industries (agri-food, chemicals, metallurgy, heavy materials)
385b	Manufacturing engineers and managers in the processing industries (agri-food, chemicals, metallurgy, heavy materials)
386a	Design, research and development engineers and managers in other industries (printing, soft materials, furniture and wood, energy, water)
387f	Environmental engineers and managers
388a	IT design, research and development engineers and managers
388c	IT project managers, IT managers
388d	Technical and commercial engineers and managers in IT and telecommunications
388e	Telecommunications engineers and managers
473b	Research and development and manufacturing methods technicians in electricity, electromechanics and electronics
473c	Manufacturing and quality control technicians in electricity, electromechanics and electronics
474b	Mechanical engineering and metalworking research, development and manufacturing methods technicians
474c	Mechanical engineering and metalworking manufacturing and quality control technicians
475a	Research, development and production methods technicians in the processing industries
475b	Production and quality control technicians in the processing industries
477d	Environmental and pollution treatment technicians
478a	Computer design and development technicians
478b	IT production and operations technicians
478d	Telecommunications and network computing technicians
479a	Technicians in public research or teaching laboratories

Notes: We use the PCS (*Professions et Catégories Socioprofessionnelles*) classification to identify employees related to R&D activities subject to social security exemptions granted by the JEI status. We consider engineers, technicians and project managers in tech-related fields. Own classification. *Source:* PCS (INSEE, 2003), available at: https://www.insee.fr/fr/information/2400059.

C.2 Descriptive Analysis of JEI Beneficiaries

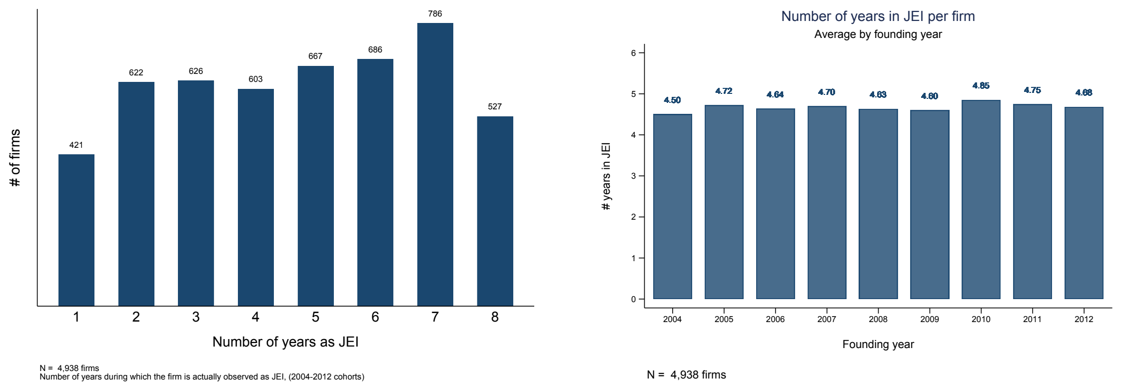
This appendix presents additional descriptive findings for the overall sample of JEI firms that receive benefits during 2004-2019.

Figure C.1: JEI firms by founding year



Notes: This figure plots the number of JEI firms receiving benefits over 2004-2019 by firm founding year.

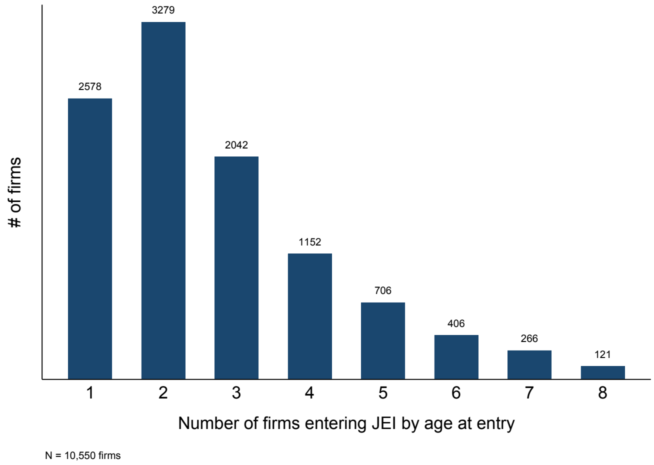
Figure C.2: Length of JEI participation



A. B.

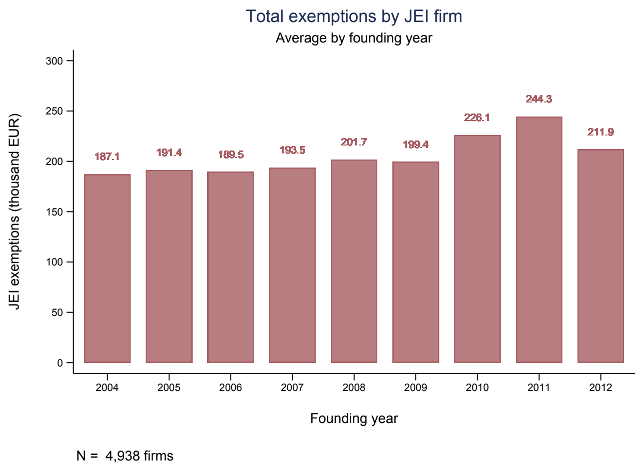
Notes: This figure shows patterns in the length of JEI participation. We focus on the 2004-2012 cohort to account for truncation and allow firms to be observed for the full 8-year window for which the JEI status can be applied. Panel A plots the number of JEI firms by the number of years they receive benefits. Panel B plots the average number of years in JEI by cohort.

