Spillovers and Selection of Ideas

Firm-Level Evidence from Innovation Networks,
Multinationals in China and Crowdfunding Platforms

Zhaoxin Pu

Dissertation
2020

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

Spillovers and Selection of Ideas

Firm-Level Evidence from Innovation Networks, Multinationals in China and Crowdfunding Platforms

---

Inaugural-Dissertation
zur Erlangung des Grades

Doctor oeconomiae publicae
(Dr. oec. publ.)

an der Volkswirtschaftlichen Fakultät der
Ludwig–Maximilians–Universität München

2020

vorgelegt von

Zhaoxin Pu

Referent:	Prof. Dietmar Harhoff, Ph.D.
Korreferent:	Prof. Dr. Monika Schnitzer
Promotionsabschlussberatung:	22. Juli 2020

Datum der mündlichen Prüfung:	9. Juli 2020
Namen der Berichterstatter:	Prof. Dietmar Harhoff, Ph.D.
	Prof. Dr. Monika Schnitzer
	Prof. Christian Fons-Rosen, Ph.D.

Für meine Eltern.

Acknowledgements

This dissertation would not have been possible without the support and encouragement of many people. First and foremost, I would like to thank my advisor Dietmar Harhoff for his guidance, generous support and continuous encouragement over the past years. I had the privilege to learn from his vast knowledge and experience in all matters related to innovation and entrepreneurship research. I am grateful to my second advisor Monika Schnitzer for her valuable advice and support throughout my doctoral studies and, moreover, for being a role model in many respects. I am indebted to Christian Fons-Rosen, who played an important role as co-author, advisor and mentor during my PhD. His rigor and clarity of thought have been a steady source of motivation and scientific improvement.

I am grateful to Fabian Gäßler for our joint project, his insightful comments and his guidance through the thickets of patent data. Furthermore, I would like to thank Stefano Baruffaldi, Anna Gumpert, Rainer Widmann, as well as participants of the Max Planck Innovation & Entrepreneurship Seminar, the LMU Innovation Workshop and many other conferences for helpful discussions and feedback that greatly improved my work.

The Max Planck Institute for Innovation and Competition and the Munich Graduate School of Economics at LMU have provided an excellent research work environment. I would like to thank the MPI administration and IT teams and the student research assistants for their great support. I gratefully acknowledge financial support from the German Science Foundation and the Elite Network of Bavaria through the CRC TRR 190 “Rationality and Competition” and the doctoral program “Evidence-Based Economics”. Furthermore, I would like to thank the Deutsche Bundesbank’s Research Data and Service Centre for providing me with data for parts of this dissertation and for their hospitality during my numerous research stays in Frankfurt.

My friends and colleagues at the MPI and the MGSE made sure that the years of my PhD were not all work but also lots of fun. I will always cherish our insightful debates and banter during lunch, our coffee breaks and after-work drinks on the bridge and the time we spent together at Schneefernerhaus. I am especially thankful for the moral support, wit and laughter of Dennis Byrski, Anastasia Driva, Ilka Gerhardts, Corinna Hartung, Jörg Hoffmann, Marie Lechler, Felix Pöge, Cristina Rujan, Lisa Simon, Stefan Sorg, Magdalena Streicher, Alexander Suyer and Matthias Wilhelm.

I am grateful to my close friends, who reliably reminded me of the importance of education and perseverance. Without their unconditional love and support, none of this would have been possible. I dedicate this thesis to them.

Zhaoxin Pu
Munich, July 2020

Contents

Preface

Knowledge, technology, and human capital are drivers of economic growth and social welfare (Jones and Romer, 2010). The ability to use, recombine and extend knowledge is essential for firms and organisations to continuously develop and prevail in today's knowledge economy (Mokyr, 2002). Global competition forces companies to innovate and develop their products and services faster. Business investment in knowledge-based capital has increased more rapidly than investment in physical capital in many OECD countries (OECD, 2017). Firms who can successfully combine and exploit knowledge and ideas from different sources will be able to gain competitive advantages.

These sources can originate from within the organisation, other industries, technology areas, or countries. Almost all new knowledge builds on prior ideas – recombinant innovation and cross-pollination of ideas are increasingly important as the burden of knowledge grows (Weitzman, 1998; Jones, 2009). High impact research is characterized by novel features inserted into otherwise conventional combinations of prior work (Uzzi et al., 2013). The growing importance of teams across disciplines and borders is also observed in international co-inventions (Branstetter et al., 2014). Firms' strategy to benefit from spillovers of ideas and their approach to selecting promising ideas are key for an innovation-oriented economy.

Globalisation is not only characterized by the integration of economies and international tradability of goods and services, but also by the increasing internationalization of R&D (Hall, 2011). Corporate R&D make up more than 60% of R&D expenditures in OECD countries (OECD, 2017). Multinational enterprises, in particular, play an important role in the international diffusion of knowledge (e.g., Almeida, 1996; Aitken and Harrison, 1999; Javorcik, 2004; Keller and Yeaple, 2009). As countries have different knowledge profiles, MNEs can tap into different sources of ideas in host countries than at home (Griffith et al., 2006; Branstetter, 2006; Alcácer and Chung, 2007). Knowledge diffusion is partially localized, hence establishing a subsidiary in a foreign country facilitates access to local ideas, especially when knowledge is tacit (Jaffe et al., 1993; Audretsch and Feldman, 1996; Almeida and Kogut, 1999; Branstetter, 2001; Keller, 2002).

This thesis sheds light on three aspects of spillovers and selection of ideas. Each chapter of the thesis uses a quasi-experiment and firm-level data to advance our understanding of firms' innovation and knowledge sourcing behavior. The chapters are self-contained and each can be read on its own. The first chapter examines spillovers of ideas across industries. It studies the innovation network of firms and examines how shocks to invention in a given sector can propagate to the rest of the economy through technological linkages between firms. The second chapter examines spillovers of ideas across borders, and from science to corporate innovation. It assesses how China's science and technology boom has changed the innovation activities of multinational firms with FDI in China. The third chapter studies the selection of ideas. It investigates how governance rules of crowdfunding platforms affect the number and quality of entrepreneurial ventures that get selected for funding.

Chapter 1, which is based on joint work with Christian Fons-Rosen, examines spillovers of ideas across industries. Innovation is often the combination of existing knowledge. Cross-pollination and recombination of ideas across industries and disciplines are important aspects for technological progress (see e.g., Weitzman, 1998). If firms are able to recombine technologically similar bodies of knowledge from different industries, a broad set of industries can create favourable conditions for an economy to benefit from recombinant innovation. What are the effects of shocks to invention in a specific industry, though – do they propagate across the network of innovating firms through technological linkages?

In this chapter, we estimate and quantify the impact of an industry-specific policy shock on the innovation undertaken by firms in technologically proximate but non-targeted industries. Recent innovation literature has documented the value of cross-disciplinary research and the benefits of being exposed to a variety of technology fields (e.g., Jones et al., 2008; Uzzi et al., 2013; Acemoglu et al., 2016). However, trade openness and globalization tend to increase a country's degree of industrial specialization. The direct effects of trade liberalization on innovation have received substantial attention in the recent literature (e.g. Bloom et al., 2016; Autor et al., 2020; Shu and Steinwender, 2019, for a review). This chapter contributes to our understanding of the *indirect* effects of sector-specific trade policies on innovation and knowledge sourcing of firms that are outside the targeted sector.

We construct a firm-level panel of 45,000 European manufacturing firms with information on innovation activity, knowledge sourcing, and technological distances across each pair of firms and sectors. We use the removal of import quotas on Chinese textiles in 2001 as an exogenous competition shock to the European textile sector to identify the induced changes in innovation of non-textile firms through technological linkages to textile firms. We use linked Orbis and PATSTAT data to calculate the pairwise technological proximity between two firms in the network. Using firm-level variation in technological overlap with textile firms, we study how patenting and knowledge sourcing by non-textile firms adjust to the changes experienced by the textile industry after the competition shock.

Our key result is that the shock induced by the removal of Chinese import quotas on textiles propagates through technological linkages across the innovation network, indirectly affecting non-textile firms. While the direct effect of the import quotas removal increases innovation by the *average* European textile firm (Bloom et al., 2016), the indirect effect on non-textile firms is negative once we account for the centrality of each textile firm in the innovation network. This negative effect increases in technological and geographical proximity to textile firms. The effect persists across all firm sizes and upholds also when we analyse the indirect policy effects at the industry level.

We further study if firms redirect their innovation focus and change their knowledge sourcing behavior. Our analysis shows that non-textile firms shift their innovation away from 'textile-intensive' technology areas and cite fewer patents from textile firms. Instead, they start seeking new sources of knowledge that are further away in both the geographical and technological space. We aggregate the data to the regional level to compare the effect sizes of the removal of import quotas on textile firms (direct effect) and on non-textile firms (indirect effects). We find that in the median region, the negative indirect effect is around three times larger than the positive direct effect on patenting.

This chapter contributes to the literature assessing the effects of trade and industry policies on firm innovation in several ways. First, we find that shocks in a specific part of the firm innovation network (textile industry) propagate to other parts of the network (non-textile industries) through technological linkages. We contribute to the literature on technological versus product market spillovers (e.g. Bloom et al., 2013) and show that these shocks diffuse through the technological proximity channel in ways that is not captured by conventional input-output connections in product markets. The evidence suggests that technologically similar knowledge transfers across industries thanks to firms' ability to recombine similar bodies of knowledge.

Second, we contribute to the literature on trade, import competition and innovation by focusing on the indirect effects on firms outside of targeted sectors. There is little evidence regarding the indirect effects of trade on innovation and knowledge sourcing of firms. We advance the literature by not only assessing the changes in the level of patenting, but also the within-firm changes in the direction of innovation as well as knowledge sourcing.

Third, we look at a particular policy episode and use an instrument, which allow us to compare the magnitudes of direct and indirect effects of a sector-specific policy shock on firm innovation. Our finding highlights the importance of these indirect effects. In sum, our results suggest that policy makers should account for direct as well as indirect effects, when designing targeted industrial or trade policies and assessing the potential impact on innovation outcomes in the overall economy.

Chapter 2 studies technology sourcing across national borders through foreign direct investments and the value of science for corporate innovation. In 2018, China was declared world's largest producer of scientific articles for the first time, overtaking the number of scientific publications from the United States (NSF, 2018). China's domestic R&D expenditures increased with an average annual growth rate of 18% between 2000 and 2015. In this chapter, I assess how the rapid development of China's science and technology capabilities has changed the innovation activities of multinational firms with FDI in China. Multinational enterprises (MNEs) are among the most innovative firms in an economy (UNCTAD, 2008; Criscuolo et al., 2010), and play an important role in the international diffusion of knowledge (e.g., Almeida, 1996; Aitken and Harrison, 1999; Javorcik, 2004; Keller and Yeaple, 2009). Foreign direct investments abroad can serve as a channel for productivity and knowledge spillovers from destination country to parent firm (Griffith et al., 2006; Branstetter, 2006). In this chapter, I investigate the effects on MNEs, if one of the largest economies in the world and an increasingly popular FDI destination country launches a major science and technology policy programme that prompted an unprecedented growth of scientific articles.

I use China's Medium- and Long-Term Plan for the Development of Science and Technology ("MLP") in 2006 as a quasi-natural experiment to identify the causal relationship between China's rise in science and German multinational firms' innovation and knowledge sourcing behavior. I assess if German MNEs with FDI in China offshore more of their R&D activities to China, as measured by the number of patents invented in China, and if they rely more on Chinese scientific knowledge in their innovation activities, as measured by the number of references to Chinese scientific articles by patent publications. To study this question, I create a unique dataset that links FDI data (from Deutsche Bundesbank), patent data (from PATSTAT) and bibliometric data (from Web of Science). The dataset captures both foreign direct investment and inventive activities of MNEs and furthermore links corporate innovation to science.

In order to relate the development of science fields targeted by China's MLP policy to MNE's innovation strategies, I construct a proprietary technology-to-science concordance matrix. Based on 9.7 million non-patent references, I calculate the linkage intensities between technology classes and scientific disciplines. The matrix reflects the relevance of a scientific discipline for patents from a certain technology class, and is used to capture the differences between firms in their exposure to the MLP policy and China's scientific progress. This allows me to explicitly model a firm's absorptive capacity and potential learning from science in the empirical analysis (see e.g., Cohen and Levinthal, 1989).

I use a triple differences specification to identify the effect of China's 2006 MLP policy on German multinationals' inventive activities. From official Chinese policy documents, I identify the focal fields targeted by the Chinese government and map these into scientific disciplines in Web of Science. The empirical strategy exploits each firm's different exposure to the MLP targeted scientific fields and accounts for differential trends between MNEs in and outside China. I find that MNEs with subsidiaries in China have a stronger growth in patents invented in China as well as in citations to Chinese scientific articles, compared to MNEs outside China, which is line with existing literature on technology sourcing through FDI (Griffith et al., 2006;

Branstetter, 2006). However, among MNEs with FDI in China, I find no evidence that firms that have a technological profile that closely relates to the MLP targeted scientific fields offshore more of their R&D activities to China, relative to their peers with a lower exposure to the MLP policy. Similarly, they do not seem to increase their reliance on Chinese science. The estimated effects are negative and statistically insignificant, that is, a positive differential effect cannot be ruled out but is at most modest.

This chapter contributes to the literature in several ways. First, it relates to the literature on technology sourcing through FDI and knowledge flows associated with international R&D. While there is an extensive body of research on technology spillovers through trade and FDI (see Keller, 2010, for a survey), most of the papers focus on the spillovers to the host country through inward FDI. Griffith et al. (2006) and Branstetter (2006) are some of the few exceptions and show that UK and Japanese MNEs with affiliates in the USA were able to benefit from knowledge and productivity spillovers through their outward FDI to the US. This chapter is, to the best of my knowledge, the first paper to examine the causal effect of China's S&T boom on multinational firms' innovation strategies in advanced economies. The Chinese context allows me to examine if technology sourcing through FDI can also be observed in a destination country that is still in a development process.

Second, the chapter speaks to the growing literature on the value of science for innovation and productivity. Recent papers in innovation and the economics of science have documented that patents closely related to science are more valuable and more novel (Ahmadpoor and Jones, 2017; Poege et al., 2019; Watzinger and Schnitzer, 2019). This paper contributes by capturing potential benefits of technological proximity to China's science boom for multinational firms. The effects of China's rise in production and exports on other countries' manufacturing and innovation have been examined by a significant body of literature (see Shu and Steinwender, 2019, for a review). But the effects of China's rise in scientific prowess on innovation in other countries have so far received little attention.

Third, a contribution of the chapter is the creation of a unique dataset that allows for detailed studies of multinational firms' foreign investment and innovation activities, as well as the link between corporate innovation and science. While concordance matrices that relate technology classes to industries exist, there is currently no concordance matrix available that relates technology classes to science at large scale. The technology-to-science matrix developed in this chapter could be more broadly applied in other papers that study the relationship between technology and science.

Chapter 3, which is based on joint work with Fabian Gaessler, shifts the focus from the spillovers of ideas to the selection of ideas. Entrepreneurs with innovative ideas in the early stages of their venture often face difficulties in attracting external finance due to the inherent uncertainty over risk and return of their undertaking (Hall and Lerner, 2010). Crowdfunding has recently emerged as a form of entrepreneurial finance that enjoys increasing popularity and that contributes to narrowing the funding gap for innovative young firms and start-ups (Agrawal et al., 2014; Belleflamme and Schwienbacher, 2014; Mollick, 2013). This chapter investigates the governance rules of crowdfunding platforms and examines how platform openness affects the number and type of projects that get selected for funding by the crowd.

The success story of Internet-based crowdfunding platforms is intriguing given that individual non-expert funders make investment decisions in the face of considerable information asymmetry about the quality of projects and the capabilities of project creators. Indeed, crowdfunding is characterized by a high default rate among financed projects (Mollick and Kuppuswamy, 2014). To prevent market failure, crowdfunding platforms may establish screening processes for project applications with the goal to maximize the number of funded projects of high quality and thus reward their funders (Belleflamme et al., 2015). A platform's decision to open up or to restrict access to its marketplace determines the number and type of project creators and funders participating on both sides of the market. As such, platform operators face a trade-off between quantity and quality when deciding on the degree of platform openness.

In this chapter, we study the causal relationship between platform openness, crowdfunding success and performance of funded projects. We empirically investigate the effects of platform openness in the reward-based crowdfunding market, focusing on the two dominating platforms Kickstarter and Indiegogo. We exploit a strategic decision at Kickstarter to switch from access control to de facto openness as a quasi-experiment. On Kickstarter, each project had to undergo a manual review and obtain approval before listing on the platform, which stood in contrast to the open platform strategy of its main competitor Indiegogo. In 2014, however, Kickstarter decided to abandon its access control and adopted an open platform policy. Taking advantage of the different geographical coverage of the two platforms, we use the synthetic control method to construct an appropriate control group for Kickstarter. We combine this with a difference-in-differences approach to disentangle the effects of platform openness on Kickstarter's market thickness and successful funding of ideas. We assess the effects of openness on project quality, specifically the quality of selected projects, using novel text analysis methods and evaluate reward delivery and demand-side feedback of projects that reached the funding stage.

First, we find that removing the manual pre-entry screening had an instant effect on market thickness, both in terms of quantity and variety of projects. However, openness had no significant effect on the projects' novelty. Second, we find that the number of projects that get selected and funded by the crowd increased in absolute, but not in relative terms – the funding success rate of projects dropped by almost a third. Third, among projects that reach their funding goal, we find that those projects without pre-screening are on average of poorer quality. They more frequently fail to deliver their rewards and are subject to more complaints. While the policy change to openness led to more market matches and higher revenues for Kickstarter in the short-run, it comes at the expense of project quality and funder experience. Overall, our results suggest that there are limits to the "wisdom of the crowd" in screening out low quality projects. Effective platform control can facilitate successful matching between entrepreneurial ventures who seek early-stage financing and crowdfunders.

This chapter contributes to the existing literature in the following ways. First, we add to the recent literature on crowdfunding, which has largely focused on the characteristics of projects and funding dynamics so far and paid little attention to the role of platform rules and procedures (Dushnitsky and Fitza, 2018). Our study helps understand how platform design affects the selection and funding of projects. Moreover, our empirical strategy allows us to study the interdependencies between platforms and highlights platform competition effects.

Second, our study relates to the literature on entrepreneurial finance, in particular the importance of due diligence for investors and the survival of entrepreneurial firms (Amit et al., 1998) If platform design determines which projects enter the platform and reach their crowdfunding goal, the "signaling" value of funding success for subsequent financing rounds of young firms becomes platform-specific (Dushnitsky and Zunino, 2019).

Third, this paper provides insights into the strategic value of quality as a competitive advantage for platforms. We show how an open strategy can affect the platform performance in the context of competition. In a thus far growing crowdfunding market environment, market thickness may take precedent over quality of projects. However, as markets mature, quality is likely to become more relevant. Platform governance can play an important role in striking the right balance between quantity and quality, ensuring that ideas worth financing get indeed selected.

In summary, this dissertation offers new insights into firms' innovation and knowledge sourcing behavior. It contributes to the understanding of spillovers and selection of ideas in light of increasingly international and interdisciplinary innovation networks. Evidence from these firm-level analyses may hopefully contribute to designing economic policies that help spurring technological progress and economic growth.

1

The Transmission of Sectoral Shocks Across the Innovation Network

1.1 Introduction

Recent innovation literature has documented the value of inventor teams with different academic backgrounds,[1] the benefits of cross-disciplinary research and the importance of recombining ideas from diverse fields.[2] One example where the recombination of ideas across industries brought ground-breaking innovation is the development of fiber optics technology. Corning, a leader in specialty glass manufacturing, was approached by the British Post Office in the 1960s to explore manufacturing of optical glass fiber that can be used for light-transmission in telecommunications. Despite having no prior experience in the telecommunications sector, Corning soon invented the first low-loss optical fiber, which paved the way for long-distance optical communication.[3]

Governments and science funding bodies worldwide acknowledge the importance of an integrative approach in tackling challenges through innovation, explicitly supporting interdisciplinary collaboration. The US defense agency DARPA is a prominent case, which funded the first materials-focused interdisciplinary laboratories in the late 1960s and recently launched a purpose-built social media platform aimed at facilitating the connection between industry and academic experts from different fields.[4] Diversification, not only of research, but also of production can create favourable conditions for an economy to benefit from cross-pollination of ideas from different sectors.

Trade openness and globalization tend to increase a country's degree of industrial specialization, though. Consistent with seminal work in trade theory,[5] by downsizing unproductive sectors and relocating factors of production towards industries with a comparative advantage, a country ends up with a higher degree of industrial production concentrated in fewer sectors. This specialization can bring about numerous upsides. However, to gauge the aggregate impact of government interventions, one should not only consider the sectors directly targeted by a trade or industrial policy, but take a broader view and also evaluate the indirect effects on the rest of the economy.

This paper aims to evaluate whether a sector-specific shock propagates across the network of innovating firms through technological linkages. When a particular industry changes its innovation strategy in response to a competition shock, does this trigger further innovation adjustments for firms in the rest of the economy that belong to non-targeted industries? How

This chapter is based on joint work with Christian Fons-Rosen.

[1] See e.g., Adams et al. (2005); Jones et al. (2008); Gruber et al. (2013).

[2] See e.g., Acemoglu et al. (2016); Jones (2009); Uzzi et al. (2013); Weitzman (1998).

[3] See Cattani (2006) for more details on Corning and the development of optical glass fiber.

[4] https://www.darpa.mil/news-events/2015-08-14; https://www.darpa.mil/news-events/2019-03-19. The German Fraunhofer-Gesellschaft is another example. Instead of being organized by scientific disciplines, its research portfolio is centred on issue-oriented questions that follow an interdisciplinary approach, such as health, security, mobility or communication. https://www.fraunhofer.de/en/about-fraunhofer/profile-structure.html. Web links last accessed on March 23, 2021.

[5] For example, the Ricardian model, the Heckscher-Ohlin model, or Krugman (1979).

does the magnitude of this indirect effect depend on the technological distance to the industry initially targeted by the shock?

We use the removal of import quotas on textiles in 2001 following China's accession to the WTO as a shock to the European textile & clothing sector to assess these questions. Using firm-pair level variation in technological overlap, we study how patenting decisions by non-textile firms adjust to the changes experienced by the textile industry after the competition shock. We use balance sheet information for our panel of 45,012 European non-textile patenting firms from combining Bureau Van Dijk's ORBIS and Amadeus data sets. Detailed patent information come from matching PATSTAT to our firm sample and include information on filings, citations and technology classes. The data allow us to calculate technological distances at the bilateral firm level.

Our main finding is that the average non-textile firm reduces patenting by 3% of a standard deviation after the textile firms, to which it is highly technologically connected, reduce their patenting by a standard deviation. Non-textile firms react twice as strongly to reductions in patenting by textile firms that are located in the same country, suggesting that geographical proximity is an amplifying factor. Larger non-textile firms experience greater reductions than smaller ones, and results are robust to accounting for industrial input-output relationships. Similar findings are obtained for quality-adjusted patent counts.

We then analyze if firms redirect their innovation focus or change their knowledge sourcing behavior, in response to textile firms' reduced patenting. We find that non-textile firms are less likely to patent in textile-intensive technology classes and tend to diversify across a wider set of fields. Furthermore, they cite fewer textile patents and start searching for new sources of knowledge in more distant geographical and technological spaces. Main results are economically and statistically very similar when we run industry-level estimations, rather than firm-level estimations.

In a final exercise, we aggregate the data to the regional NUTS3 level to gauge the relative effect sizes of the removal of import quotas on textile firms (*direct effects*) and non-textile firms (*indirect effects*). In the median region, textile firms file 4.1% more patents after the shock; at the same time, non-textile firms reduce their patenting by 1.0%. To account for the fact that there are many more non-textile firms than textile firms in an economy, we convert these percentage changes into absolute changes. We find that in the median NUTS3 region, the negative indirect effect is around three times larger than the positive direct effect on patenting.

Our paper speaks to several literature strands. First, in the literature on direct effects of competition on innovation, Bloom et al. (2016) explore China's opening up to globalization and find a positive effect on innovation by European textile firms. By contrast, Autor et al. (2020) find a negative effect on U.S. firms across a large variety of sectors. The finding of Aghion et al. (2005) that competition and innovation have an inverted-U relationship can help reconciling these differing results. Other references on the direct effect of import competition on innovation and productivity include Pavcnik (2002) using Chilean data and Amiti and Konings (2007) on Indonesia. Papers that incorporate the innovation dimension but rather look at the effects of exports include Bustos (2011) using Argentinean data and Aw et al. (2011) using Taiwanese data. Following a more macro approach, Cai et al. (2017) develop a theoretical model of trade, innovation and knowledge diffusion to study the role of country and sector heterogeneity on aggregate R&D and welfare. For an extensive literature review on the general relationship between trade liberalization and innovation, we refer to Shu and Steinwender (2019). Our paper differs from the above in that we look at the *indirect* effects of specific trade policies on innovation and knowledge sourcing of firms in other sectors by highlighting the importance of technological linkages.

Second, our paper relates to the spillovers literature. In particular, our study relates to work by Bloom et al. (2013), who develop an empirical methodology to identify the separate effects of technology and product market spillovers and find that the social rate of return to R&D exceeds the private return. Lychagin et al. (2016) find that geographic and technological spillovers contribute more to productivity than product market spillovers. Zacchia (2019) shows that individual relationships between inventors of different companies drive knowledge spillovers between firms. He constructs a network of companies and causally estimates the spillover effect, which suggest that the marginal social return of R&D performed by a firm amounts to approximately 112% of the marginal private return. We construct a network of innovating firms based on their technological linkages, and study how shocks in one focal point of the network (firms in the textile sectors) propagate to other sectors in the network. We relate to Bloom et al. (2013)'s notion of a firm's position in technology space, and examine how spillover effects via the technological proximity network vary with geographical distance. We advance the literature by looking at a particular policy episode and by using an instrument, which allows us to compare the magnitudes of direct and indirect effects of a sector-specific policy shock on firm innovation. Our empirical strategy which rests on firm-level variation allows us to gauge the average firm-level response, but also to quantify the relative importance of the indirect effects at the industry and regional level.

Third, we speak to the literature on the recombination of ideas. Seminal theoretical work by Cohen and Levinthal (1989) and Bernstein and Nadiri (1989) put at the forefront how firms enjoy knowledge spillovers coming from innovation undertaken in other firms. Weitzman (1998) provides micro-foundations for the knowledge production function, which is modelled as a function of reconfigured old ideas. On the empirical front, Jones et al. (2008) show that across all scientific disciplines, highest-impact papers are produced by teams that increasingly span university boundaries. Based on 18 million scientific publication, Uzzi et al. (2013) show that the highest-impact papers are those that insert novel features into otherwise conventional combinations of prior work. Acemoglu et al. (2016) analyze citation properties of 1.8 million U.S. patents to provide evidence on knowledge sharing across technological elds. When there is more past upstream innovation for a particular technology class to build on, then that technology class innovates more. Griffith et al. (2017) quantify the relative cost for a firm to access new ideas created within its own organization versus elsewhere, based on patent citation time lags, and find that accessing knowledge generated by another firm is much more costly than accessing new ideas across national borders or technology areas. In line with this literature, we follow the notion that firms build on previous knowledge generated by other firms or industries. While Griffith et al. (2017) classify patents by industry in which they are used so to capture economic relatedness, we focus on technology areas. We explicitly want to characterize the firms according to their *technological* profile and linkages, and not their relatedness in the product market.

Fourth, we relate to the literature on industrial policy and the specialization of economic activity. In a theoretical framework, Liu (2019) analyzes industrial policy when sectors are vertically linked through an input-output network. Market imperfections in one sector compound through backward linkages to upstream sectors. He shows that the 'centrality' of sectors in the production network matters, when policy makers decide on which industry to target. We follow a similar idea; instead of production networks, we look at innovation networks and assign more weight to firms that are technologically more 'central' in our empirical analysis. Feldman and Audretsch (1999) examine the effect of the composition of economic activity on innovation. Their descriptive analysis supports the thesis that a diverse set of economic activities is more conducive to innovation than a specialization in a narrow set of areas.

The remainder of this paper is structured as follows: Section 1.2 describes the empirical strategy. Section 1.3 provides details on the dataset, the construction of variables and shows descriptive statistics. Section 1.4 then presents the econometric analysis and a discussion of the results. Section 1.5 concludes.

1.2 Empirical Strategy

Our empirical specification estimates within-firm changes in patenting and knowledge sourcing of non-textile firms (i) as a function of patenting changes of textile firms (j), weighted by the pairwise technological distance to each of these textile firms ($tech_{ij}$). Consider a basic firm-level equation for patents of non-textile firm i in a non-targeted manufacturing industry s, country c, and year t as:

$$lnPat_{isct}^{NTXT} = \beta \sum_{j} tech_{ij} lnPat_{jt}^{TXT} + \gamma_i + \gamma_{st} + \gamma_{ct} + u_{isct} \tag{1.1}$$

The level of observation is at the (non-textile) firm-year level. We take the first difference to sweep out firm fixed effects and estimate:

$$\Delta lnPat_{isct}^{NTXT} = \beta \sum_j tech_{ij} \Delta lnPat_{jt}^{TXT} + \Delta\gamma_{st} + \Delta\gamma_{ct} + \Delta u_{isct} \tag{1.2}$$

where Δ denotes the long five-year difference operator. The dependent variable is the within-firm five-year log change in patents by non-textile firms.[6] The regressor of interest is the change in patenting by textile firms, weighted by their technological proximity to the respective non-textile firm ($tech_{ij}$). The variation comes from the fact that technological proximity differs for each pair of non-textile firm (i) and textile firm (i). We discuss the calculation of the pairwise technological proximity between two firms in detail in Section 1.2.2 below. The empirical specification includes country-year and industry-year fixed effects to absorb country-specific and industry-wide shocks. We use overlapping five-year differences (e.g. 2001-1996, or 2002-1997), in order to maximize the use of our data, and cluster the standard errors at the four-digit industry (SIC4) level.

1.2.1 Instrumental Variable Estimation

A possible concern is an omitted variable bias by which textile and non-textile firms are likely to face an unobserved common technology shock because their patents belong to related technology fields. Technological proximity across a pair of textile and non-textile firms can lead to similar changes in innovation strategy as a reaction to a common technology shock. Another concern may be reverse causality whereby changes in non-textile firms cause a in the behavior of textile firms.

We therefore use an instrumental variable approach to address these concerns. We use the removal of import quotas on textile and clothing from China as an instrument for changes in the innovation output of European textile firms, in the spirit of Bloom, Draca, and Van Reenen (2016) (BDVR). Following China's accession to the WTO in 2001, these import quotas were abolished and caused a competition shock for the textile firms in Europe, affecting their domestic production as well as their patenting activity. The underlying quotas vary at the four-digit industry level, and reflect the toughness of quotas at their level in 2000 prior to their abolishment. Our instrument for the potentially endogenous regressor in equation 1.2 is the technology-weighted level of quotas across the 2,380 textile firms, where again weights differ for each of our non-textile firms. Equation 1.3 presents the first stage and equation 1.2 the second stage of the IV estimation approach.

$$\sum_j tech_{ij} \Delta lnPat_{jt}^{TXT} = \lambda \sum_j tech_{ij} QUOTA_{jt}^{TXT,2000} + \Delta\gamma_{ct} + \Delta\gamma_{st} + \Delta u_{isct} \tag{1.3}$$

Identification hence comes from the instrumented change in patenting of textile firms and

[6] In alternative specifications, we additionally consider dependent variables that reflect changes in knowledge sourcing and in the direction of patenting.

from variation across non-textile firms in their technological exposure to the textile sector. The exclusion restriction here is that shocks to the patenting and knowledge sourcing of *non-textile firms* are uncorrelated with the level of *textile quotas* that were determined in the 1950s-70s. This seems plausible, especially when considering that differences in quotas across four-digit textile industries reflect historic bargaining power of the respective industry in richer western economies when the quotas were introduced. For a more detailed discussion of the quotas instrument, we refer to BDVR.

A priori, there are two possible outcomes for the sign of the coefficient on the instrumental variable in equation (2.3). If $\lambda > 0$, it means that the larger the quota reduction, the stronger the import competition from China and the stronger is the *increase* in domestic textile patenting. While $\lambda < 0$ would imply that the larger the quota reduction, the stronger the *drop* in domestic textile patenting.[7]

1.2.2 Technological Proximity between Firms

A central element of our empirical specification is the pairwise technological proximity ($tech_{ij}$) between two firms. We calculate the technological proximity between any non-textile firm and any textile firm in the sample based on the overlap in their patent portfolio. For each firm, we determine its patent portfolio as of 2001 and construct a vector of patent shares across technology classes, which reflects the firm's technological profile. We then calculate the pairwise un-centred correlation between any two firms' patent portfolio vectors.[8]

$$tech_{ij} = \frac{\sum_c PAT_{ic} * PAT_{jc}}{\sqrt{\sum_c PAT_{ic}^2} * \sqrt{\sum_c PAT_{jc}^2}} \in (0,1) \tag{1.4}$$

In order to determine the technological profile of a firm's patent portfolio, we need to assign each patent to a unique patent technology class. In our main specifications, we use a technology classification that builds on 34 technology areas (TF34), aggregated from the IPC codes following the proposal by Schmoch (2008), which unambiguously assigns each patent to one of these technology areas. In an alternative specification, we use the Cooperative Patent Classification (three-digit CPC codes) to classify each patent into more granular technology classes. Under the three-digit CPC scheme, there are 126 distinct technology classes.[9]

We use these technological linkages to define the innovation network between firms, that is,

[7]This first stage differs in multiple dimensions from the baseline estimation in BDVR. Appendix A.1 describes these differences in great detail.

[8]For similar approaches, see Jaffe (1986), Bloom, Schankerman, and Van Reenen (2013), and Lychagin, Pinkse, Slade, and Reenen (2016).

[9]We consider only the main technology area or main three-digit CPC code of a patent, even if patents may be assigned to multiple technology classes. CPC codes distinguish between the position and hence importance of the different codes associated with one patent (i.e. cpc-position = F first or L later). We use this information to determine the unique technology class for each patent.

to establish extensive and intensive connections between firm-pairs – whether two firms are connected and how strongly connected they are. We hold the network constant and study how shocks in one focal area of the network (here firms in the textile sector) propagate to firms in other (non-textile) sectors in the network through their technological linkages.[10] The notion of the network mechanism is as follows, if inventions in textiles increase, then spillovers from textile to non-textile firms will be higher and inventions will increase more for those non-textile firms that are technologically closest. Ideas can propagate across industries thanks to firms' ability to recombine technologically similar bodies of knowledge.

1.2.3 Testing for Alternative Mechanisms

Input-Output Relationships. A concern is that textile and non-textile firms that are technologically close may also be closely linked through vertical input-output industrial relationships, which may confound our regression results. One could for example think of financial dependence of a non-textile firm on the textile industries if its main customers are in the textile industries. The strength of a firm's vertical relationship with the textile industries depends on two aspects: first, the share of production outputs it supplies to the textile industries; second, the share of inputs it receives from the textile industries. We therefore determine each non-textile firm's input and output exposure to the textile industries and test whether these factors drive any of our baseline results.

Industry Level Estimations. In addition to our within-firm analysis, we use industry-level regressions for non-textile SIC4 industries to better understand the indirect effects of the textile quota removal at a more aggregate level. In our baseline firm-level estimation, all non-textile firms in the sample are given the same weight, as we use a non-weighted estimation specification. This could lead to the following scenario: suppose a four-digit industry consists of one very large firm that dominates the patenting activity of the industry and many smaller patenting firms. If all the small firms reduce patenting but the one very large firm increases patenting, at the firm level, our regression analysis would suggest that on average patenting decreases. However, at the industry level, aggregate patents may actually increase.

$$\Delta lnPat_{kt}^{NTXT} = \beta_1 \sum_{l,l \neq k} tech_{kl} \Delta lnPat_{lt}^{TXT} + \beta_2 \sum_{l,l \neq k} IO_{kl} \Delta lnY_{lt}^{TXT} + \Delta \gamma_{SIC2D} + \Delta u_{kt} \tag{1.5}$$

Also here, we account for possible vertical relationships through interdependencies in sales between a non-textile industry k and a textile industry l. As an additional regressor we include changes in textile sales at the four-digit level, weighted by the industry-pair specific input link, respectively output link, from a SIC4 input-output matrix (IO_{kl}). The term γ_{SIC2D} captures industry fixed-effects at the two-digit SIC code level. In absence of a second instrument for

[10] A fixed network seems a reasonable assumption here, as firms change their technological focus slowly over time.

changes in textile sales, and to avoid two endogenous regressors, we estimate equation (1.5) using OLS.