Figure C.3: Firms with JEI status by age at entry



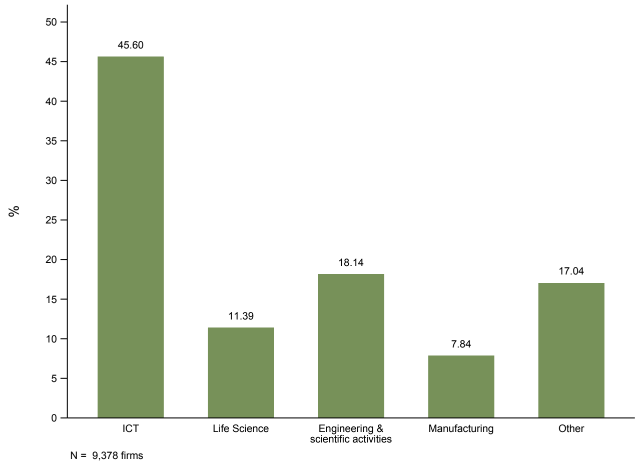
Notes: This figure plots the number of JEI beneficiaries (2004-2019) by age at entry into the scheme.

Figure C.4: Average JEI exemptions across firms by cohort



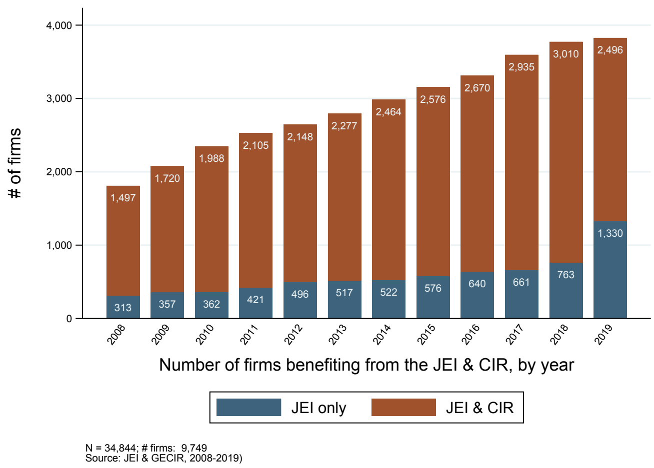
Notes: This figure plots the average amount of exemptions per JEI beneficiary by firm founding year. We restrict analysis to the 2004-2012 cohort to account for truncation and allow firms to be observed for the full 8-year window for which the JEI status can be applied.

Figure C.5: Sectoral distribution of JEI firms



Notes: We group firms into five main sectors based on their primary economic activity, as described in Table C.1. We take into account the earliest available NAF code for each JEI firm. The figure plots the corresponding sectoral distribution of JEI beneficiaries (2004-2019).

Figure C.6: JEI beneficiaries receiving an R&D tax credit



Notes: This figure plots the number of firms benefiting from the JEI scheme by year, distinguishing between those firms that additionally benefit from the French R&D tax credit (CIR) and those that do not. We focus on those firms receiving benefits during 2008-2019 period only due to GECIR data availability.

Table C.3: JEI exemptions by firm age

Panel A. JEI exemptions

	N	Mean	SD
Age = 1	2,578	11,073	14,121
Age = 2	5,539	25,925	31,091
Age = 3	6,695	37,601	42,517
Age = 4	6,650	46,264	51,462
Age = 5	5,993	53,244	63,158
Age = 6	5,111	59,195	69,227
Age = 7	4,222	64,676	76,826
Age = 8	3,341	68,665	88,582

Panel B. Share exemptions in total wage-related expenses

	N	Mean	SD
Age = 1	1,540	0.1677	0.0708
Age = 2	4,053	0.1950	0.0776
Age = 3	5,248	0.1705	0.0643
Age = 4	5,329	0.1620	0.0656
Age = 5	4,793	0.1503	0.0650
Age = 6	4,145	0.1397	0.0672
Age = 7	3,509	0.1344	0.0692
Age = 8	2,754	0.1249	0.0704

Notes: Panel A reports average JEI exemptions (€) by firm age. All 10,550 JEI beneficiaries observed over 2004-2019 are included. An average firm aged 4 received an exemption of approximately €46,000 from paying social security contributions in the respective fiscal year. Panel B shows the proportion of exemptions relative to the total wage-related expenses incurred by beneficiaries in a given year. For this analysis, we focus on the subset of firms for which wage information is available. We retrieve total wage expenses and social security charges from firms' income statement information in FARE (2009-2019). On average, a JEI beneficiary saves 15.7% of its total wage bill in a given year. *Data sources:* JEI & FARE, own calculations.