Regional Effects. For example, employees of textile firms affected by the industrial policy could relocate to non-textile firms with a similar technological focus in the same local labor market. In order to test for potential alternative channels at the regional level, we assess if geographical distance plays a role in explaining our results. More precisely, we estimate models where we only consider textile firms located in the same country or within a 50km radius of the non-textile firm. Contrasting these estimates to model estimates including textile firms in foreign countries or further away than 50km allows us to gauge if some labor reallocation mechanism is likely driving our results.

1.3 Data & Descriptive Analysis

For our firm-level analysis, we link Bureau van Dijk's (BvD) ORBIS and Amadeus databases to the PATSTAT database. The ORBIS database is the largest cross-country firm-level database available and includes both public and private firms from all industries. Among others, it includes firm-level data on financial accounts, industry codes, and address data. We use the 2016 Fall version of BvD ORBIS, which includes all historical ORBIS vintages from 2005-2016.[11] We complement ORBIS with the 2006 vintage of BvD's Amadeus database, which includes firm financial data from 1995-2006, in order to improve data coverage for the late 1990s. Amadeus is a similar database of the same data provider BvD, covering firms in Europe rather than globally.

As Kalemli-Ozcan et al. (2015) discuss, it is advisable to combine different BvD vintages to obtain a consistent coverage of firms over time. We link ORBIS to Amadeus at the firm-year level via BvD's unique firm identifier (BvD-ID), while accounting for duplicate accounts, different currencies and accounting standards as well as possible BvD-ID changes over time. For the harmonization and cleaning of the ORBIS and Amadeus data, we broadly follow Kalemli-Ozcan et al. (2015). In the following, we describe the sample of manufacturing firms and the construction of variables used for our econometric analysis.

Sample of European Manufacturing Firms with Patenting Activity. Consistent with the previous literature (e.g. Bloom et al., 2016; Autor et al., 2020) our analysis focuses on firms in the manufacturing sector. We use the four-digit SIC industry information in ORBIS to identify all firms that belong to the manufacturing sector in any of the 13 European countries of Austria, Denmark, France, Germany, Italy, the Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland, and the UK. From these approximately 2 million firms, we identify those that

[11] For representativeness of ORBIS data, see Kalemli-Ozcan et al. (2015).

were active as of 2001, the treatment year when the textile import quotas were abolished. We use the incorporation date information in ORBIS where available, and otherwise deduce from non-missing entries for revenue or employees whether a firm has been active as of 2001. Keeping only these, we are left with about 1.6 million manufacturing firms. This includes patenting and non-patenting firms.

For our empirical analysis, we are interested in those firms that innovate and undertake patenting. We use ORBIS' embedded BvD-to-PATSTAT link to merge the firm data to patent data.[12] About five percent of the above 1.6 million manufacturing firms have a link to the PATSTAT database. In order to calculate a firm's technological proximity to other firms based on its patent filings in the pre-period, we further need to impose the condition that firms in our sample patented at least once before 2001. The above steps result in a final sample of 45,012 non-textile and 2,380 textile firms. Figure 1.1 shows the concentration of textile firms for each region in our sample and the distribution of non-textile firms by country. The largest concentration of textile firms is in Southern Europe (Spain, Portugal and Italy) and in Poland. Nonetheless, the northern part of the United Kingdom and some areas in Norway, France and Germany also have a higher ratio of textile firms.

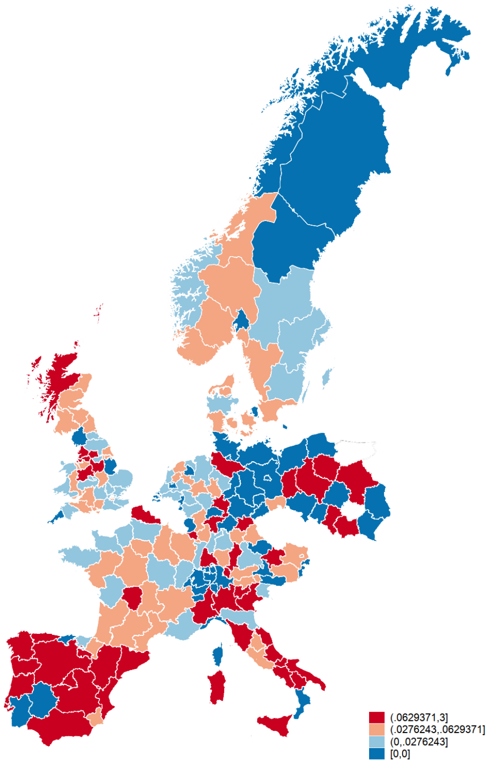
Patent Filings. Our main dependent variable is the change in the number of patents filed by a firm. We consider patents at the DOCDB patent family level. In our paper, we refer to patent families interchangeably as patents. Our sample of 45,012 non-textile firms filed around 615,000 patent families during the years 1996-2005, while the textile firms filed approximately 10,000 patent families during the same period.

As we aggregate patent applications to the patent family level, we need to take a few decisions as to how we unify patent attributes at the patent family level. The year is determined by the filing year of the patent member that was filed first within the family. For the technology class, we consider the modal technology area under the TF34 scheme. In the event of ties, we use the numerically lowest technology area. When using the CPC scheme, we prioritize the so-called "F" codes. CPC codes distinguish between the position and hence the importance of different technology classes associated with one patent (i.e. cpc-position = "F" first or "L" later). Where this information is available, we prioritise the "F" codes, and consider the modal code in case there are multiple "F" codes.

We use technology class information also for assessing changes in the direction and diversity of a firm's patenting activities. Focusing on the patents filed by textile firms in our sample, we identify those technology classes where textile firms are more actively patenting. For each non-textile firm, we then calculate the fraction of patents it files in these relatively 'textile-

[12] For the matching of ORBIS and PATSTAT, Bureau van Dijk uses string similarity matching between a company name from ORBIS and the name of the patent applicant from PATSTAT, mapping BvD-IDs to each PATSTAT person ID. Additional information like address information is used to enhance the matching precision.

Figure 1.1: Number of Textile vs. Non-Textile Firms



Country	Non-Textile Firms	(in %)
AT	839	1.86
CH	3,114	6.92
DE	15,366	34.14
DK	661	1.47
ES	1,655	3.68
FR	5,525	12.27
GB	8,672	19.27
IT	5,164	11.47
NL	1,849	4.11
NO	571	1.27
PL	92	0.20
PT	67	0.15
SE	1,437	3.19
Total	45,012	100.00

Notes: The map shows the relative number of textile firms to non-textile firms in each NUTS2 region in the 13 European countries of our sample. Red-shaded regions have a relatively high share of textile firms, blue-shaded regions have a relatively low share of textile firms.

intensive' technology classes and assess whether there is any redirection of patenting towards or away from these technology classes.[13] Furthermore, we calculate a Herfindahl-Hirschman-Index (HHI) that reflects the level of specialization of a firm's patent portfolio in terms of filings by technology class. We use this index to assess if firms diversify more or less as part of their broader adjustments in patenting and knowledge sourcing decisions.

The technological proximity network is calculated based on firms' pre-period technological profiles. That is, we consider patents filed prior to 2001 for the calculation of the pairwise technological proximity between two firms.

Import Quotas Data for the IV Approach. For the import quotas on textiles that were abandoned following China's entry into the WTO, we use data from BDVR. The quotas variable reflects the toughness of quotas at their initial level in 2000 prior to China's WTO

[13]More precisely, we weight the number of patents a firm files in a given technology class by multiplying it with an indicator that reflects the technology class 'textile-intensity'. The indicator value ranges between 0 and 1 and is measured as the share of patents in this technology class filed in the pre-period prior to 2001 by textile firms (vs. non-textile firms).

entry and varies at the four-digit industry level. It is calculated as the proportion of (import value-weighted) HS 6-digit product categories that were covered by a quota within that four-digit industry. The removal of these quotas had a direct effect on textile firms. Based on the four-digit industry codes of each firm, we can calculate a firm-specific measure of the intensity of the quota reduction.[14]

Geocoding & Geographical Distance between Two Firms. We have detailed address data from ORBIS, which allows us to geocode the location of the firms in our sample. We use the HERE Geocoder API and, where available, a firm's street name, ZIP code, city name and NUTS codes to obtain the corresponding longitude and latitude geo-coordinates.[15] We then compute the geographical distance between any non-textile firm and any textile firm in our sample, using the STATA package "geodist".

Patent Citation Data. As a measure of patent value, we use the number of forward citations received from EPO patents within five years after the first filing date (Harhoff et al., 1999). We also use patent citation data to study the knowledge sourcing behavior of firms. Based on patent-to-patent citations we form dyads consisting of the citing firm and the cited firm. In terms of citing patents, we only consider citations from EPO patents that were filed with the EPO directly or under the PCT, so as to have consistent citation behavior.[16] In terms of cited patents, we consider all patents, irrespective of the office they were filed at. Due to the restriction to EPO citing patents, our sample is somewhat smaller in those specifications where we use this second variant of measuring technological proximity based on citation data.

As set out above, we calculate the pairwise technological distance between any two firms in our sample. Similarly, we also calculate the pairwise geographical distance between any two firms in our sample. We can then analyze the technological versus geographical distance to the cited firms and knowledge sources, and evaluate if firms have to 'travel further' in the knowledge or in the geographical space when a given industry becomes less pivotal as a source of knowledge.

Vertical Industrial Input-Output Relationship. We want to account for potential confounding factors associated with vertical linkages between textile and non-textile firms. In order to capture a non-textile firm's vertical input and output exposure to the textile industries, we use a SIC4 industry-level input-output matrix, as conventionally used in the literature,[17] and

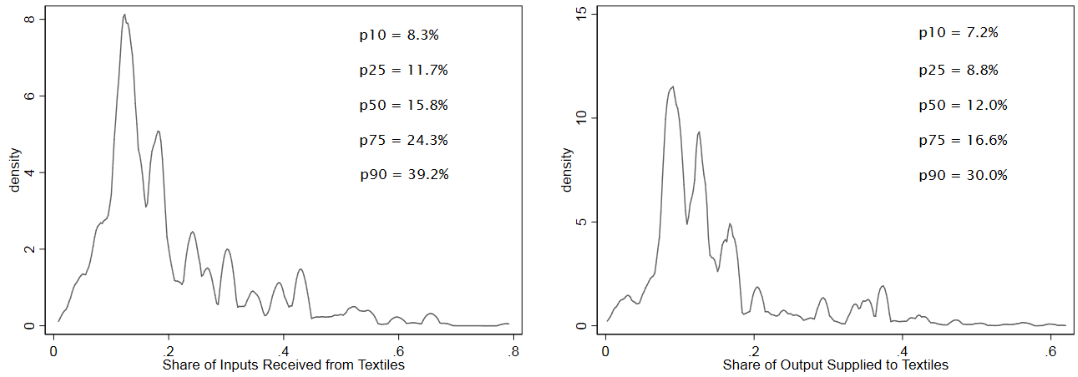
[14]A firm can have multiple primary and secondary SIC 4-digit industry codes. We follow the approach of BDVR, applying a two-third weight to primary codes, a one-third weight to secondary codes, and equal weighting within these groups.

[15]For ca. 97% of the firms in our sample, we know the address at the street level; for 1% the info is at the ZIP or city level; for 2% we have no address info.

[16]See e.g., Alcácer et al. (2009) and Bacchiocchi and Montobbio (2010) for a discussion of different patent citation behavior between USPTO and EPO.

[17]E.g. Javorcik (2004), Liu (2019).

Figure 1.2: Vertical Production Links in the Input/Output Matrix



Notes: This figure plots the distribution of the vertical input & output exposure measures across the non-textile firms. The left panel shows the histogram for the share of inputs received from textile industries, the right panel shows the histogram for the share of output supplied to textile industries.

combine it with a firm's industry profile. If a firm has multiple industry SIC codes, we weight them as before: two-third weight to primary codes, a one-third weight to secondary codes, and equal weighting within these groups. For each firm, we calculate the share in output that it supplies to the textile industries, as well as the share in inputs that it sources from the textile industries. Figure 1.2 displays histograms for these two measures of the vertical input-output relationship to textile industries. Their correlation with our technological proximity measure is positive and small (0.064 and 0.080, respectively).

Table 1.1 provides descriptive statistics for our sample of non-textile firms. The median firm employs 52.5 workers, has annual revenues of 7.4 million Euros and total assets of 4.6 million Euros. It files 0.2 patents per year and the patent stock amounted to 2.0 patents as of 2001. The sample is skewed in terms of both firm size and patenting activity.

Table 1.1: Summary Statistics

	mean	p5	p25	p50	p75	p95	count
No. of Employees	538.31	2.00	16.75	52.50	169.40	1,138.00	29,510
Revenue (th.)	120,285.28	250.00	2,150.00	7,372.25	26,896.50	240,868.70	29,338
Total Assets (th.)	13,8861.30	98.85	1,177.23	4,597.00	19,871.80	240,714.80	24,389
Patent Stock	29.03	0.20	1.00	2.00	7.20	53.00	45,012
Patent Filings p.a.	1.35	0.00	0.00	0.20	0.40	3.00	45,012
No. of Primary SIC Codes	1.50	1.00	1.00	1.00	2.00	4.00	45,012
No. of Secondary SIC Codes	1.07	0.00	0.00	0.00	1.00	5.00	45,012
Observations	45,012						

Notes: This table presents summary statistics for our sample of non-textile firms. The first three financial variables are provided by Bureau Van Dijk and approximately a third of the sample has missing values. Patent stock and patent filings come from PATSTAT, the industry SIC codes come from Bureau van Dijk; these variables are available for all firms.

1.4 Results

1.4.1 Baseline Results

OLS and Reduced Form Estimations. Table 1.2 displays the baseline OLS estimation results. The dependent variable is the within-firm 5-year log change in patents by non-textile firms. The regressor of interest is the tech-weighted average patenting change by each of the 2,380 textile firms. As the technological proximity differs for each pair of firms, the weights vary for each non-textile firm. All columns control for country-specific macro shocks by including a full set of country dummies interacted with a full set of time dummies. In Column (1) we find an elasticity of 1%, meaning that a decrease in (tech-weighted) textile patenting by one standard deviation is associated with a 1% of a standard deviation decrease in non-textile patenting. Results remains stable in Columns (2) and (3) after adding a full set of 2-digit and 4-digit industry dummies interacted with a full set of time dummies, respectively.

Table 1.2: OLS Results

DV: ΔlnPat of NTXT firms	(1)	(2)	(3)
ΔlnPat of TXT firms	0.0102***	0.0112***	0.0098***
	(0.004)	(0.004)	(0.003)
Country-Year FE	Yes	Yes	Yes
Industry(SIC-2D)-Year FE	No	Yes	No
Industry(SIC-4D)-Year FE	No	No	Yes
No. of clusters	471	471	471
Observations	225,060	225,060	225,060
Unique Firms	45,012	45,012	45012

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the SIC4 industry level. This table presents the baseline OLS results for the full panel of non-textile firms. The dependent variable is the change in the log of patents by non-textile firms. The regressor of interest is the change in the log of patents by textile firms, weighted by the technological proximity between a given non-textile firm and each textile firm. To clarify, given that we have 2,380 textile firms in our sample, each regressor is a weighted average of 2,380 changes in the log of patenting, and the closer the technological proximity to the non-textile firm, the greater the weight assigned. Equation (1) is the baseline specification that controls for country-year fixed effects. Equations (2) and (3) additionally include a full set of year dummies interacted with a full set of industry 2-digit and 4-digit dummies, respectively.

Table 1.3 presents the reduced form in which the toughness of quota removals is directly regressed on the same dependent variable as in Table 2, namely, the within-firm 5-year log change in patents by non-textile firms. Results remain stable across all three columns: a standard deviation increase in quota toughness is associated with a 2% of a standard deviation decrease in patenting by a given non-textile firm.

Table 1.3: Reduced Form Results

DV: △lnPat of NTXT firms	(1)	(2)	(3)
Toughness of quotas in 2000	−0.0196***	−0.0191***	−0.0176***
	(0.003)	(0.003)	(0.003)
Country-Year FE	Yes	Yes	Yes
Industry(SIC-2D)-Year FE	No	Yes	No
Industry(SIC-4D)-Year FE	No	No	Yes
No. of clusters	471	471	471
Observations	225,060	225,060	225,060
Unique Firms	45,012	45,012	45,012

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the SIC4 industry level. This table presents the reduced form results for the full panel of non-textile firms. The dependent variable is the change in the log of patents by non-textile firms. The regressor of interest is the quota toughness prior to Chinas accession to the WTO for each textile firm, weighted by the technological proximity between a given non-textile firm and each textile firm. To clarify, given that we have 2,380 textile firms in our sample, each regressor is a weighted average of 2,380 quota changes, and the closer the technological proximity to the non-textile firm, the greater the weight assigned. Equation (1) is the baseline specification that controls for country-year fixed effects. Equations (2) and (3) additionally include a full set of year dummies interacted with a full set of industry 2-digit and 4-digit dummies, respectively.

Instrumental Variable Estimations. Table 1.4 presents IV results using Chinas WTO accession. The uneven numbered columns report the first stage results and the even numbered columns report the second stage results. In this panel of non-textile firms, the endogenous variable is the tech-weighted average change in patents by textile firms, and the instrument is the tech-weighted toughness of quotas faced by textile firms in 2000. The observed negative coefficient in Column (1) implies that the removing of tougher quotas is related to stronger reductions in textile patenting.[18] Column (2) presents the second stage results that show a strong and significant effect of (instrumented) reductions in *textile* patenting on reductions in *non-textile* patenting with a magnitude of 3.6%. Similar to previous tables, the next columns incorporate industry-year dummies leading to only minor changes.

The IV results in Table 1.4 indicate that the OLS coefficient appears downward biased. The test statistics for under-identification (Kleibergen-Paap rk LM statistic) and weak identification (Kleibergen-Paap rk Wald statistic) show that the first stage is very strong in all cases. Given the robust IV results, we prefer the third specification with country-year and SIC4 industry-year fixed effects and will consider the model in Column (6) as our baseline specification going forward. Appendix Table A-1 shows that our baseline results also hold when we condition on a sample of firms for which we have financial data in ORBIS.

[18]To reconcile this negative coefficient with the results in Bloom et al. (2016), in which the removal of tougher quotas leads to more patenting by textile firms, we refer to Appendix A.1 and Table A.1.

Table 1.4: IV Results - First and Second Stage

DV: ΔlnPat of NTXT firms						
	(1)	(2)	(3)	(4)	(5)	(6)
Method	1st stage	2nd stage	1st stage	2nd stage	1st stage	2nd stage
Toughness of quotas in 2000	-0.5341^{***}		-0.5325^{***}		-0.5340^{***}	
	(0.004)		(0.004)		(0.004)	
ΔlnPat of TXT firms		0.0367^{***}		0.0360^{***}		0.0329^{***}
		(0.006)		(0.006)		(0.006)
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry(SIC-2D)-Year FE	No	No	Yes	Yes	No	No
Industry(SIC-4D)-Year FE	No	No	No	No	Yes	Yes
Underidentification test		59.8		54.1		51.9
Weak identification test		14,991.6		14,484.1		15,609.7
No. of clusters	471	471	471	471	471	471
Observations	225,060	225,060	225,060	225,060	225,060	225,060
Unique Firms	45,012	45,012	45,012	45,012	45,012	45,012

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the SIC4 industry level. This table presents the instrumental variable estimation results for the full panel of non-textile firms. The dependent variable is the same as in Tables 2 and 3. The instrument is the weighted change in the quota toughness prior to Chinas accession to the WTO for each textile firm, as described in Table 3. The endogenous regression is the weighted change in the log of patents by textile firms, as described in Table 2. Columns (1), (3), and (5) present first stage results, while columns (2), (4), and (6) present second stage results. Equations (1) and (2) are the baseline specification that control for country-year fixed effects. Equations (3) and (4) additionally include a full set of year dummies interacted with a full set of industry 2-digit dummies. Equations (5) and (6) instead include a full set of year dummies interacted with a full set of industry 4-digit dummies. The table reports test statistics for underidentification (Kleiberger-Paap rk LM statistic) and weak identification (Kleibergen-Paap rk Wald statistic).

1.4.2 Heterogeneity & Robustness

Geographical Heterogeneity. In the following, we describe how the intensity of the observed reduction in non-textile patenting depends on geographical distance. For illustration purposes, Column (1) of Table 1.5 repeats the result of Column (6) in Table 1.4 - our baseline result - and is based on a tech-weighted average of *all* textile firms across our sample of European firms. Columns (2) and (3) split these textile firms into two groups based on whether they are located in the same country as the non-textile firm or not. For the construction of the tech-weighted average textile patent changes in the regressor, Column (2) includes textile firms in other European countries and Column (3) only includes textile firms within the same country as the non-textile firm. The estimated coefficient is substantially larger in the latter case (0.0575 vs. 0.0324), suggesting that geographical proximity amplifies the impact that textile firms have on non-textile firms, conditional on a given technological distance. One could expect to possibly find heterogeneous effects within the same country, for example of local labor markets allowing for the reallocation of affected employees across firms with a similar technological focus. Columns (4) and (5) split the textile firms into being closer or further away than 50 kilometres, respectively. Coefficients are similar and do not seem to be consistent with any major labor reallocation mechanism. The number of observations in

Table 1.5: Geographical Heterogeneity

DV: ΔlnPat of NTXT firms	(1)	(2)	(3)	(4)	(5)
ΔlnPat of TXT firms: all	0.0329***				
	(0.006)				
ΔlnPat of TXT firms: other countries		0.0324***			
		(0.006)			
ΔlnPat of TXT firms: same country			0.0575***		
			(0.011)		
ΔlnPat of TXT firms: same country, <50km				0.0570***	
				(0.019)	
ΔlnPat of TXT firms: same country, >50km					0.0516***
					(0.011)
Country-year FE	Yes	Yes	Yes	Yes	Yes
Industry(SIC-4D)-year FE	Yes	Yes	Yes	Yes	Yes
Underidentification test	51.9	53.3	47.5	26.6	46.8
Weak identification test	15,609.7	6,757.8	1,177.6	203.3	1,162.5
No. of clusters	471	471	471	452	462
Observations	225,060	225,060	225,060	204,400	216,135
Unique Firms	45,012	45,012	45,012	40,880	43,227

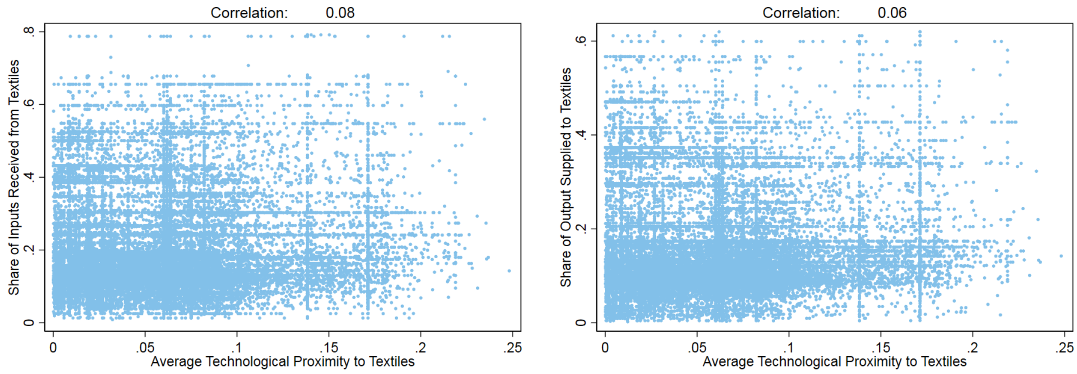
Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the SIC4 industry level. This table introduces geographical variation into the analysis. Presented are the second stages of IV estimations where each regressor differs in its geographical scope. The dependent variable is the same as in Tables 2 and 3. The endogenous regressor is the weighted change in the log of patents by textile firms, as described in Table 2. In Column (1) the regressor encompasses all textile firms at its weighted average; in Columns (2) and (3) it is limited to the sample of textile firms outside, or respectively inside, the same country as the non-textile firm. Considering only textile firms in the same country, Column (4) restricts the sample to those within a 50km radius of the non-textile firm; finally, Column (5) restricts the sample to the ones outside of a 50km radius of the non-textile firm. All specifications include country-year fixed effects as well as a full set of year dummies interacted with a full set of industry 4-digit dummies.

these last two columns is smaller, as there are some non-textile firms for which we know the country but not the exact address to calculate the geographical distance.

Accounting for Input-Output Relationships. A particular concern might be that our results are driven by industrial input-output relationships that correlate with our technological proximity measures. Figure 1.3 plots for each non-textile firm the average technological proximity to the textile industry against both the share of inputs received from textiles (left panel) and the share of production outputs supplied to textiles (right panel). The correlation between these vertical connectedness measures and our technological proximity measure is positive, albeit small at 0.08 and 0.06, respectively, which mitigates the possibility that our results are driven by this alternative mechanism.

Nonetheless, Table 1.6 explicitly accounts for these input-output relationships in our econometric framework. As before, Column (1) is the baseline. Columns (2) and (3) split non-textile firms by the median output exposure to the textile sector, while Columns (4) and (5) redo the same exercise for the median input exposure. We reassuringly do not observe any substantial variation in the magnitude of the estimated coefficient, meaning that the

Figure 1.3: Technological Connectedness vs. Vertical Connectedness to Textile Firms



Notes: This figure shows the correlation between technological proximity and vertical input & output exposure measures for the sample of non-textile firms: share of inputs received from textile industries (left) and share of outputs supplied to textile industries (right).

industrial network channel seems close to orthogonal to our story. Columns (6) and (7) compare non-textile firms with weak input and output links (below 25th percentile on both dimensions) to the textile industry to the ones with strong input and output links (above 75th percentile on both dimensions). While both results remain statistically and economically significant, if anything, the firms with the weak links to textile firms have a larger estimated coefficient.

Table 1.6: Accounting for Input-Output Relationships

DV: ΔlnPat of NTXT firms	Base	Output		Input		Input & Output	
	(1)	(2) <p50	(3) >p50	(4) <p50	(5) >p50	(6) <p25	(7) >p75
ΔlnPat of TXT firms	0.0329***	0.0297***	0.0331***	0.0336***	0.0313***	0.0493**	0.0207*
	(0.006)	(0.010)	(0.007)	(0.009)	(0.008)	(0.022)	(0.013)
Country-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry(SIC-4D)-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Underidentification test	51.9	48.0	25.1	35.3	24.1	25.8	25.4
Weak identification test	15,609.7	9,071.9	10,542.9	9,882.1	11,028.0	6,551.1	4,663.3
No. of clusters	471	303	288	296	298	167	168
Observations	225,060	110,220	107,260	109,760	108,255	25,515	41,120
Unique Firms	45,012	22,044	21,452	21,952	21,651	5,103	8,224

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the SIC4 industry level. In this table we account for the impact of industrial input-output relationships in the production network. The dependent variable and regressor of interest are as in Table 1.2. Column (1) repeats the baseline IV result of Table 1.4 Column (6). Columns (2) and (3) split the non-textile firms by the median output exposure to the textile sector, while Columns (4) and (5) split them by the median input exposure to textiles. Columns (6) and (7) estimate the effect for the non-textile firms with the weakest vs. the strongest input and output exposure to textiles (below 25th percentile vs. above 75th percentile on both dimensions).

Firm Size Heterogeneity, Patent Quality & Lag Specifications. We also check for firm size heterogeneity and find that the effect increases with the size of the firm. Appendix Table A-2 reports the IV results by the quartile of size; the effect for the largest firms (4th quartile) is approximately four times larger than for the smallest firms (1st quartile). The distribution of patent quality is known to be skewed, with only a subset of patents having significant market value. We therefore re-estimate our baseline model, restricting the patent count to the subset filed at the European Patent Office (EPO) and using a cite-weighted patent count. Appendix Table A-3 shows that the main result above still holds, meaning that the effect is not driven by changes in patent quality. We further test alternative models that use lagged regressors. The estimated coefficients in Appendix Table A-4 show robust results, with larger effects for the two-year lag of the reduction in textile patents.

Patenting Direction and Citation Behavior. Up to now, we have seen that non-textile firms reduce their patenting after an exogenous reduction in patenting by textile firms, and it does not seem to be driven by input-output relationships. The magnitude of the effect is increasing in the technological proximity to the patenting undertaken by textile firms. One mechanism consistent with this result is that non-textile firms have lost a source of knowledge for their innovation activities. We explore this possible channel in Table 1.7. We identify those technology classes that have the largest share of patents originating from textile firms relatively. Column (1) is our baseline result and in Column (2) we find that the reduction in patenting by textile firms leads non-textile firms to move away from these textile-intensive technology classes and refocus towards technology areas where textile firms are less prevalent. At the same time, in Column (3) the HHI concentration index decreases, meaning that non-textile firms diversify more across a larger set of technology classes.

Table 1.8 also addresses the refocusing of innovation efforts by non-textile firms from a different angle. Consistent with the storyline of previous findings, the result of Column (1) suggests that non-textile firms are less likely to cite patents of textile firms after the China WTO accession. The remaining columns then analyze whether non-textile firms have to look in more distant locations (both technologically and geographically) to partially substitute for the knowledge lost from textile firms. The negative estimated coefficients in Columns (2) and (3) imply that non-textile firms start citing more technologically distant patents, with results mainly driven by large firms. Columns (4) and (5) report that, for the case of large firms, non-textile firms also start citing patents from more geographically distant firms. Overall, the results of this table are consistent with non-textile firms trying to mitigate the loss of textile firms as a source of knowledge by exploring new sources that are in more distant technological and geographical areas.

Table 1.7: Patenting Direction: 'Textile-Intensity' and Diversity of Tech Classes

DV: Δ in	(1) PAT filings (baseline)	(2) Share of PAT in 'txt-int.' classes	(3) HHI (1/diversity)
ΔlnPat of TXT firms	0.0329***	0.1421***	0.0354***
	(0.006)	(0.012)	(0.008)
Country-year FE	Yes	Yes	Yes
Industry(SIC-4D)-year FE	Yes	Yes	Yes
Underidentification test	51.9	51.9	51.9
Weak identification test	15,609.7	15,609.7	15,609.7
No. of clusters	471	471	471
Observations	225,060	225,060	225,060
Unique Firms	45,012	45,012	45,012

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the SIC4 industry level. Column (1) is our baseline result and the dependent variable in Column (2) is the share of patents granted in textile-intensive technology areas. Column (3) has the HHI concentration index as dependent variable. All specifications include country-year fixed effects as well as a full set of year dummies interacted with a full set of industry 4-digit dummies.

Table 1.8: Citation Behavior

DV: Δ in	% Cit to TXT	TechDist		GeoDist	
	(1) all	(2) all	(3) large firms	(4) all	(5) large firms
ΔlnPat of TXT firms	0.0637**	−0.0070	−0.0884**	0.0032	−0.0988**
	(0.031)	(0.030)	(0.041)	(0.028)	(0.040)
Country-Year FE	Yes	Yes	Yes	Yes	Yes
Industry(SIC-4D)-Year FE	Yes	Yes	Yes	Yes	Yes
Underidentification test	36.5	30.7	19.4	31.3	20.4
Weak identification test	2,084.4	1,131.4	1,344.5	1,268.2	1,274.7
No. of clusters	220	186	151	182	148
Observations	11,649	8,163	4,518	7,902	4,366
Unique Firms	5,567	3,895	1,742	3,760	1,684

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the SIC4 industry level. This table addresses the refocusing of citation behavior. In Column (1), the dependent variable is the share of citations made to textile firms. Columns (2) and (3) look at the technological distance to cited patents, for the full sample and for the 25% of largest non-textile firms, respectively. Columns (4) and (5) repeat the exercise but with geographical distance to cited patents instead. All specifications include country-year fixed effects as well as a full set of year dummies interacted with a full set of industry 4-digit dummies.

Alternative Measure of Technological Proximity. In addition to calculating the pairwise technological proximity of two firms based on the overlap of their patent filings, we also consider calculating it based on the overlap in their patent citation behavior. The idea is that technological proximity could also be understood as two firms building on and referring to the same type of pre-existing technology. Based on their respective patent references, we can calculate the share of citations a firm makes to a certain technology class. Similar to above, we can construct a vector of relative patent citations made to the different patent technology classes for each firm, and again calculate the pairwise un-centred correlation between any two firms' citation behavior vectors. Appendix Table A-5 shows that the results are robust; the estimated effect is smaller in size but qualitatively unchanged. We further check for robustness by using the more fine-grained CPC scheme, instead of the TF34 scheme, for the calculation of the technological proximity measure. We find in non-reported results that the results also hold.

1.4.3 Some 'Macro' Considerations: Industry-Level Analysis & Regional Effects

Industry-Level Regressions Accounting for Vertical Relationships. In all the previous tables we conducted a within-firm analysis of non-textile companies in which all firms are given the same weight. Industry-level estimations might differ, though, because more weight is assigned to larger firms within the industry and if, for example, large and small firms respond differently to changes in textile patenting. From a policy perspective it is relevant to understand the evolution of non-textile patenting at the industry level in addition to the firm level. In Table 1.9, we run industry-level regressions which implicitly give more weight to large firms. We assign each firm to its main SIC4 industrial category, and then estimate the 5-year within-industry changes in non-textile patenting instead of the within-firm changes.[19]

Table 1.9 presents OLS estimations for a panel of SIC4 industries over time, and all estimations include a complete set of SIC2 dummies interacted with a complete set of year dummies. Column (1) presents the baseline result where the main regressor of interest is now the 5-year log change in tech-weighted patenting in textile industries; the new technological distances are constructed across industry pairs as opposed to firm pairs. A decrease in (tech-weighted) textile patenting by one standard deviation is associated with a 3.66% of a standard deviation decrease in non-textile patenting. This result, in which the estimated coefficient is three times larger than in our firm-level sample, is consistent with our previous finding in Table A-2 that large firms reduce their non-textile patenting by a greater degree than small firms.

The remaining columns aim to account for vertical linkages. In Column (2) we add a regressor capturing changes in sales by the textile sector, weighted by how much a non-textile industry supplies its output to the textile sector. In Column (3) we add a similar regressor,

[19]Note that the log of industry-level sum of patents is not the same as the sum of the firm-level log of patents (summed to the industry level), as per Jensen's Inequality.

Table 1.9: Industry-Level (SIC4) Regressions: OLS

DepVar: ΔlnPat of NTXT industries (SIC4)	(1) Base	(2) Output (alpha)	(3) Input (sigma)	(4) Input & Output
ΔlnPat of $tech_{ij}$-weighted TXT industry	0.0366*	0.0360*	0.0366*	0.0358*
	(0.020)	(0.020)	(0.020)	(0.020)
ΔlnY of IO_{ij}output-weighted TXT industry		4.8856		7.1522
		(6.236)		(8.039)
ΔlnY of IO_{ij}input-weighted TXT industry			1.0258	−4.0132
			(6.966)	(8.978)
Industry(SIC-2D)-Year FE	Yes	Yes	Yes	Yes
Observations	2,065	2,065	2,065	2,065
Unique SIC4 Industries	413	413	413	413

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the SIC4 industry level. This table presents industry-level regressions, including 2-digit industry-year fixed effects in all specifications. The dependent variable is the change in the log of patents by non-textile firms in each 4-digit industry. The regressor of interest is the change in the log of patents by textile firms, weighted by the technological proximity between a given non-textile 4-digit industry and each textile 4-digit industry. Similar to Table 1.6, Columns (2)-(4) control for industrial input-output relationships.

but now weighted by how relevant textile products are as inputs for each of the non-textile industries. Column (4) incorporates both variables. Estimation results for our regressor of interest are unchanged, supporting the idea that industrial linkages cannot explain our finding.