C.3 Additional Results

Table C.4: JEI participation and firm acquisition (probit)

Variable	(1) Probit *Acquired*	(2) Probit *Acquired*	(3) Probit *Acquired*
JEI	0.360***	0.348***	0.358***
	(0.072)	(0.073)	(0.074)
from Paris (0/1)	0.207**	0.169*	0.154*
	(0.071)	(0.071)	(0.073)
ICT (0/1)		0.229***	
		(0.068)	
Life Science (0/1)		−0.165	
		(0.131)	
Constant	−1.946***	−2.010***	−1.784***
	(0.120)	(0.124)	(0.125)
Founding year FE	Yes	Yes	Yes
Sector FE	No	No	Yes
N	4,433	4,433	4,433
# JEI	2,677	2,677	2,677

Notes: The table reports results from a probit estimation investigating the relationship between firms' probability to be acquired and JEI participation. The outcome variable is a binary indicator that takes the value 1 if a firms has been acquired within its first 12 years since establishment. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C.5: JEI participation and firm acquisition (limited sample)

Variable	(1) OLS *Acquired*	(2) OLS *Acquired*	(3) Probit *Acquired*	(4) Probit *Acquired*
JEI	0.040***	0.032***	0.041***	0.418***
	(0.007)	(0.006)	(0.007)	(0.078)
from Paris (0/1)	0.028**	0.020*	0.023*	0.184*
	(0.010)	(0.009)	(0.010)	(0.077)
Constant	0.023***	0.018***	0.023***	−1.850***
	(0.005)	(0.005)	(0.005)	(0.129)
Founding year FE	Yes	Yes	Yes	Yes
Sector FE	No	Yes	No	Yes
N	3,787	3,787	3,787	3,787
# JEI	2,244	2,244	2,244	2,244

Notes: The table investigates the relationship between firms' probability to be acquired and JEI participation. The outcome variable is a binary indicator that takes the value 1 if a firms has been acquired within its first 12 years since establishment. We include only the 2006-2011 cohorts to alleviate potential concerns regarding truncation. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C.6: JEI participation and firm acquisition (8y-window)

Variable	(1) OLS *Acquired (8y)*	(2) OLS *Acquired (8y)*	(3) OLS *Acquired (8y)*
JEI	0.017^{***}	0.016^{***}	0.016^{***}
	(0.004)	(0.004)	(0.004)
from Paris (0/1)	0.019^{**}	0.017^{**}	0.016^{*}
	(0.006)	(0.006)	(0.006)
ICT (0/1)		0.015^{**}	
		(0.005)	
Life Science (0/1)		-0.004	
		(0.007)	
Constant	0.007^{**}	0.003	0.008^{**}
	(0.003)	(0.003)	(0.003)
Founding year FE	Yes	Yes	Yes
Sector FE (5 groups)	No	No	Yes
N	4,433	4,433	4,433
# JEI	2,677	2,677	2,677

Notes: The table investigates the relationship between firms' probability to be acquired and JEI participation. The outcome variable is a binary indicator that takes the value 1 if a firms has been acquired within its first 8 years since establishment. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C.7: JEI participation and firm acquisition (age at acquisition)

Variable	(1) *Age at acquisition*	(2) *Age at acquisition*
JEI	0.077	0.197
	(0.372)	(0.383)
from Paris (0/1)	-0.721^{+}	-0.698^{+}
	(0.375)	(0.386)
Constant	9.651^{***}	9.554^{***}
	(0.309)	(0.322)
Founding year FE	Yes	Yes
Sector FE (5 groups)	No	Yes
N	256	256
# JEI	193	193
Mean acq. age	9.49	9.49

Notes: The table explores whether there is a relationship between JEI participation and the timing of acquisition, among those firms found to be acquired within the observation period. We identify acquisition based on data regarding firm linkages (LIFI) and information from Crunchbase (up to 2022). The outcome variable is a binary indicator that takes the value 1 if a firms has been acquired throughout the observation period. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C.8: JEI participation and the probability of patenting (ages 5-8)

	Any patent (0/1)				Any EP patent (0/1)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Age = 5	Age = 6	Age = 7	Age = 8	Age = 5	Age = 6	Age = 7	Age = 8
JEI	0.038^{***}	0.037^{***}	0.029^{***}	0.023^{***}	0.024^{***}	0.028^{***}	0.020^{***}	0.024^{***}
	(0.006)	(0.006)	(0.005)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)
from Paris (0/1)	-0.020^{***}	0.002	-0.010^{*}	−0.006	-0.008^{+}	−0.004	-0.007^{+}	−0.005
	(0.006)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)
Constant	0.027^{***}	0.020^{***}	0.014^{***}	0.011^{***}	0.012^{***}	0.009^{***}	0.006^{**}	0.002
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Founding year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4,425	4,416	3,865	3,337	4,425	4,416	3,865	3,337
# JEI	2,670	2,663	2,298	1,958	2,670	2,663	2,298	1,958

Notes: This table reports regression results from estimating a linear probability model investigating the relationship between firms' likelihood of patenting and JEI participation. The outcome variables are binary indicators that take the value 1 if the firm files a patent in a given year (Columns 1-2) and an EP patent (Columns 3-4). EP patents labels patent families that include an EP application. The year is determined by the first filing date within each patent family. The outcome variables are coded with 0 values for closed firms to account for survival bias. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C.9: JEI participation and patent output (ages 5-8)