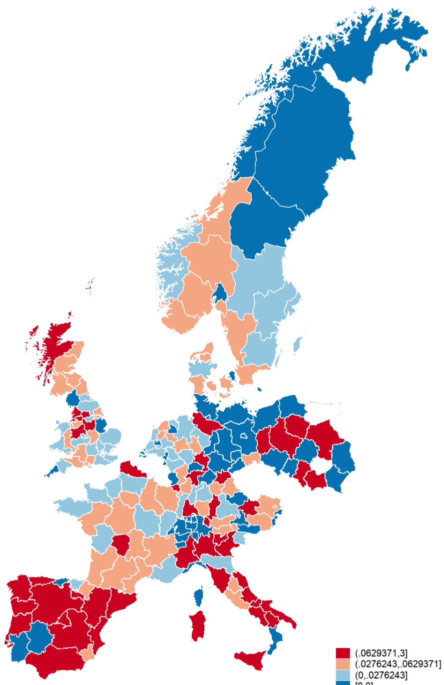
Aggregate Magnitudes at the Regional Level. In a final exercise, we aggregate the data to the regional NUTS3 level to gauge the relative effect sizes of the removal of import quotas on textile firms (direct effects) and non-textile firms (indirect effects). In Figure 1.4, we plot the sample countries by the concentration of textile firms, on the one hand, and by the relative magnitude of the direct effect in each region, on the other hand. Chart (a) shows the same map as in Figure 1.1. Chart (b) displays the relative magnitude of the direct effect of the removal of import quotas on textile firms compared to the indirect effect on non-textile firms.

While there is a positive correlation with Panel (a), strong direct effects are not confined to those regions with a high concentration of textile firms but observable across Europe. This can be explained by two factors. First, the intensity of quotas removal differed across 4-digit industries within the textile sector. Second, the various degrees of the indirect effect on non-textile firms reflect their different technological exposure to textile firms.

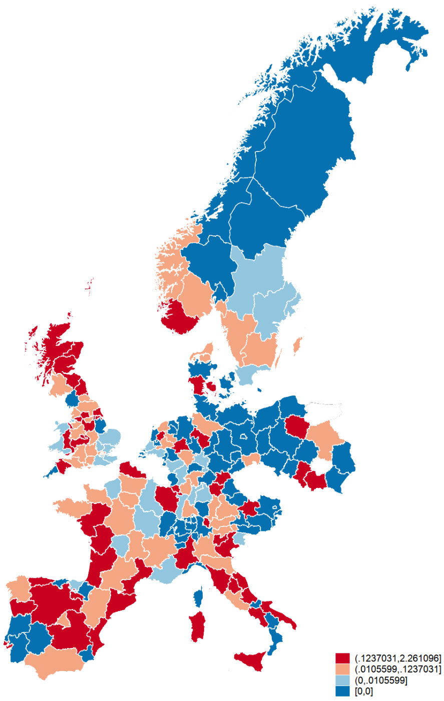
Table 1.10 compares the direct and indirect effects of the removal of import quotas on textiles from China in a five-year window. We limit our sample to NUTS3 regions with a presence of textile firms (247 regions) and condition on firms with non-missing financial data in ORBIS. In Panel (A) we start with a micro approach by documenting the predicted percentage change in annual patent filings at the firm level. To estimate the direct effect, we generate a panel of textile firms over time and estimate a reduced form specification of the change in the log

Figure 1.4: Aggregate Effects at the Regional (NUTS2) Level

(a) Number of Textile vs. Non-Textile Firms



(b) Direct vs. Indirect Effects



Notes: The Figure maps the regions at the NUTS2 regional level for illustrative reasons, while our underlying analysis in Table 1.10 is undertaken at the more disaggregate NUTS3 regional level.

of patents by textile firms on the toughness of quotas in 2000.[20] The predicted value implies a 5.7% (4.6%) increase in patenting by the mean (median) textile firm after China's entry into the WTO. For an estimation of the indirect effect, we again use the panel of non-textile firms to which we have been referring to throughout the paper. The predicted value of the indirect effect states that the mean (median) non-textile firm reduces its patenting by 1.0% (0.7%) after China's entry. Overall, it is reassuring that the direct effect is much larger than the indirect effect (between six and seven times).

Table 1.10: Direct vs. Indirect Effects at the Regional (NUTS3) Level

Predicted change in patent filings over 5-years ($\hat{y}$)	Mean	Median
Panel A: Percentage change at the firm level		
Direct effect	0.057	0.046
Indirect effect	-0.010	-0.007
Panel B: Percentage change at the NUTS3 level		
Direct effect	0.048	0.041
Indirect effect	-0.010	-0.010
Panel C: Level change at the NUTS3 level		
Direct effect	0.458	0.122
Indirect effect	-2.419	-0.341
Ratio of indirect to direct effect	5.3 times	2.8 times

Notes: This table compares the predicted change on patenting by textile firms (direct effect) and on patenting by non-textile firms (indirect effect), as a consequence of the removal of import quotas on Chinese textiles, over a 5-year period.

In Panels (B) and (C) we aggregate the data to the regional NUTS3 level, the former displaying percentage changes and the latter presenting level changes. The mean (median) NUTS3 region experiences a 4.8% (4.1%) increase in patents by textile firms after the shock (*direct effect*) and, at the same time, a 1.0% (1.0%) reduction in non-textile patenting (*indirect effect*). Magnitudes are reassuringly similar to firm-level estimates in Panel (A), implying that our firm-level results are not driven by few very large firms or by specific regions.

Both panels (A) and (B) present percentage changes, within firm and within region, respectively. In both cases, the direct effect is an order of magnitude larger than the indirect effect. To account for the fact that there are many more non-textile than textile firms in an economy, Panel (C) displays predicted level changes in the *absolute* number of patents after the China shock. In the average (median) NUTS3 region, textile firms produce an additional 0.458 (0.122) patents (*direct effect*), while non-textile firms reduce their patenting by 2.419 (0.341) units (*indirect effect*). Consequently, once we account for the relative abundance of non-textile firms in an economy, the indirect effect can be around three to five times larger than the direct effect.

[20] See Appendix Table A.1 for a replication of the BDVR results with the textile firms in our sample and a reconciliation with our first stage results.

1.5 Conclusion

Recent academic work seems to provide different viewpoints on industrial policy: while the innovation literature tends to emphasize the benefits of cross-pollination of ideas across a variety of sectors and the benefits of being exposed to other technology classes, traditional work in the trade literature rather highlights the gains from specialization. This paper addresses this debate by estimating and quantifying the impact of an industry-specific policy shock on the innovation undertaken in technologically proximate but non-targeted industries.

We construct a firm-level manufacturing panel of 13 European countries with information on innovation activity, knowledge sourcing, and technological distances across each pair of firms and sectors. We use the removal of import quotas on Chinese textiles in 2001 as an exogenous competition shock to the European textile sector to help identify the induced changes in innovation of non-textile firms through technological linkages to textile firms (Bloom et al., 2013).

Our key result is that the shock induced by the removal of Chinese import quotas on textiles propagates through technological linkages across the network of innovating non-textile firms. While the direct effect of this removal increases innovation by the average European textile firm (Bloom et al., 2016), the indirect effect is negative once we account for the centrality of each textile firm in the knowledge network. This negative effect increases in geographical and technological proximity to textile firms. The results are robust when accounting for vertical input-output linkages to the textile sector. The effect persists across all firm sizes and upholds when aggregating the panel to the 4-digit industry level, suggesting that results are not driven by a few large firms. Moreover, our analysis shows that non-textile firms shift their innovation away from 'textile-intensive' technology areas, cite fewer patents from textile firms, and instead turn to new sources of knowledge that are further away in both the geographical and technological space.

Our results highlight the importance of accounting not only for the direct effect, but also for indirect effects, when evaluating the implications of industrial or trade policies that theoretically aim at targeting specific sectors. The absolute magnitude of the indirect effects is sizeable once we account for the larger fraction of non-textile compared to textile firms. In the median NUTS3 region, the indirect effect is around three times larger than the direct effect. By emphasizing the fact that not all textile firms have the same centrality index in the technological proximity network and by explicitly modelling knowledge spillovers across firms and industries, we aim to incorporate some of the general equilibrium effects of an industry-specific shock.

There are a number of avenues that could be explored in future work. It would be interesting to complement our study with an assessment of inventor mobility following the policy shock. Our data do not allow us to undertake a detailed assessment of employment effects. To the extent possible, we test for potential labor reallocation within the local labor market as an alternative mechanism through our geographical analysis. Our estimated coefficients based on firms within a 50km radius (as proxy for the local labor market) versus firms outside a 50km radius are very similar to our baseline. This tentatively speaks against a major labor reallocation of inventors. Future research using worker-level data, which may account for labor mobility, could provide additional insights. We are also considering extending our empirical analysis with a theoretical model that allows for increase in Chinese import competition to have conflicting direct effects on European firms and knowledge spillovers across industries that depend on technological proximity. A general equilibrium model could help parsing out potential mechanisms related to labor reallocation effects and allow for counterfactual analysis. Our empirical strategy uses a policy shock that targeted the European textile sector. The textile sector is not among the most innovation intensive sectors, hence our results can be considered a 'lower bound'. Further analysis exploring policy shocks in the context of other industries would be valuable.

2

Multinational Innovation and China's Science and Technology Boom

2.1 Introduction

"China declared world's largest producer of scientific articles" – this was the headline of a Nature article in January 2018 and summarized a key finding of the recent report by the US National Science Foundation (NSF, 2018). For the first time, China published more scientific articles than any other country in the world, overtaking the United States. China's domestic R&D expenditures increased with an average annual growth rate of 18% between 2000-2015, accounting for one-third of total growth of worldwide R&D expenditures. This reflects the rapid development of China's science and technology capabilities. At the same time, we observe an internationalization of R&D. Firms have been increasingly conducting R&D abroad, including in emerging economies.[1] Multinational enterprises (MNEs) are among the most innovative firms in an economy and play an important role in the international diffusion of knowledge.[2] As seminal work by Griffith et al. (2006) and Branstetter (2006) shows, foreign direct investments and basing inventors abroad can be an effective way for technology sourcing. Foreign subsidiaries can serve as a channel for productivity and knowledge spillovers from destination country to parent firm.

In this paper, I assess how China's science and technology boom since the mid-2000s affected the innovation activities of multinational firms in an advanced economy like Germany. I exploit China's major science and technology policy programme in 2006 as a quasi-natural experiment to identify the causal relationship between China's rise in science and German multinational firms' innovation and knowledge sourcing behavior. The national S&T strategy set its goal to transform China into a major centre of innovation by 2020 and a global leader in science and technology by 2050, and identified a concrete list of specific subjects and initiatives for targeted investment. China's Medium- and Long-Term Plan for the Development of Science and Technology 2006-2020 (or short, MLP) prompted one of the major S&T booms in recent history and an unprecedented growth of scientific articles from China.

While China's innovation processes and policies have received increasing scholarly interest,[3] to the best of my knowledge, this is the first paper to examine the causal effect of China's S&T boom on firms' innovation strategies in advanced economies. I assess if German MNEs with FDI in China offshore more of their R&D activities to China, as measured by the number of patents invented in China, and if they rely more on Chinese scientific knowledge in their innovation activities, as measured by the number of references to Chinese scientific articles by patents.

This paper creates a unique dataset that links FDI, patent and bibliometric data, which captures not only foreign investment and inventive activities of multinational firms but also relates

[1] See e.g., Hall (2011); Zhao (2006); Goldberg et al. (2010); OECD (2017); NSF (2018).

[2] See e.g., UNCTAD (2008); Criscuolo et al. (2010).

[3] See e.g., Freeman and Huang (2015); Boeing et al. (2016); Howell (2018); Jiang et al. (2019); Jia et al. (2019).

corporate innovation to science. I match three main databases: administrative data on foreign direct investments by German firms from Deutsche Bundesbank's "Microdatabase Direct Investment", patent data from PATSTAT, and scientific article data from Web of Science. The data is novel in at least two regards. First, while recent papers have used linked FDI and patent data to study multinational innovation by US firms,[4] the linkage of patent data to administrative FDI data is novel in the European context. Second, I further combine it with novel data on non-patent references to scientific articles, which have recently been used to document the relationship between science and technology development.[5] For the first time, I map this type of data specifically to patents by multinational firms, which enables me to shed light on the link between academic research and corporate innovation for some of the most innovation intensive firms.

To relate the development of science fields targeted by China's MLP policy to MNE's innovation strategies and knowledge sourcing behavior, I construct a technology-to-science concordance matrix. Based on 9.7 million non-patent references, I calculate the linkage intensities between technology classes and scientific disciplines. The matrix reflects the relevance of a scientific discipline for patents from a certain technology class, and is used to capture the differences between firms in their exposure to the MLP policy and China's scientific progress. This allows me to explicitly model a firm's absorptive capacity and potential learning from science in the empirical analysis. While concordance matrices that relate technology classes to industries exist, there is currently no concordance matrix available that relates technology classes to science at large scale. This paper presents a technology-to-science concordance matrix that could be more broadly applied in other papers studying the relationship between technology and science.

I use a triple differences specification to identify the effect of China's MLP policy on German multinationals' inventive activities, based on the 2006 MLP policy shock, firm's different exposure to the MLP targeted scientific fields, and considering MNEs in and outside China. I find no evidence that MNEs with FDI in China offshore more of their R&D activities to China if they have a technological profile that closely relates to the MLP targeted scientific fields, relative to their peers with a lower exposure to the MLP policy. Similarly, I find no evidence that the high MLP exposure MNEs in China increase their citations to Chinese scientific knowledge by more than their low MLP exposure peers. In both instances, the estimated effect is negative and statistically insignificant. A positive differential effect on the MNEs with high MLP exposure cannot be ruled out, but is at most modest at 7% at the top end of the 95% confidence interval. These results are in line with existing literature on technology sourcing through FDI, as MNEs with subsidiaries in China have a stronger growth in patents invented in China as well as in citations to Chinese scientific articles, compared to MNEs outside China. However, among MNEs with FDI in China, there is no difference as to whether a firm has a technological

[4] See e.g., Bilir and Morales (2020); Berry and Kaul (2015).

[5] See e.g., Ahmadpoor and Jones (2017); Poege et al. (2019); Watzinger and Schnitzer (2019).

profile that closely relates to the MLP targeted scientific fields or not. The results are robust to a number of alternative specifications.

This paper relates to the literature on technology sourcing through FDI and knowledge flows associated with international R&D. The idea that there are spillover benefits from FDI and that multinational firms can tap into remote knowledge through the establishment of subsidiaries is not new and has been subject of a large literature in both economics and international business (e.g., Cantwell, 1989; Almeida, 1996; Aitken and Harrison, 1999; Chung and Alcácer, 2002; Javorcik, 2004; Haskel et al., 2007; Keller and Yeaple, 2009).[6] Knowledge diffusion is partially localized, hence establishing a subsidiary in a foreign country facilitates access to local knowledge, especially when knowledge is tacit (Jaffe et al., 1993; Audretsch and Feldman, 1996; Almeida and Kogut, 1999; Branstetter, 2001; Keller, 2002).

This paper is closely related to that of Griffith et al. (2006) who provide evidence for technology sourcing through R&D activities abroad. They find that UK firms with more inventors based in the United States benefited more from the growth in US R&D stock in the 1990s in terms of firm-level productivity. Branstetter (2006) shows that foreign subsidiaries can serve as a channel of knowledge spillover from destination country to investing firm, based on citation patterns of Japanese MNEs with affiliates in the US. While Griffith et al. (2006) examine the special relationship between the UK and the US, I shed light on the relationship between Germany and China – two of the largest economies in the world that view each other as important trading and political partners. The Chinese context allows me to exploit a policy-induced growth in science and technology. Furthermore, it allows me to examine if knowledge sourcing through outward FDI can also be observed in a destination country that is still in a development process.

This paper is also related to the literature on the internationalization of R&D and the organization of innovation in a multinational firm. Collaborations in form of patent co-inventions and inventor mobility represent important channels for the transmission of knowledge, and help explain the geographical localization of knowledge flows and spillovers within firm boundaries (Singh, 2005; Breschi and Lissoni, 2005). Branstetter et al. (2014) document the surge of Chinese and Indian patents granted at the USPTO and find that the majority of these are by local inventors working for MNEs from advanced industrial economies and in international collaboration. In a recent paper on the organization of innovation in the global firm, Bilir and Morales (2020) estimate the productivity gains of R&D in affiliates and in the headquarters. They find that headquarter R&D is a substantially more important determinant of firm performance than affiliate R&D.

As Hall (2011) points out, there is clear evidence that R&D is becoming more internationalized, however, data on internationalization is often not ideal. Most statistics on the internationalization of R&D take on an inward perspective reflecting on the percent of inland R&D expenditures

[6] See Keller (2010) for a survey on the literature on technology spillovers through trade and FDI.

received from firms abroad (OECD, 2017; European Commission, 2016). The firm-level data in this paper allow to gauge the share of inventive activities that domestic firms undertake abroad. The outward perspective adds relevant evidence to the public policy debate regarding offshoring of R&D activities, which is commonly met with concern regarding the loss of high value jobs and of technological capabilities in the home country. The paper contributes by showing that China's science and technology boom is not associated with a growing share of offshore versus domestic inventions in the multinational firm.

This paper speaks to the growing literature on the value of science for innovation and productivity. Based on references from patents to scientific publications, recent papers demonstrate that patents closely related to science are more valuable (Ahmadpoor and Jones, 2017), while the quality of the underlying scientific contributions drive the value of inventions (Poege et al., 2019). Science-based patents are also found to be more novel and more valuable in monetary terms (Watzinger and Schnitzer, 2019). This suggests that access to science can be important for corporate innovation. Kantor and Whalley (2019) provide causal evidence for the importance of proximity to scientific research for productivity growth based on historical data from agriculture. This paper contributes by capturing potential benefits of technological proximity to China's science boom for multinational firms. In particular, I explicitly account for firms' different R&D profile and hence capabilities to absorb and use academic knowledge (see e.g., Cohen and Levinthal, 1989; Griffith et al., 2004).