Panel A: All patents

	Log # patents			
	(1) Age = 5	(2) Age = 6	(3) Age = 7	(4) Age = 8
JEI	0.034***	0.030***	0.028***	0.026***
	(0.006)	(0.006)	(0.005)	(0.005)
from Paris (0/1)	−0.012*	0.006	−0.010*	−0.008
	(0.006)	(0.007)	(0.005)	(0.005)
Constant	0.022***	0.019***	0.011***	0.009***
	(0.004)	(0.004)	(0.003)	(0.003)

Panel B: EP patents

	Log # EP patents			
	(1) Age = 5	(2) Age = 6	(3) Age = 7	(4) Age = 8
JEI	0.020***	0.023***	0.019***	0.022***
	(0.004)	(0.004)	(0.004)	(0.004)
from Paris (0/1)	−0.006+	−0.001	−0.007*	−0.004
	(0.004)	(0.005)	(0.003)	(0.004)
Constant	0.009***	0.007**	0.005*	0.002
	(0.002)	(0.002)	(0.002)	(0.002)

Panel C: Citation-weighted patents

	Log # cit. weighted patents			
	(1) Age = 5	(2) Age = 6	(3) Age = 7	(4) Age = 8
JEI	0.020***	0.017***	0.015**	0.020***
	(0.006)	(0.005)	(0.005)	(0.005)
from Paris (0/1)	−0.010+	0.002	−0.009**	−0.006
	(0.006)	(0.005)	(0.003)	(0.004)
Constant	0.013***	0.008**	0.008*	0.001
	(0.003)	(0.003)	(0.003)	(0.002)
Founding year FE	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes
N	4,425	4,416	3,865	3,337
# JEI	2,670	2,663	2,298	1,958

Notes: This table reports OLS regression results from investigating the relationship between firms' innovation output as reflected by their patenting activity and JEI participation. The outcome variables are the log number of patents, the log number of EP patents, and the log number of citation-weighted patents, respectively. All dependent variables are computed as $log(1+\#\text{patents})$. EP patents labels patent families that include an EP application. Patents are weighted by the number of forward citations received from subsequent EP applications within 5 years since their first filing. The outcome variables are coded with 0 values for closed firms to account for survival bias. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C.10: JEI participation and type of employment (tech workers)

	Log # non-tech employees				Log # tech employees			
	(1) Age = 5	(2) Age = 6	(3) Age = 7	(4) Age = 8	(5) Age = 5	(6) Age = 6	(7) Age = 7	(8) Age = 8
JEI	0.092^{**}	0.120^{***}	0.129^{***}	0.105^{**}	0.208^{***}	0.207^{***}	0.241^{***}	0.235^{***}
	(0.031)	(0.034)	(0.036)	(0.037)	(0.031)	(0.034)	(0.036)	(0.038)
from Paris (0/1)	0.152^{***}	0.188^{***}	0.160^{***}	0.188^{***}	0.000	−0.006	0.054	0.003
	(0.036)	(0.040)	(0.043)	(0.045)	(0.036)	(0.040)	(0.043)	(0.045)
Constant	1.175^{***}	1.104^{***}	1.028^{***}	0.956^{***}	0.886^{***}	0.908^{***}	0.878^{***}	0.878^{***}
	(0.025)	(0.028)	(0.029)	(0.030)	(0.026)	(0.028)	(0.029)	(0.030)
Founding year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	3,839	3,735	3,709	3,685	3,839	3,735	3,709	3,685
# JEI	2,386	2,313	2,293	2,272	2,386	2,313	2,293	2,272

Notes: This table explores the relationship between JEI participation and employment, distinguishing between tech and non-tech workers. The dependent variables are the log transformed with offset 1. They are coded with 0 values for closed firms to account for survival bias. We identify R&D employees based on information regarding employees' occupations following the definition of Harrigan, Reshef, and Toubal, 2023. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C.11: JEI participation and type of employment (ages 9-12)

	Log # non-R&D employees				Log # R&D employees			
	(1) Age = 9	(2) Age = 10	(3) Age = 11	(4) Age = 12	(5) Age = 9	(6) Age = 10	(7) Age = 11	(8) Age = 12
JEI	0.065	0.049	0.003	−0.015	0.287^{***}	0.203^{***}	0.150^{**}	0.128^{*}
	(0.043)	(0.049)	(0.055)	(0.069)	(0.038)	(0.043)	(0.050)	(0.063)
from Paris (0/1)	0.162^{**}	0.146^{*}	0.151^{*}	0.141	0.026	0.012	0.038	0.056
	(0.053)	(0.060)	(0.070)	(0.088)	(0.047)	(0.053)	(0.063)	(0.081)
Constant	1.005^{***}	0.942^{***}	0.894^{***}	0.913^{***}	0.657^{***}	0.673^{***}	0.667^{***}	0.680^{***}
	(0.034)	(0.038)	(0.042)	(0.054)	(0.029)	(0.033)	(0.038)	(0.048)
Founding year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	3,140	2,559	1,949	1,315	3,140	2,559	1,949	1,315
# JEI	1,875	1,510	1,117	768	1,875	1,510	1,117	768

Notes: This table explores the relationship between JEI participation and employment, distinguishing between R&D and non-R&D staff. The dependent variables are the log number of non-R&D workers (Columns 1-4) and the log number of R&D-workers (Columns 5-8) evaluated at ages 9-12. They are coded with 0 values for closed firms to account for survival bias and computed with offset 1. We identify R&D employees based on information regarding employees' occupations and not on firm self-reported values. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C.12: JEI participation and sales (ages 1-8)

	Log sales							
	(1) Age = 1	(2) Age = 2	(3) Age = 3	(4) Age = 4	(5) Age = 5	(6) Age = 6	(7) Age = 7	(8) Age = 8
JEI	-1.403^{***}	-1.277^{***}	-0.978^{***}	-0.380^{**}	-0.409^{**}	0.008	0.089	0.143
	(0.217)	(0.147)	(0.126)	(0.121)	(0.143)	(0.162)	(0.179)	(0.192)
from Paris (0/1)	0.119	−0.103	−0.060	−0.006	0.000	0.093	0.132	0.091
	(0.240)	(0.182)	(0.157)	(0.149)	(0.171)	(0.190)	(0.209)	(0.226)
Constant	2.821^{***}	3.544^{***}	4.095^{***}	4.202^{***}	3.727^{***}	3.021^{***}	2.394^{***}	1.836^{***}
	(0.117)	(0.098)	(0.090)	(0.093)	(0.112)	(0.130)	(0.143)	(0.153)
Founding year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE (5 groups)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,769	3,381	3,964	4,128	4,397	4,394	4,390	4,360
# JEI	599	1,652	2,216	2,378	2,647	2,649	2,646	2,627

Notes: This table examines the relationship between JEI participation and sales measured at ages 1-8. The dependent variables is log sales computed with offset 1. Values for closed firms are coded with 0 to account for survival bias. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.13: JEI participation and employment (ages 1-8)

	Log # employees							
	(1) Age = 1	(2) Age = 2	(3) Age = 3	(4) Age = 4	(5) Age = 5	(6) Age = 6	(7) Age = 7	(8) Age = 8
JEI	0.089^{*}	0.120^{***}	0.168^{***}	0.200^{***}	0.181^{***}	0.168^{***}	0.185^{***}	0.186^{***}
	(0.035)	(0.026)	(0.025)	(0.026)	(0.030)	(0.034)	(0.036)	(0.038)
from Paris (0/1)	0.036	0.048	0.059^{*}	0.069^{*}	0.098^{**}	0.110^{**}	0.103^{*}	0.092^{*}
	(0.042)	(0.030)	(0.028)	(0.030)	(0.035)	(0.039)	(0.042)	(0.045)
Constant	0.769^{***}	1.094^{***}	1.251^{***}	1.370^{***}	1.349^{***}	1.296^{***}	1.203^{***}	1.119^{***}
	(0.027)	(0.021)	(0.021)	(0.022)	(0.025)	(0.027)	(0.029)	(0.030)
Founding year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,394	3,168	3,876	4,007	4,149	4,153	4,226	4,248
# JEI	616	1,664	2,230	2,384	2,539	2,544	2,567	2,571

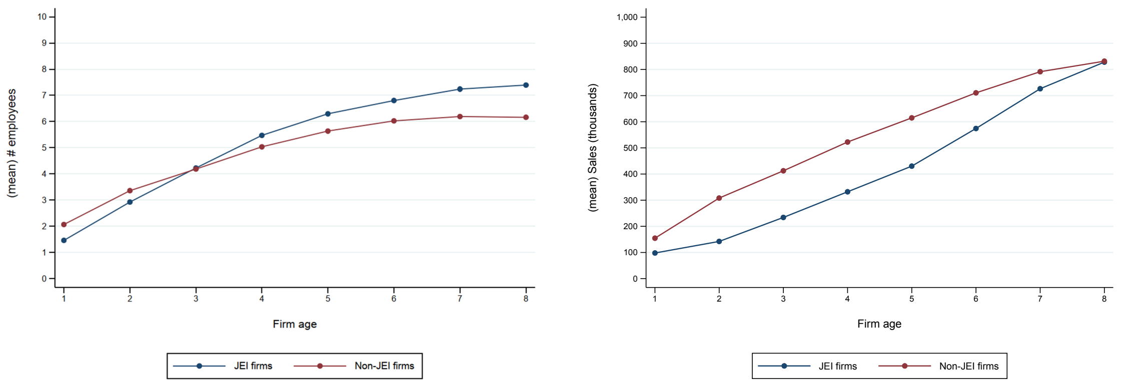
Notes: This table examines the relationship between JEI participation and employment measured at ages 1-8. The dependent variables is the log number of employees computed with offset 1. Values for closed firms are coded with 0 to account for survival bias. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C.14: JEI participation and firm size indicators (ages 9-12)