A significant body of literature has examined the effects of China's rise in production and exports to other countries' manufacturing and innovation.[7] Other papers have documented China's R&D explosion and rise in global science (see e.g., Boeing et al., 2016; Freeman and Huang, 2015; NSF, 2018). But the effects of China's rise in scientific prowess on innovation in other countries have so far received little attention. This study seeks to fill this gap in the literature, being set in a context where China transitions from a position of global "production bench" to a rising science and technology nation.

The remainder of this paper is structured as follows: Section 2.2 provides details on China's S&T policy shock from 2006. Section 2.3 describes the data and shows a set of stylized facts on multinational firms, FDI and their global inventive activities. Section 2.4 describes the empirical strategy. Section 2.5 then presents the econometric analysis and results. Section 2.6 concludes.

[7] See Shu and Steinwender (2019) for a review.

2.2 Proximity to Foreign Science: China's MLP 2006 as a Quasi-Natural Experiment

I use China's major science and technology (S&T) programme in 2006 as a quasi-natural experiment to study whether German multinational firms with a subsidiary in China increased innovation following the major science and technology boom in China. In the following, I provide details on China's pivotal S&T strategy from 2006, the policy-making process and argue that this setting is suited for assessing the impact of FDI destination countries' science and technology development on multinational innovation. Moreover, I discuss the growing importance of China as a FDI destination country for German multinational firms.

2.2.1 China's National Science & Technology Strategy: MLP 2006

In early 2006, the State Council of China issued the "Medium- and Long-Term Plan for the Development of Science and Technology (2006-2020)" (hereafter MLP) – the national strategy to transform China into a major centre of innovation by 2020 and a global leader in science and technology by 2050.[8] The Chinese government put the strengthening of "indigenous innovation" capability at the centre of the plan, in its aim to reduce China's reliance on foreign technology and to leapfrog into leading positions in priority fields of strategic relevance. The MLP constitutes a milestone in China's innovation strategy and laid out China's science and technology policy for the coming decade (Gu et al., 2009; Liu et al., 2011).

The major development goals include the establishment of world-class research institutions, emphasizing the promotion of China's strength in basic research and frontier technology development in the areas of equipment manufacturing, information technology, energy- and resource efficient technologies, as well as biotechnology and national defence. The MLP identifies a number of specific initiatives for investment in science and technology that are organized along the sections Main Areas and Priority Topics, Major Special Projects, Frontier Technologies, and Basic Research in the document.[9] Besides these specifically targeted fields, the plan sets out measures that concern the regulatory and institutional framework in support of an efficient national innovation system.

Special government funding plans complement the MLP strategy. In order to turn the plan into action, the Chinese government formulated a range of "Measures for the Management of Specific Funds", such as the so-called 863 Plan for National High-Tech Research and Develop-

[8] "China's National Medium- to Long-term Plan for the Development of Science and Technology (2006-2020)", issued by the State Council of China, 2006. The original document in Chinese is available at http://www.gov.cn/jrzg/2006-02/09/content_183787.htm. An English translation of the document is available at https://www.itu.int/en/ITU-D/Cybersecurity/Documents/National_Strategies_Repository/China_2006.pdf. Web links last accessed on 24 Feb 2020.

[9] For a more detailed discussion of the MLP, see e.g., Gu et al. (2009) or Liu et al. (2011).

ment or the 973 Plan for National Basic Research and Development.[10] These specific funds were to avoid overlaps with other science and technology funding plans, suggesting that these were dedicated resources for the MLP targeted fields.[11]

The development of the MLP and the identification of key topics was led by a steering group of 26 experts, consisting of the three presidents of the Chinese Academies of Sciences (CAS), Social Sciences (CASS) and Engineering (CAE), the president of the National Science Foundation of China and government officials at or above the ministerial level.[12] Chaired by the then Prime Minister Wen Jiabao, the steering group initially identified 20 key R&D issues and commissioned studies on these strategic topics to be evaluated from both a scientific and socio-economic perspective. More than 2,000 experts were involved in further consultation and review of the reports during discussion forums and research retreats, some of which lasted more than a month.[13] Overall, the process of the MLP, especially in the early stages, can be regarded as expert- and scholarly-led, with some participants comparing the research meetings to a temporary think tank.[14]

The plan was drafted under the lead of the Ministry of Science and Technology and presented in early 2006 by the State Council. The MLP framework translated into 79 supporting policies for implementation, which account for more than one-third of the innovation policies that China's government had issued over the past 25 years (Liu et al., 2011). The level of coordination across multiple agencies for the development of these policies is unprecedented in the history of China's S&T policy, and reflects the importance and the top-down implementation of the MLP by the State Council.[15]

The MLP serves as a quasi-natural experiment in this study given its nature of an exogenous top-down innovation policy in a state capitalism model, where the central government takes influence in promoting areas of strategic national interest (Xu, 2011). Based on the accounts

[10]The "Measures for the Management of Specific Funds for Scientific Research in the Public Welfare Industry" (http://www.gov.cn/ztzl/kjfzgh/content_883846.htm) and for the "National Science and Technology Support Plan" (http://www.gov.cn/ztzl/kjfzgh/content_883833.htm) support the implementation of the third section of the MLP 'Main Areas and Priority Topics', the 863 Plan is a special funding plan for the 'Frontier Technologies' mentioned in the fifth section of the MLP (http://www.gov.cn/ztzl/kjfzgh/content_883839.htm), while the 973 Plan is associated with the initiatives set out in the sixth section 'Basic Research' (http://www.gov.cn/ztzl/kjfzgh/content_883822.htm). It is also officially reported that the state has invested tens of billions in the 'Major Special Projects' set out in the fourth section of the MLP (https://www.sciping.com/17994.html). Web links last accessed on 24 Feb 2020.

[11]http://www.gov.cn/ztzl/kjfzgh/content_883846.htm, Article 12, last accessed on 24 Feb 2020.

[12]http://www.gov.cn/zwgk/2005-08/12/content_22217.htm, last accessed on 24 Feb 2020.

[13]http://www.cas.cn/xw/kjsm/gndt/200906/t20090608_641015.shtml, last accessed on 24 Feb 2020.

[14]Zuye Zou, the Director of Beijing Municipal Science and Technology Commission, for example commented: "We dont have any weathervane. We just research the scientific properties of the projects, study the global development trend, and study the problems in China. We dare to reveal the facts. Experts in such a great amount make me feel like I am in a temporary think tank. Everyone has his or her own opinions. Everyone has the right to agree or disagree with others. We design questionnaires, conduct surveys, and then analyze them. The formation of consensus within the group is also the result of multiple collisions." http://www.cas.cn/xw/kjsm/gndt/200906/t20090608_641015.shtml, last accessed on 24 Feb 2020.

[15]In particular, the National Development and Reform Commission (NDRC), the Ministry of Science and Technology and the Ministry of Finance have had lead roles in developing and implementing the concrete policies.

set out above, the MLP policy can be plausibly assumed exogenous for German multinational firms that are at the centre of the empirical analysis.[16] There is no evidence that multinational firms in general, or German MNEs in particular, had influence on the Chinese government for the MLP to favourably promote those fields that would have benefited them most.

2.2.2 China as a Popular Destination for German FDI

The top destination countries for German foreign direct investments are still the US, France and other neighbouring European countries. However, over the last decade, China has become a FDI destination country of growing importance for German multinational firms, both in terms of investment volume as well as number of German multinationals with subsidiaries in China. According to official statistics by the Deutsche Bundesbank, Germany's stock of outward FDI in China amounted to 77 billion in 2016, which is about eight times as high as in 1999.[17] While in 1999 there were less than 200 German investors with subsidiaries in China, the number increased to more than 1000 German parent firms with nearly 2200 Chinese subsidiaries in 2016; together they employed 748,000 employees and had a turnover of 280 billion in China (Deutsche Bundesbank, 2019). This study focuses on the German multinational firms that had already established a FDI presence in China prior to the MLP policy in 2006. With its focus on Germany and China, this paper explores the investment and innovation relationship between two of the largest economies in the world. The following chapter describes the data for the empirical analysis in detail.

2.3 Data & Descriptive Analysis

For the empirical analysis, I combine data from three main sources. First, I use Deutsche Bundesbank's "Microdatabase Direct Investment" to obtain data on foreign direct investments of German firms abroad, particularly in China. Second, I use the European Patent Office's PATSTAT database to collect patent and inventor information for these firms. Third, I use bibliographic information from the Web of Science database to measure the scientific strength of China relative to the rest of the world. A fourth database, Bureau van Dijk's ORBIS, is used to establish the link between Deutsche Bundesbank's firm level FDI data and PATSTAT's patent data. I provide more details on each of the databases below, before I proceed and present some stylized facts on multinationals, FDI and their global inventive activities.

[16] Other papers that have used the MLP policy as an exogenous policy and quasi-experiment for their research design include Jia et al. (2019) and De Rassenfosse and Raiteri (2017).

[17] This includes primary investments (direct participation) in non-holding companies plus secondary investments (indirect participation) via holding companies. Direct FDI according to the Bundesbank 'asset/liability-principle' corresponds to attributable equity capital plus debts to shareholders or affiliated enterprises. Germany's global stock of outward FDI totalled 1,153 billion in 2016, with 477 billion to EU countries and 326 billion to the USA.

2.3.1 Data Sources

Microdatabase Direct Investment (MiDi): Foreign Direct Investments by German Firms Abroad. The MiDi database of the Deutsche Bundesbank (the German Central Bank) contains annual data on foreign direct investments by German companies since 1999. Originally collected as administrative data for the calculation of official German outward and inward FDI aggregate statistics, it provides a high quality and reliable source of microdata on German FDI for research purposes.[18] In accordance with the German Foreign Trade and Payments regulation ("Aussenwirtschaftsverordnung"), German companies are legally required to report to the Deutsche Bundesbank if they have: (i) direct ownership of at least 10% of the shares or voting rights in a foreign subsidiary with a total balance sheet of more than 3 million Euros; (ii) indirect ownership, or through a mixture of direct and indirect shares, with a majority stake of at least 50% in a foreign subsidiary with a total balance sheet of more than 3 million Euros.

Due to the compulsory nature of reporting, the response rate is regarded as "close to complete" according to the Bundesbank (Drees et al., 2018). MiDi is a 'below firm-level' panel dataset that captures direct and indirect investment relations between a German parent company and its foreign subsidiaries at the 'DE parent firm x year x foreign subsidiary' level. In this paper, I use the 2018 edition of the MiDi database, which covers the years 1999-2016 and includes data on around 13,000 German multinational firms with outward FDI and their linked subsidiaries. MiDi includes information on the country of investment, the number of foreign subsidiaries, the year of entry, participation shares and detailed balance sheet information for the reported subsidiaries.[19] As such, MiDi allows the researcher to draw a comprehensive picture of the multinational firms' activities and ownership structure worldwide. A more detailed description of the MiDi database is provided by Drees et al. (2018).

PATSTAT: The Inventive Activities of Multinational Firms. In terms of patents, I consider all patent families from the European Patent Office's PATSTAT database, with first filing date between 1985 and 2016, which are ultimately owned by a German parent firm in the MiDi. When assessing patent related variables, the unit of analysis is the DOCDB patent family, to which I also interchangeably refer as "patents". The PATSTAT DOCDB database is maintained and updated by the EPO on a weekly basis and includes records from more than 90 patent offices worldwide. For each patent, I determine the country of invention via the inventor

[18]DOI of the MiDi database: 10.12757/Bbk.MiDi.9916.04.05.
The Deutsche Bundesbank collects data on foreign direct investments since 1976. MiDi as a standardized micro-dataset of individual FDI relations, constructed based on this data, is available for the years after 1999.

[19]The raw MiDi database includes data on ca. 17,000 reporting entities. I exclude the ca. 2,000 private and public households that also report to MiDi. In order to obtain a consistent sample for my analysis, I further drop those firms that are legally not required to but voluntarily report to MiDi and only keep those who were not affected by changes to the legal reporting requirement rules in 2002 and in 2007, which reduces the sample by another 2,000 reporting entities.

location information on the patent publication. The inventor team of a patent can come from a single country or multiple countries. I distinguish between fully German, fully Chinese and German-Chinese co-inventor teams as well as international inventor teams with Chinese involvement.[20] The country of invention is not necessarily the country of patent filing nor the country of patent ownership. I am mainly interested in the location of the actual underlying R&D activity within the multinational firm, which is best captured by the inventor location.[21]

Linking MiDi and Patent Data (via Bureau van Dijk's ORBIS). In order to concurrently study the international innovation and FDI activities of multinational firms, I received special permission to match the Deutsche Bundesbank MiDi database with patent data. The link is established via Bureau van Dijk's ORBIS database. On the one side, each German parent company in the MiDi is assigned to a unique ID that can be linked to databases by Bureau Van Dijk (BvD). This record linkage was conducted by the Deutsche Bundesbank Research Data and Service Centre based on supervised machine learning, in a wider effort to link various data sources on company data that are used within Deutsche Bundesbank (for details, see Schild et al., 2017). On the other side, for the patent data, I also have a mapping between PATSTAT and Bureau van Dijk's ORBIS database. Using ORBIS' embedded BvD-to-PATSTAT link, I can assign all patents to the respective patent holding firm. Combining both sides, I can then determine the patent portfolio for each of the German multinationals in MiDi.

The MiDi data has a BvD-link only for the German parent companies, but not for the foreign subsidiaries. Hence, I cannot directly match the subsidiaries in the MiDi to ORBIS/PATSTAT and obtain the patents owned by the subsidiaries. Due to data protection rules, guest researchers at Deutsche Bundesbank are not permitted to see firm names, the region of the affiliates, addresses or any other information that could help identify the entities. Hence, I was not able to undertake a record linkage at the subsidiary level myself. Instead, I use Bureau van Dijk's Ownerhsip database to identify the subsidiaries linked to German parent firms that appear in MiDi. With the resulting subsidiaries' BvD-ID, I can then again use ORBIS' BvD–to–PATSTAT link to retrieve relevant patent information from PATSTAT for the subsidiaries' patents. This way, I identify the patent portfolio of the whole group that belongs to the German parent firm. Figure B-1 in the Appendix illustrates the core aspects of the data linkage between MiDi and PATSTAT/ORBIS.

MiDi contains around 13,000 multinational parent firms with outward FDI and their linked affiliates. For 80% of these German parent firms Deutsche Bundesbank's record linkage can

[20] For the determination of the inventor country information for each DOCDB patent family, I prioritized information contained in patent application documents from the EPO, USPTO, WIPO and German patent office, and consider the application document for which information is most complete and earliest.

[21] See e.g., Li (2017) for a discussion of the differences between patent-applying, patent-inventing and patenting-owning FDI. The spatial organization of patent ownership within a MNE or international patent filing strategies relate to the IP, R&D tax and profit shifting literature, which are not discussed here (see e.g., Gaessler et al., 2018).

assign a unique BvD-ID, such that MiDi data can be linked to Bureau van Dijk's databases. Indeed, of these MiDi parent firms with BvD-ID assignment, I find 99% in my ORBIS sample. This results in a baseline sample of 10,167 German multinational firms, for which I can establish a link between MiDi and ORBIS/PATSTAT.

I combine MiDi and PATSTAT data at the 'firm x year x country' level, where the year corresponds to the first filing year of the patent family, which is closest to the timing of the actual inventive activity. Country corresponds to the country of investment for MiDi data. Multiple affiliates of the same parent firm in one country are aggregated into one unit. With respect to patent data, country corresponds to the country of invention. Patents that have an international inventor team can be assigned to multiple countries. This allows me to determine for each firm and year the intra-firm, cross-country spatial organization of innovation, including the number of patents that were invented with Chinese involvement.

Web of Science: Measuring China's Scientific Strength. I use scientific publication data from Clarivate Analytics' Web of Science (WoS) database to measure China's relative scientific strength over time in the different subjects. WoS is the largest bibliographic database of scientific literature with 43 million research articles published between 1980 and 2016. It provides for each scientific publication all key information such as authors, affiliations, journal, subject field, citations etc. The Web of Science subject classification scheme comprises 251 subject fields in natural sciences and engineering, social sciences, arts and humanities. For each subject field and year (1999-2016), I calculate China's share in top 1% cited papers within the field, which allows me to measure changes in the relative scientific strength of China over time.

Technology-to-Science Concordance: Link between Patents and Bibliometric Data. Recent literature has documented the important relationship between science and technology development as documented in patents (Ahmadpoor and Jones, 2017; Poege et al., 2019; Watzinger and Schnitzer, 2019). Patents not only cite other patents, but also frequently non-patent literature (NPL) such as scientific articles, that relate to the underlying invention. Patent references to scientific articles can be considered an indicator for interaction between science and corporate innovation. I use the number of references to Chinese scientific articles by the sample firms' patents to assess if and to which extent the multinational firms have turned to and drawn on Chinese science in their R&D activities.

While there exist concordance matrices that relate technology classes to industries (Schmoch, 2008; Dorner and Harhoff, 2018), there is - to the best of my knowledge - currently no concordance matrix publicly available that maps the linkage intensities between technology classes and science at large scale.[22] I construct a technology-to-science concordance matrix, based on 9.7 million non-patent references made by 1.4 million DOCDB patent families to scientific articles prior to 2006. I use linked data between DOCDB patents and referenced scientific publications in Web of Science provided by Poege et al. (2019). The matrix reflects the relevance of a scientific discipline for patents from a certain technology class. I use the matrix to capture the differences between firms in their potential to benefit from China's scientific progress. Section 2.4.2 provides more details on the calculation of the technology-to-science concordance table.

2.3.2 Stylized Facts on MNEs, FDI and Global Inventive Activities

Based on the novel linkage between FDI and patent data, I present a set of stylized facts on the FDI and innovation activities of German multinational firms in this section. This study focuses on changes in the innovation and knowledge sourcing behavior of MNEs in response to China's science and technology policy. I hence restrict my sample to firms that have been in MiDi and have patented before 2006, as an indication that these are generally innovating firms. Of the approximately 4,300 German multinational firms in MiDi as of 2006, 57% have patented before. This is a much higher patenting rate than in the overall population of firms. This reflects the well-known fact that multinational firms are much more innovative than non-multinational firms (UNCTAD, 2008; Criscuolo et al., 2010). In the following, I document three facts for the sample of 2,424 patenting MNEs in MiDi, which inform the subsequent empirical strategy.

1. Offshoring of R&D and FDI are not always geographically connected

A substantial share of offshore innovation activities happens in countries where the MNE has no FDI. The mean multinational firm in the sample has FDI in 3.8 countries and inventive activities in 1.5 countries outside Germany. I explore how often only FDI, only innovation or both activities are observed in a country, using the linked panel data at the 'firm x year x country' level.[23] I find that the geographical overlap between FDI and inventive activities abroad is fairly small at 6%. For 69% of the observations, a firm has FDI in the country abroad, but no inventions. In 25% of the cases, we observe foreign inventions despite no FDI presence in that country.[24]

[22]Neuhäusler et al. (2019) provide a probabilistic concordance scheme for the assignment of scientific publications to technology fields, but focus on patents with German inventors and Germany-based authors of scientific articles. Callaert et al. (2006) explore the interaction intensity between technology domains and science fields, however consider science fields at a fairly aggregate level and focus on a subset of EPO and USPTO non-patent references.

[23]For each firm and year, the data includes the union of FDI and innovation countries, while the panel is not balanced with respect to the countries.

[24]That is, the firm has no FDI of substantial value that would need to be reported to the Deutsche Bundesbank according to the legal reporting requirements.

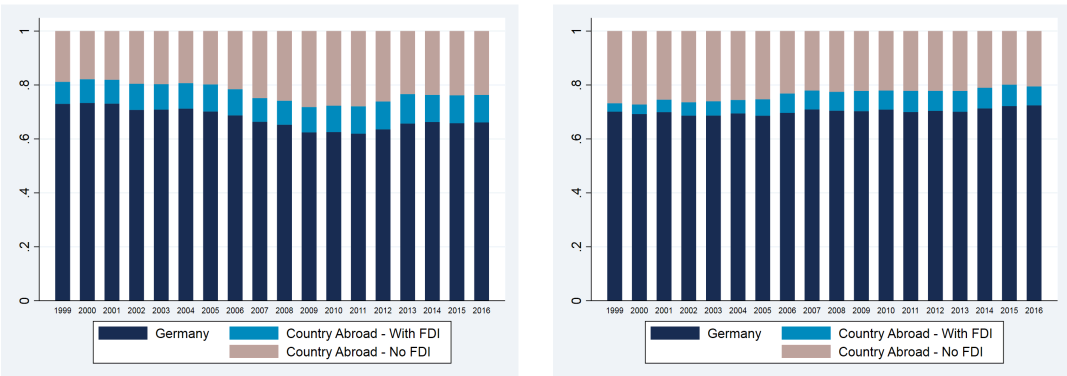
This finding is relevant, as it shows a geographical disconnect between FDI and offshoring of R&D activities. Past literature has relied on patent data to identify international locations of R&D of a firm, implicitly equalling the existence of an inventor at a location to a firm having an affiliate there.[25] With the merged FDI and patent dataset in this study, I can show that there is no perfect overlap between these two. This suggests that solely using patent data or FDI data will lead to missing out on international activities of MNEs. This highlights the advantage of the linked patent and administrative FDI data in this study.

2. There is an increase in the internationalization of R&D, mostly driven by large patent-intensive firms.

About a third of the patents of the MNEs and their subsidiaries are invented abroad. This share has been increasing since the early 2000s, reaching a temporary peak at nearly 40% in 2011, as Figure 2.1 panel (a) shows. Of the patents invented abroad, one third were invented in countries with FDI; the other two thirds were invented in countries where the MNE had no affiliate. This suggests that MNEs without FDI in China may still have patents invented in China.

The percentages in panel (a) are calculated on the overall number of patents held by the MNEs in the sample. If instead, one calculates the share of domestic versus foreign inventions at the individual firm level and takes the average across the MNEs, then we actually do not observe such a pronounced increase in the internationalization of R&D, as shown by Figure 2.1 panel (b). This suggests that within-firm changes towards offshoring of R&D are not very strong, but the aggregate trend is driven by few larger, more patent-intensive firms.

Figure 2.1: Offshoring Inventions



(a) Total

(b) Firm-Level (Average)

Source: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, Microdatabase Direct Investment (MiDi), 1999-2016, PATSTAT, own calculations.

[25]E.g. Griffith et al. (2006), or Blit (2017) for R&D locations and affiliate locations within the US.

3. Firms with FDI in China are generally more internationally active and invested.

As of 2006, 14% of the 2,424 patenting MNEs in the Bundesbank's MiDi had a subsidiary in China. Another 14% entered China after 2006. The remaining MNEs stayed out of China during the entire studied period of 1999-2016. Comparing the firms with FDI in China relative to the firms that never entered China, it becomes apparent that the former are larger, have FDI in more countries, generate more patents and are generally more internationally active. This is not surprising as we know from the literature that MNEs test neighbouring foreign markets first before tapping into a geographically and culturally more distant country like China (Egger et al., 2014).

German multinational enterprises have a long-standing presence in China. Foreign direct investments in the 1990s mostly reflect low-cost production opportunities in China though, rather than motives related to knowledge sourcing, which became more important after 2006 (UNCTAD, 2008). A challenge for any empirical analysis would be that FDI entry is an endogenous choice for a firm, and may thus be correlated to technological shocks that pose a threat to identification.

Before I proceed with detailing the empirical strategy for this paper, I hence impose the following restrictions on the sample to mitigate this problem: I use pre-2006 information to determine the set of firms for which the decision to enter (or stay out of) China is not influenced by China's science and technology boom induced by the MLP. I only consider MNEs that already had a subsidiary in China as of 2006 and MNEs that never entered China, so as to avoid sample selection in relation to the MLP. Firms that entered China post-2006 may have chosen to set up a subsidiary there in response to the science and technology boom induced by the MLP. I further require MNEs without FDI in China to have FDI outside OECD or EU countries, as an indication for their degree of internationalization. This results in a sample of 959 German patenting MNEs, 349 with and 610 without FDI in China.[26] The majority of the MNEs considered are in manufacturing, especially in machinery and equipment, chemicals, metal products, medical, precision and optical instruments, motor vehicles, rubber and plastic products and electrical machinery, or in wholesale and commission trade.[27] Table B-1 in the Appendix provides further descriptive statistics on the sample.

[26]I also undertook the analysis using the wider set of MNEs with no FDI in China, without imposing any conditions. This gave similar results.

[27]A detailed table with the distribution of firms by industry codes is not feasible due to data confidentiality rules of the RDSC of the Deutsche Bundesbank.

2.4 Empirical Strategy

This study examines whether China's rise in science and technology increased innovation for multinational firms in China, and to what extent the inventive activities rely on Chinese scientific knowledge. The empirical specification uses a triple differences model to estimate within firm-changes in innovation and knowledge sourcing. Consider the following regression equation,

$$lny_{it} = \beta\left[MLP_i^{High} \times China_i^{2006} \times Post2006\right] + \alpha_i + \lambda_t + \gamma_{China,t} + \delta_{MLP,t} + \epsilon_{it} \qquad (2.1)$$

where the outcome variable y_{it} measures the number of patents invented in China by the multinational firm i in year t. In addition to estimating the effects on the innovation activities in China, I also use the number of citations to Chinese scientific articles as an outcome variable to capture possible changes in the knowledge sourcing behavior of the multinational firms in China. The model is estimated using data for the period 1999-2016, with seven years before and eleven years after the MLP policy.

The main independent variable is the interaction between MLP_i^{High}, a binary variable that is one if firm i has a high exposure to MLP targeted disciplines (details are described further below), $China_i^{2006}$, which is one if firm i has FDI in China as of 2006, and $Post2006$, which is an indicator variable for the years 2006 and onward. In the baseline, $China_i^{2006}$ is a simple dummy; I also present results where the variable reflects the size of the FDI presence in China as of 2006. Instead of a dummy variable, I replace the $China_i^{2006}$ term with the investment volume or the number of employees in China. The two-way interaction terms of a triple differences model, $ChinaxPost$, $MLPxPost$, as well as $ChinaxMLP$ are technically included, but drop out through the inclusion of firm fixed effects α_i and year fixed effects λ_t. The terms $\gamma_{China,t}$ and $\delta_{MLP,t}$ are group-year fixed effects in relation to the two categorical variables $China$ and MLP. I first set forth the rationale for the triple differences model (DDD), before I proceed with details on the definition and construction of the MLP exposure variable.

2.4.1 Triple Differences Estimation

The empirical strategy generally rests on the idea that firms with a technology profile that closely corresponds to the MLP targeted scientific fields (i.e. higher MLP exposure) are more likely to benefit from China's science and technology boom induced by the MLP. As discussed in Section 2.2, the selection of focal fields in the MLP can be considered exogenous for multinational firms in China. This suggests using a difference-in-differences (DiD) approach for identification, with the varying degree of exposure to the MLP policy as the source of variation in treatment intensity. However, in the simple DiD framework, one may have the concern that unobservable factors unrelated to China's actual MLP policy, such as global technological trends, might differentially affect the innovation activities of high MLP exposure firms relative to low MLP exposure firms. The triple differences approach (DDD) allows high and low MLP exposure firms to have differential evolutions in the outcome variable, but this difference has to be the same for MNEs inside and outside China.

The intuition for identification is as follows: The DDD model compares the changes in innovation of high MLP exposure firms before and after 2006 (first difference) to changes in innovation of low MLP exposure firms (second difference). It then compares this differential impact (of high versus low MLP exposure) for MNEs with FDI in China, relative to the differential impact for MNEs outside of China (third difference). In a two-period setting, this would be:

$$\text{MNEs in China:} \qquad (y_{C,H}^{Post} - y_{C,H}^{Pre}) - (y_{C,L}^{Post} - y_{C,L}^{Pre}) \tag{2.2}$$

$$\text{MNEs outside China:} \qquad (y_{N,H}^{Post} - y_{N,H}^{Pre}) - (y_{N,L}^{Post} - y_{N,L}^{Pre}) \tag{2.3}$$

In a multi-period setting, this results in equation (2.1).[28]

The proposed triple differences approach resolves a number of potential concerns. First, firms with high MLP exposure may have different growth rates in innovation, both inside and outside of China, e.g. due to global trends in science and technology. The DDD model accounts for this, because the β coefficient on the triple interaction term will capture only the differential impact (high versus low MLP exposure) for MNEs with FDI in China relative to the differential impact for MNEs outside China. If firms with high MLP exposure have higher growth rates in innovation both inside and outside of China, then β would be zero. The MLP exposure reflects a firm's technological profile, as discussed in the section below. This means that technology-specific trends are largely controlled for in the DDD approach.

Second, there might be China-specific confounding factors that affect multinationals with FDI in China, but not multinationals outside China. Firms with FDI in China may have systematically higher growth rates in patenting, regardless of their technological profile and hence exposure to the MLP, e.g. due to China-wide policies and trends such as IP laws that are unrelated to the MLP. However, these China-wide effects are expected to affect MLP high and low multinationals in China equally, and would hence be differenced out by the DDD method.

Third, a remaining concern may be that the MLP of the Chinese government targeted science and technology fields that were associated with high growth rates in innovation. That is, one may be concerned about an unobserved factor (such as global trends in certain technology areas) that influences both the inventive activities of the MNEs in China as well as the Chinese government's policy choice. I employ a triple differences approach with a rich set of fixed effects that addresses this concern. As I compare the differential growth for multinationals inside versus outside China, endogeneity is not a concern as long as the MLP did not pick out

[28] See e.g., Imbens and Wooldridge (2007). The triple differences estimation strategy in this paper resembles the triple differences approach by Bernard et al. (2019) in certain aspects.

scientific fields that are more important to MNEs in China than MNEs outside China.

In fact, based on the account of the MLP policy-masking process in Section 2.2 and further analysis that I discuss in more detail in Section 2.4.2, there is no evidence that the MLP favoured the technological and scientific profile of multinational firms in China relative to firms outside China. We also do not observe that MNEs with a stronger linkage to MLP targeted scientific fields anticipated the policy and 'self-selected' themselves into China.

2.4.2 Firm's Exposure to MLP Targeted Scientific Disciplines

I now turn to the details of the MLP exposure variable. MLP exposure is a time-invariant variable that is calculated for each firm. The variation in firms' exposure to MLP targeted scientific disciplines comes through the following three elements: First, firms differ in their technological profile. Second, technology-science linkage intensities vary across technology classes. Third, the MLP targets only certain scientific fields, which in turn also differ in their growth over time. The variable is defined as follows.

$$
\begin{gathered}
MLPexp_i = \sum_{j}^{J=34} \sum_{d}^{D=251} k_j^{2005} \times c_{jd}^{2005} \times I_d^{MLP}, \\
k_j \in [0,1], \text{where } \sum_{j=1}^{J} k_j = 1 \\
c_{jd} \in [0,1], \text{where } \forall j : \sum_{d=1}^{D} c_{jd} = 1
\end{gathered}
\tag{2.4}
$$

Determining a Firm's Technology Profile. For each multinational firm in the sample, I determine the firm-specific technology profile based on its past patent filings, which corresponds to the first term, k_j^{2005}, in equation (2.4). I determine a firm's pre-2006 patent portfolio and construct a vector of patent shares across patent technology classes (denoted with j). The relative patent filings in the respective technology classes represent the firm-specific technology profile. Each DOCDB patent family filed by the German parent firm or by one of its subsidiaries needs to be assigned to a unique patent technology class. I use the technology classification following Schmoch (2008), which assigns each patent to one of 34 different technology classes.

Technology-to-Science Concordance Matrix. The second term in the equation, c_{jd}^{2005}, corresponds to the linkage intensity between technology class j and scientific discipline d from a proprietary technology-to-science concordance matrix, which I calculate on the basis of 9.7 million non-patent literature (NPL) references. The matrix has the dimension 34 (which is the number of technology classes J) x 251 (which is the number of scientific disciplines D).[29] It captures the share of NPL citations that patents from technology class j make to scientific articles from discipline d. Note that the underlying patents and NPL references include the population of pre-2006 DOCDB patent family-to-scientific article pairs in the linked PATSTAT/Web of Science database, and are not confined to patents filed by German multinational firms. This is done with intent, as I want to capture the general relevance of a scientific discipline for a technology class and a firm's potential to benefit from China's scientific progress, and not what German multinational firms have actually cited in the past. As such, the technology-to-science concordance matrix is universal and not subject to endogenous citing behavior of certain firms in the sample. Figure B-2 and Table B-2 in the Appendix illustrate the technology-to-science concordance matrix and the fact that technology classes vary in the degree they build on knowledge from the different scientific disciplines.

Scientific Disciplines Targeted by the MLP. Lastly, the third term in the equation, I_d^{MLP}, indicates if scientific discipline d was targeted by the Chinese MLP policy. The MLP consists of 96 subjects and initiatives identified as focal by the Chinese government, which I map into scientific disciplines in Web of Science. This results in 68 scientific disciplines that are targeted by the MLP.[30] Appendix Table B-3 illustrates this mapping based on some examples.

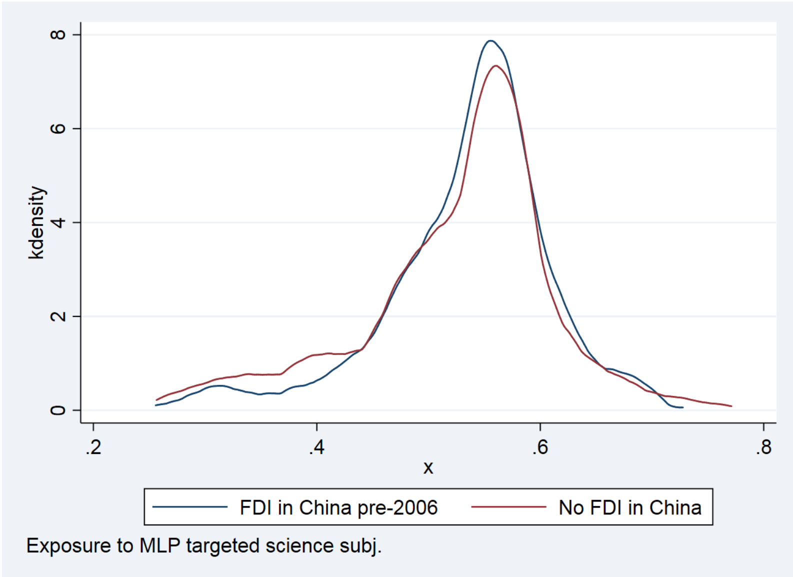
Taken together, the MLP exposure variable relates a firm's technology profile to the scientific disciplines targeted by China's MLP policy. Or in more technical terms, the MLP exposure variable is the technology-to-science concordance matrix, conditional on scientific disciplines targeted by the MLP and weighted according to the relative importance of individual technology classes for a given firm. Through the MLP exposure variable, I explicitly account for heterogeneity in firms' R&D profile and hence capabilities to absorb and use academic knowledge (see e.g., Cohen and Levinthal, 1989; Griffith et al., 2004).

Figure 2.2 shows the distribution of the MLP exposure variable for MNEs in China and outside China. A Kolmogorov-Smirnov test shows that the distribution of the MLP exposure variable for MNEs in China is not statistically different from the distribution for MNEs outside China, which supports the view that the Chinese government did not favour scientific fields that are more important to MNEs that were already in China as of 2006. Also, there is no evidence that firms anticipated the MLP policy and 'self-selected' themselves into China prior to 2006. In the baseline specification, I use a binary version of the variable and split the firms at the median into a high and a low MLP exposure group. In alternative specifications, I use a continuous version of the MLP variable.

[29] I follow Web of Science's classification scheme of scientific disciplines.

[30] A scientific discipline is considered targeted by the MLP if it is associated with at least one MLP policy item. Certain focal disciplines, such as biology, environmental, electrical and mechanical engineering, or environmental science are associated with multiple MLP items. However, I do not apply double-counting or weighting, as a multiple mention in the MLP policy document does not necessarily translate into multiple funding or growth for the scientific discipline. In the absence of precise funding data by scientific discipline, I use an indicator variable approach here.

Figure 2.2: MLP Exposure



Source: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, Microdatabase Direct Investment (MiDi), 1999-2016, Web of Science, PATSTAT, own calculations.