	Log # employees				Log Sales			
	(1) Age = 9	(2) Age = 10	(3) Age = 11	(4) Age = 12	(5) Age = 9	(6) Age = 10	(7) Age = 11	(8) Age = 12
JEI	0.187***	0.150**	0.124*	0.114^{+}	0.183	−0.054	−0.262	−0.424
	(0.041)	(0.046)	(0.054)	(0.067)	(0.218)	(0.249)	(0.291)	(0.360)
from Paris (0/1)	0.047	0.043	0.076	0.087	−0.052	−0.038	−0.187	−0.001
	(0.050)	(0.056)	(0.067)	(0.085)	(0.260)	(0.296)	(0.351)	(0.435)
Constant	1.033***	0.977***	0.935***	0.935***	1.136***	0.692***	0.351	0.310
	(0.032)	(0.036)	(0.040)	(0.051)	(0.172)	(0.194)	(0.223)	(0.277)
Founding year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	3,659	2,999	2,268	1,526	3,695	2,969	2,226	1,490
# JEI	2,153	1,750	1,288	872	2,171	1,728	1,260	849

Notes: This table examines the relationship between JEI participation and firm size indicators measured at ages 9-12. Dependent variables are the log number of employees and log sales computed with offset 1. They are coded with 0 values for closed firms to account for survival bias. Robust standard errors are shown in parentheses. Stars denote $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure C.7: Evolution of firm size across JEI and non-JEI firms



A. Employment

B. Sales

Notes: This figure presents the evolution of firm size indicators by age across JEI and non-JEI firms. Panel A plots the average number of employees, while Panel B plots average sales (thousands). Both the number of employees and sales are coded with values 0 for closed firms to account for survival bias.

List of Figures

List of Tables

Bibliography

Aghion, P., Dechezleprêtre, A., Hemous, D., Martin, R., & Van Reenen, J. (2016). Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. *Journal of Political Economy*, *124*(1), 1–51.

Ahmadpoor, M., & Jones, B. F. (2017). The dual frontier: Patented inventions and prior scientific advance. *Science*, *357*(6351).

Akcigit, U., & Kerr, W. R. (2018). Growth through heterogeneous innovations. *Journal of Political Economy*, *126*(4), 1374–1443.

Amoroso, S., Aristodemou, L., Criscuolo, C., Dechezleprêtre, A., Dernis, H., Grassano, N., Moussiegt, L., Napolitano, L., Nawa, D., Squicciarini, M., & Tübke, A. (2021). World Corporate Top R&D investors: Paving the way for climate neutrality. A joint JRC and OECD Report. EUR 30884 EN, Luxembourg: Publications Office of the European Union, JRC126788.

Angelucci, S., Hurtado-Albir, F. J., & Volpe, A. (2018). Supporting global initiatives on climate change: The EPO's "Y02-Y04S" tagging scheme. *World Patent Information*, *54*, S85–S92.

Arkolakis, C., Ramondo, N., Rodríguez-Clare, A., & Yeaple, S. (2018). Innovation and production in the global economy. *American Economic Review*, *108*(8), 2128–2173.

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In *The rate and direction of inventive activity: Economic and social factors* (pp. 609–626). Princeton University Press.

Barbieri, N., Marzucchi, A., & Rizzo, U. (2020). Knowledge sources and impacts on subsequent inventions: Do green technologies differ from non-green ones? *Research Policy*, *49*(2), 103901.

Baumol, W. J. (2005). Education for innovation: Entrepreneurial breakthroughs versus corporate incremental improvements. *Innovation Policy and the Economy*, *5*, 33–56.

Berestycki, C., Carattini, S., Dechezleprêtre, A., & Kruse, T. (2022). Measuring and Assessing the Effects of Climate Policy Uncertainty. *OECD Economics Department Working Papers No. 1724. OECD Publishing*.

Bernard, A. B., Fort, T. C., Smeets, V., & Warzynski, F. (2024). Heterogeneous globalization: Offshoring and reorganization. *Working Paper*.

Bilir, L. K. (2014). Patent Laws, Product Life-Cycle Lengths, and Multinational Activity. *American Economic Review*, *104*(7), 1979–2013.

Bilir, L. K., & Morales, E. (2020). Innovation in the Global Firm. *Journal of Political Economy*, *128*(4), 1566–1625.

Bircan, C., Javorcik, B., & Pauly, S. (2021). Creation and Diffusion of Knowledge in the Multinational Firm. *EBRD*.

Branstetter, L. G., Chen, J.-R., Glennon, B., & Zolas, N. (2021). Does Offshoring Production Reduce Innovation: Firm-Level Evidence from Taiwan. *NBER Working Paper No. w29117*.

Branstetter, L. G., Fisman, R., & Foley, C. F. (2006). Do stronger intellectual property rights increase international technology transfer? Empirical evidence from US firm-level panel data. *Quarterly Journal of Economics*, *121*(1), 321–349.

Bronzini, R., & Piselli, P. (2016). The impact of R&D subsidies on firm innovation. *Research Policy*, *45*(2), 442–457.

Clancy, M., & Dyèvre, A. (2023). Big Firms Have Different Incentives. *New Things Under the Sun*.

Cohen, W. M. (2010). Fifty years of empirical studies of innovative activity and performance. *Handbook of the Economics of Innovation*, *1*, 129–213.

Cohen, W. M., & Klepper, S. (1996). Firm size and the nature of innovation within industries: The case of process and product R&D. *The Review of Economics and Statistics*, 232–243.

Conti, A. (2018). Entrepreneurial finance and the effects of restrictions on government R&D subsidies. *Organization Science*, *29*(1), 134–153.

Danzer, A., Feuerbaum, C., & Gaessler, F. (2020). Labor supply and automation innovation. *Max Planck Institute for Innovation & Competition Research Paper 20-09*.

Dechezleprêtre, A., Glachant, M., Haščič, I., Johnstone, N., & Ménière, Y. (2011). Invention and Transfer of Climate Change–Mitigation Technologies: A Global Analysis. *Review of Environmental Economics and Policy*, *5*(1), 109–130.

Dechezleprêtre, A., Martin, R., & Mohnen, M. (2022). Knowledge Spillovers from Clean and Dirty Technologies. *Working Paper*.

Dechezleprêtre, A., Ménière, Y., & Mohnen, M. (2017). International patent families: From application strategies to statistical indicators. *Scientometrics*, *111*, 793–828.

de Rassenfosse, G., Dernis, H., Guellec, D., Picci, L., & De La Potterie, B. V. P. (2013). The worldwide count of priority patents: A new indicator of inventive activity. *Research Policy*, *42*(3), 720–737.

Dernis, H., & Khan, M. (2004). Triadic Patent Families Methodology. *OECD Science, Technology and Industry Working Papers, 2004/02*.

Drees, D., Schild, C.-J., & Walter, F. (2018). Microdatabase Direct Investment 1999-2016. *Data Report 2018-06 - Metadata Version 5. Deutsche Bundesbank Research Data and Service Centre*.

Fan, J. (2021). Talent, Geography, and Offshore R&D. *Working Paper*.

Farre-Mensa, J., Hegde, D., & Ljungqvist, A. (2020). What is a patent worth? Evidence from the US patent "lottery". *The Journal of Finance*, *75*(2), 639–682.

Fort, T. C., Keller, W., Schott, P. K., Yeaple, S., & Zolas, N. (2020). Colocation of Production and Innovation: Evidence from the United States. *Working Paper. Dartmouth College*.

Gautier, E., & Wolff, F.-C. (2020). Les aides à l'innovation ont-elles un effet sur les salaires et l'emploi des jeunes entreprises innovantes ? *BPI France Working Paper*.

Griliches, Z. (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature*, *28*(4), 1661–1707.

Guellec, D., & Pottelsberghe de la Potterie, B. (2005). Measuring the internationalisation of the generation of knowledge. *Handbook of Quantitative Science and Technology Research*, 645–662.

Gumpert, A., Manova, K., Rujan, C., & Schnitzer, M. (2023). Multinational Firms and Global Innovation. *Working Paper*.

Haeussler, C., Harhoff, D., & Mueller, E. (2014). How patenting informs VC investors – The case of biotechnology. *Research Policy*, *43*(8), 1286–1298.

Hall, B. H. (2011). The internationalization of R&D. *SSRN Working Paper No. 2179941*.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics*, 16–38.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). The NBER patent citation data file: Lessons, insights and methodological tools. *NBER Working Paper 8498*.

Hall, B. H., & Lerner, J. (2010). The financing of R&D and innovation. *Handbook of the Economics of Innovation*, *1*, 609–639.

Hallépée, S., & Houlou Garcia, A. (2012). Évaluation du dispositif JEI. *DGCIS Évaluation. Direction générale de la compétitivité, de l'industrie et des services*.

Haltiwanger, J., Jarmin, R. S., Kulick, R., & Miranda, J. (2016). High growth young firms: Contribution to job, output, and productivity growth. In *Measuring entrepreneurial businesses: Current knowledge and challenges* (pp. 11–62). University of Chicago Press.

Haltiwanger, J., Jarmin, R. S., & Miranda, J. (2013). Who creates jobs? Small versus large versus young. *Review of Economics and Statistics*, *95*(2), 347–361.

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.

Harrigan, J., Reshef, A., & Toubal, F. (2023). Techies and Firm Level Productivity. *NBER Working Paper No. 31341*.

Helpman, E., Melitz, M. J., & Yeaple, S. R. (2004). Export Versus FDI with Heterogeneous Firms. *American Economic Review*, *94*(1), 300–316.

Henderson, R. (1993). Underinvestment and incompetence as responses to radical innovation: Evidence from the photolithographic alignment equipment industry. *RAND Journal of Economics*, 248–270.

Hernández, H., Grassano, N., Tübke, A., Amoroso, S., Csefalvay, Z., & Gkotsis, P. (2020). *The 2019 EU Industrial R&D Investment Scoreboard*. EUR 30002 EN. Publications Office of the European Union. JRC118983.

Hottenrott, H., & Richstein, R. (2020). Start-up subsidies: Does the policy instrument matter? *Research Policy*, *49*(1), 103888.

Howell, S. T. (2017). Financing innovation: Evidence from R&D grants. *American Economic Review*, *107*(4), 1136–64.

IEA. (2021). Net Zero by 2050. *International Energy Agency (IEA)*.

IEA. (2023). Energy Technology Perspectives 2023. *International Energy Agency (IEA)*.

INSEE. (2003). Nomenclature des Professions et Catégories Socioprofessionnelles (PCS).

IRENA. (2017). Accelerating the Energy Transition through Innovation. *International Renewable Energy Agency (IRENA) Working Paper*.

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

Javorcik, B. S. (2004). The composition of foreign direct investment and protection of intellectual property rights: Evidence from transition economies. *European Economic Review*, *48*(1), 39–62.

Kerr, W. R., Lerner, J., & Schoar, A. (2014). The consequences of entrepreneurial finance: Evidence from angel financings. *The Review of Financial Studies*, *27*(1), 20–55.

Krieger, J. L., Schnitzer, M., & Watzinger, M. (2023). Standing on the Shoulders of Science. *Harvard Business School Working Paper 21-128*.

Lerner, J. (1999). The Government as Venture Capitalist: The Long-Run Impact of the SBIR Program. *The Journal of Business*, *72*(3), 285–318.

Liu, J. (2023). Multinational Production and Innovation in Tandem. *Working Paper. New York University*.

Marx, M., & Fuegi, A. (2020). Reliance on science: Worldwide front-page patent citations to scientific articles. *Strategic Management Journal*, *41*(9), 1572–1594.

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

Moser, P., Ohmstedt, J., & Rhode, P. W. (2018). Patent citations—an analysis of quality differences and citing practices in hybrid corn. *Management Science*, *64*(4), 1926–1940.

Moutaabbid, A. (2016). Le dispositif JEI a bénéficié à 3,500 entreprises en 2015, essentiellement des secteurs du numérique et des activités scientifiques. *Le 4 Pages de la DGE*, (66).

Nelson, R. R. (1959). The Simple Economics of Basic Scientific Research. *Journal of Political Economy*, *67*(3), 297–306.

OECD. (2008). The Internationalisation of Business R&D: Evidence, Impacts and Implications. *OECD Publishing. Paris*.

OECD. (2011). Fostering Innovation for Green Growth. *OECD Green Growth Studies. OECD Publishing*.

OECD. (2021a). Eligibility of current and capital expenditure for R&D tax relief in selected OECD and non-OECD countries.

OECD. (2021b). No Net Zero Without SMEs: Exploring the Key Issues for Greening SMEs and Green Entrepreneurship. *OECD SME and Entrepreneurship Papers No. 30. OECD Publishing*.

Popp, D., Pless, J., Haščič, I., & Johnstone, N. (2021). Innovation and entrepreneurship in the energy sector. *Working Paper*.

Probst, B., Touboul, S., Glachant, M., & Dechezleprêtre, A. (2021). Global trends in the invention and diffusion of climate change mitigation technologies. *Nature Energy*, *6*(11), 1077–1086.

Quantin, S., Bunel, S., & Lenoir, C. (2021). Evaluation du dispositif Jeune Entreprise Innovante (JEI): Un exemple d'application du modele d'analyse de sensibilité de Rosenbaum. *Working Paper*.

Raffo, J. (2020). MATCHIT: Stata module to match two datasets based on similar text patterns. *Statistical Software Components S457992, Boston College Department of Economics*.

Ramondo, N., Rappoport, V., & Ruhl, K. J. (2016). Intrafirm trade and vertical fragmentation in us multinational corporations. *Journal of International Economics*, *98*, 51–59.

Santoleri, P., Mina, A., Di Minin, A., & Martelli, I. (2022). The causal effects of R&D grants: Evidence from a regression discontinuity. *Review of Economics and Statistics*, 1–42.

Schild, C.-J., Schultz, S., & Wieser, F. (2017). Linking Deutsche Bundesbank Company Data using Machine-Learning-Based Classification. *Technical Report 2017-01. Deutsche Bundesbank Research Data and Service Centre*.

Schmoch, U. (2008). Concept of a technology classification for country comparisons. *Final Report to the World Intellectual Property Organisation, WIPO*.

Schneider, C., & Veugelers, R. (2010). On Young Highly Innovative Companies: Why They Matter and How (Not) to Policy Support Them. *Industrial and Corporate Change*, *19*(4), 969–1007.

Shane, S. (2001). Technological opportunities and new firm creation. *Management Science*, *47*(2), 205–220.

Söderblom, A., Samuelsson, M., Wiklund, J., & Sandberg, R. (2015). Inside the black box of outcome additionality: Effects of early-stage government subsidies on resource accumulation and new venture performance. *Research Policy*, *44*(8), 1501–1512.

Squicciarini, M., Dernis, H., & Criscuolo, C. (2013). Measuring Patent Quality: Indicators of Technological and Economic Value. *OECD Science, Technology and Industry Working Papers, 2013/03. OECD Publishing*.

Tintelnot, F. (2017). Global production with export platforms. *Quarterly Journal of Economics*, *132*(1), 157–209.

Trajtenberg, M., Henderson, R., & Jaffe, A. (1997). University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and New Technology*, *5*(1), 19–50.

Vysoká, L., Dörr, R., Sarris, S., & Gáthy, G. (2021). *Technology Transfer and Commercialisation for the European Green Deal*. EUR 30694. Publications Office of the European Union. JRC124354.

Yeaple, S. R. (2003). The role of skill endowments in the structure of us outward foreign direct investment. *Review of Economics and Statistics*, *85*(3), 726–734.

Yeaple, S. R. (2013). The Multinational Firm. *Annual Review of Economics*, *5*(1), 193–217.

Zhao, B., & Ziedonis, R. (2020). State governments as financiers of technology startups: Evidence from Michigan's R&D loan program. *Research Policy*, *49*(4), 103926.