2.5 Results

2.5.1 Baseline Results

Table 2.1 shows regression results from estimating equation (2.1). Columns (1) and (2) show the results for the main dependent variable, with the log number of Chinese based inventions. The β on the triple interaction term MLP x China x Post is negative and insignificant at -0.024 log points (or approximately -2%). This indicates that China's science and technology boom induced by the MLP did not cause a faster growth in Chinese inventions for high MLP exposure firms relative to low MLP exposure firms with FDI in China. The model in column (1) includes firm fixed-effects, year fixed-effects as well as China-year and MLP-year fixed effects; with China-year and MLP-year fixed effects not driving the results as the results in column (2) show. Confidence intervals at the 95% level are reported underneath the point estimates in square brackets. These suggest that a positive effect cannot be ruled out, however, any positive effect will be of no more than 0.073 log points (or approximately 7%) at the top end of the 95% confidence interval.

Columns (3) and (4) use the log number of patent citations to Chinese scientific articles as dependent variable. Also here, the coefficient on the triple interaction term is negative and insignificant, suggesting that the MLP did not change the MNEs to rely and cite more Chinese scientific articles. Again, a positive effect cannot be ruled out, albeit any effect is insignificant and at most modest at 0.019 log points at the top end of the 95% confidence interval. The results do not change if one were to use a inverse hyperbolic sine transformation of the dependent variable instead of the log. The following section further explores the robustness of the baseline results.

Table 2.1: Triple Differences Estimation: Baseline

	Inventions in CN		Citations to CN articles	
	(1)	(2)	(3)	(4)
MLP=1 × China=1 × Post=1	-0.0238	-0.0237	-0.0581	-0.0567
	[-0.120, 0.073]	[-0.120, 0.073]	[-0.135, 0.019]	[-0.134, 0.020]
MLP=1 × Post=1		0.0037		−0.0168
		[-0.033, 0.040]		[-0.049, 0.015]
China=1 × Post=1		0.0961***		0.1016***
		[0.030, 0.162]		[0.038, 0.165]
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
China x Year FE	Yes	No	Yes	No
MLP x Year FE	Yes	No	Yes	No
Observations	15,617	15,617	15,617	15,617
Unique Firms	959	959	959	959
Adj. R-squared	0.68	0.68	0.69	0.68

Source: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, Microdatabase Direct Investment (MiDi), 1999-2016, own calculations.
Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the firm level. Confidence intervals at the 95% level are reported in square brackets. This table presents the baseline results of a linear triple differences estimation for panel data at the firm x year level. The dependent variable is the log number of patents fully or partly invented in China in models (1) and (2), and the log number of patent citations to Chinese scientific articles in models (3) and (4). The preferred baseline specifications are (1) and (3), which include a full set of firm fixed-effects, year fixed-effects as well as a full set of year dummies interacted with China category-dummies and MLP category-dummies, respectively.

Testing for Pre-Trends. First, a potential concern is that high MLP exposure MNEs experience a stronger growth in innovation than low MLP exposure MNEs in China, but not outside China. In other words, there may be pre-existing differences in the time trends of MNE patents invented in China. To address this issue, I allow β_t to vary prior to the 2006 MLP, using 2005 as the baseline year in (2.1).

$$lny_{it} = \beta_t \left[MLP_i^{High} \times China_i^{2006} \times YearPre2006\right] + \alpha_i + \lambda_t + \gamma_{China,t} + \delta_{MLP,t} + \epsilon_{it} \quad (2.1)$$

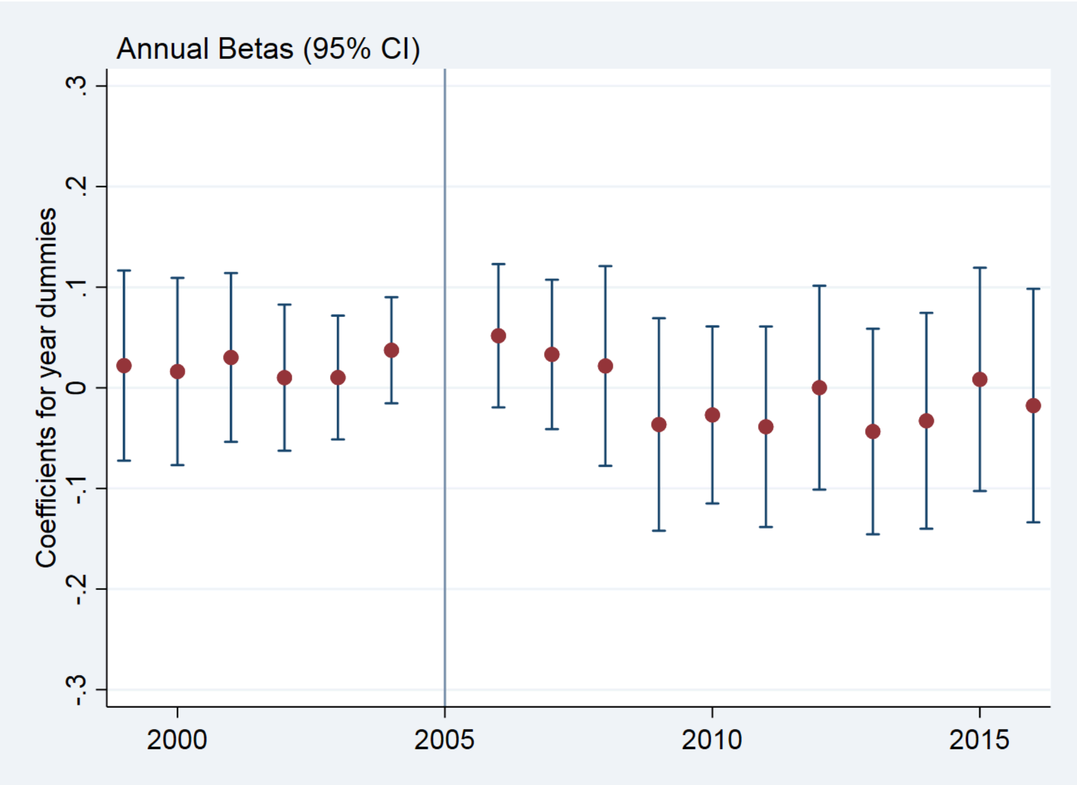
This test reveals no pre-trends in the data. A placebo test with the MLP policy intentionally incorrectly set to 2001, the year of China's accession to the WTO, also confirms that there are no pre-trends.

Measuring Annual Treatment Effects. In addition to average effects, I estimate annual treatment effects to evaluate possible timing aspects of the changes in Chinese invented patents held by the MNEs using the specification in (2.2).

$$lny_{it} = \beta_t \left[MLP_i^{High} \times China_i^{2006} \times YearPost2006 \right] + \alpha_i + \lambda_t + \gamma_{China,t} + \delta_{MLP,t} + \epsilon_{it} \quad (2.2)$$

Figure 2.3 displays the estimated annual β coefficients, which hover all around zero and do not expose a specific pattern.

Figure 2.3: Triple Differences Estimation: Annual Betas



Source: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, Microdatabase Direct Investment (MiDi), 1999-2016, Web of Science, PATSTAT, own calculations. *Notes:* This figure shows the results from estimating the baseline triple differences model, where β_t is allowed to vary by year. The annual betas prior to 2006 test for pre-trends as per equation (2.1), the annual betas after 2006 test for possible timing effects of the treatment as per equation (2.2).

2.5.2 Robustness

'China' as Continuous Variable Accounting for FDI Presence. In the baseline, China is a dummy variable indicating whether a multinational firm had any FDI in China as of 2006. In Table 2.2, I extend the analysis to account for variation in the FDI volume in China. Instead of a binary variable, I replace the term $China_i^{2006}$ with the average stock of investments in million Euros (columns 1 and 4) or with the number of employees the MNE had in China in the pre-period (columns 2 and 5). Across the MNEs with FDI in China, the mean (median)

FDI investment was at 13.3 (2.59) million Euros, with a mean (median) number of employees in China of 286.1 (71.1).

Again, the triple interaction term is negative, and in the case of using the FDI stock even significant at the 10% level.[31] Specifying 'China' as a continuous variable allows me to estimate the two-way interaction term 'China x Post', which is positive in all and significant in most models (1)-(2) and (4)-(5). This suggests that MNEs with a large FDI presence in China increased their Chinese innovation activities by more than MNEs with little or no FDI in China. This result is in line with findings in the literature that there is technology sourcing through FDI (e.g., Griffith et al., 2006; Branstetter, 2006). Exploring the predictive margins, a firm at the 90th percentile of the distribution of employees in Chinese subsidiaries increased the number of inventions in China by 0.084 log points more than a firm at the 10th percentile of the distribution.

'MLP' as Continuous Variable Rather than Binary Variable. Furthermore, I consider a continuous version of the MLP variable in Table 2.2. While the baseline version uses a binary variable that splits the firms into a high and a low MLP exposure group at the median, the models in columns (3) and (6) use the continuous measure of the MLP, the distribution of which is shown above in Figure 2.2 in Section 2.4.2. The estimates are consistent with the baseline result of a negative insignificant coefficient. The two-way interaction term 'MLP x Post' is not significantly different from zero, which suggests that firms do not differ in their growth in inventions in China or citations to Chinese articles due to differences in their exposure to MLP targeted subjects.

In non-reported results, I also consider a categorical specification of the MLP exposure variable with an alternative cut-off to the median used to define the indicator variable, or a finer categorization than a binary one, e.g. by quartiles or deciles. The results are consistent with the baseline.

[31] Comparing the results in columns (1) and (2), the point estimate for the triple interaction term using 'China FDI' (-0.0047, based on the stock of investments in China in million Euros) is about eleven times larger than if using 'China Employment' (-0.0004, based on the number of employees in China). For the interpretation of the magnitude one needs to consider the respective scale of the 'China' variable, though. The 'MLP x China Employment' effect would be of equivalent size as the 'MLP x China FDI' effect as captured by the respective triple interaction terms, if one employee corresponds to ca. €90,000. Note, this should not be interpreted as the marginal effect of one additional employee or one additional unit of investment on the outcome variable.

Table 2.2: Triple Differences: Continuous Variables 'China' and 'MLP'

	Inventions in CN			Citations to CN articles		
	(1)	(2)	(3)	(4)	(5)	(6)
MLP=1 x China FDI x Post=1	−0.0047*			−0.0030**		
	[-0.009,0.000]			[-0.006,-0.000]		
China FDI x Post=1	0.0057**			0.0039***		
	[0.001,0.010]			[0.001,0.006]		
MLP=1 x China Employ. x Post=1		−0.0004			−0.0001	
		[-0.001,0.000]			[-0.000,0.000]	
China Employ. x Post=1		0.0005**			0.0002	
		[0.000,0.001]			[-0.000,0.000]	
MLP (cont.) x China=1 x Post=1			−0.2916			−0.6019*
			[-0.849,0.266]			[-1.256,0.052]
MLP (cont.) x Post=1			0.2564			−0.0078
			[-0.148,0.661]			[-0.341,0.325]
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
China x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
MLP x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,617	15,617	15,617	15,617	15,617	15,617
Unique Firms	959	959	959	959	959	959
Adj. R-squared	0.69	0.71	0.68	0.69	0.69	0.69

Source: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, Microdatabase Direct Investment (MiDi), 1999-2016, Web of Science, PATSTAT, own calculations.

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the firm level. Confidence intervals at the 95% level are reported in square brackets. This table presents linear triple differences estimation results with alternative measures of the terms 'China' and 'MLP'. Models (1) and (4) use the average stock of foreign direct investments in € million the multinational firm had in China in the pre-period (1999-2005). Models (2) and (5) use the average number of employees the multinational firm had in China in the pre-period. Models (3) and (6) differ from the baseline in Table 2.1 in that they consider a continuous measure of the 'MLP' variable instead of a binary variable. China-year and MLP-year fixed effects are defined as in the baseline Table 2.1, hence variation in the two-way interaction term of a continuous 'China' (or continuous 'MLP') variable with the Post-dummy is not fully absorbed.

Heterogeneous Treatment Effects. In Table 2.3, I re-estimate the baseline DDD regression, splitting the sample of multinational firms at the median of turnover abroad, patent intensity (as per patent stock in 2006), years of patenting and whether they had any prior collaboration with Chinese inventors prior to 2006. The dependent variable is always the log number of patents invented in China. The idea is to test whether firms with larger international business activities and more experience in patenting and international collaboration, i.e. firms with presumably higher absorptive capacity, experience a differential effect. Indeed, the results in Table 2.3 show for these larger and more innovative firms a less negative effect, which suggests that they benefited relatively more from China's MLP policy. However, also here the triple interaction term is not significantly different from zero, and negative in all specifications but column (8). For firms with prior collaboration in form of German-Chinese co-inventions before 2006 (column 8), a positive albeit insignificant effect of 0.252 log points is estimated. This suggests that the MLP policy effect on the intensive margin of co-inventions in China is positive, while the overall average effect is negative insignificant as in the baseline.

Accounting for Quality and Actual Growth of MLP Targeted Scientific Disciplines. So far, the MLP exposure variable uses information on whether a given discipline was targeted by the Chinese MLP policy or not. But it does not account for the actual quality level or growth rates of the respective scientific fields in China. The idea is that among MLP targeted disciplines, there might still be differences in the actual quality level and improvements over time. Disciplines may receive substantial funding and increase their scientific output following the MLP policy, however, they may still not be competitive at an international level. A country's distance from the technological frontier can be considered a measure for the potential for technology transfer (Griffith et al., 2004). I therefore estimate the model with a modified definition of the MLP exposure term that aims to account for these aspects.

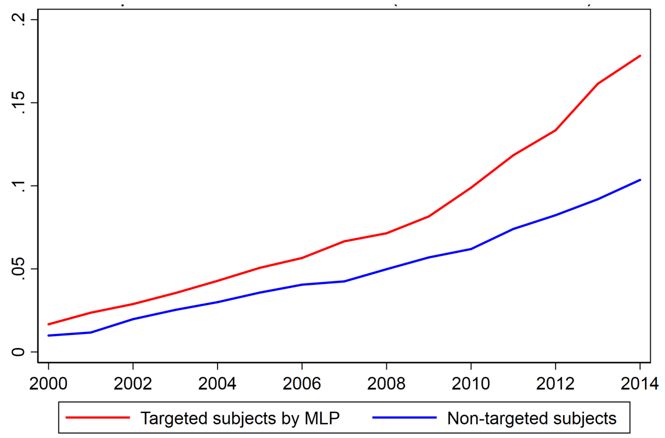
A good measure to capture China's distance to the scientific frontier in the respective discipline is to consider the share of top 1% cited articles that stems from Chinese scientists. On the basis of 47 Million Web of Science articles for all countries worldwide and their forward citations, I identify for each scientific discipline and each year the top 1% cited articles worldwide and determine the share that is attributable to China. Figure 2.4 shows the share of top 1% articles from China for subjects targeted by the MLP and for non-targeted subjects. Figure B-4 in the Appendix shows how the relative strength of Chinese science varies across scientific fields and changes over time.

Table 2.3: Triple Differences: Sample Split Analysis

Sample Split by	Turnover Abroad		Patent Stock		Years of Patenting		Prior Collab w/ CN	
	<p50	>p50	<p50	>p50	<p50	>p50	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
China=1 × Post=1 × MLP=1	-0.0274	-0.0011	-0.0214	-0.0136	-0.0078	-0.0028	-0.0375	0.2518
	[-0.091, 0.036]	[-0.134, 0.132]	[-0.048, 0.006]	[-0.170, 0.143]	[-0.045, 0.029]	[-0.171, 0.165]	[-0.108, 0.033]	[-2.174, 2.677]
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
China x Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MLP x Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,475	8,142	7,379	8,238	7,414	8,203	15,064	553
Unique Firms	480	479	481	478	473	486	928	31
Adj. R-squared	0.79	0.62	0.14	0.69	0.23	0.70	0.41	0.76

Source: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, Microdatabase Direct Investment (MiDi), 1999-2016, Web of Science, PATSTAT, own calculations. *Notes:* ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the firm level. Confidence intervals at the 95% level are reported in square brackets. This table presents sample split regressions using the triple differences estimation strategy in the baseline. In all models, the dependent variable it the log number of patents invented in China. Models (1) and (2) split the sample of multinational firms at the median of their turnover generated abroad. Models (3) and (4) estimate the regression separately for below and above median patent-intensive firms, as measured by their patent stock as of 2006 (note that all firms in the sample undertake patenting). Models (5) and (6) split the firms at the median years of patenting experience, measured by the years since first patent filing. Models (7) and (8) estimate the regression separately for firms that have not co-invented with Chinese inventors prior to 2006 and firms that have prior collaboration experience with the Chinese.

Figure 2.4: Top 1% Cited Articles from China (in % of Worldwide)



Source: Web of Science, own calculations.

I modify the definition of the MLP exposure variable by including an additional term $top1\%science_d$ in equation (2.3), which captures China's relative strength in each discipline d. First, I weight each MLP targeted scientific discipline by its relative strength in the pre-treatment period. I use the pre-period (1999-2005) average of China's share in top 1% articles to capture the pre-treatment quality level of Chinese science. In a second alternative, I calculate the geometric average growth rate of top 1% articles for each discipline in China in the post-period (2006-2016). I use the average annual growth rate to weight each MLP targeted scientific discipline by its relative improvement. Figure B-3 in the Appendix shows the distribution for the variants of the MLP exposure measure.

$$MLPexp_i = \sum_{j}^{J=34} \sum_{d}^{D=251} k_j^{2005} \times c_{jd}^{2005} \times I_d^{MLP} \times top1\%science_d \tag{2.3}$$

Table 2.4 shows the results using these two alternative MLP exposure measures that account for the quality and growth of the disciplines. Again, the triple interaction term is negative and not significantly different from zero for both dependent variables of inventions in China and citations to Chinese articles. The effects in columns (1) and (3) are very close to the baseline in Table 2.1. The effects in columns (2) and (4) are smaller in absolute terms, i.e. less negative, yet supporting the baseline results.

Alternative Dependent Variables. A patent with Chinese involvement can stem from a co-invention between German and Chinese inventors, a fully Chinese inventor team, or an international team of inventors from China and elsewhere. Columns (1) – (3) in Table 2.5 use the log of these three types of Chinese invention participation as dependent variable. In all cases, there is an insignificant negative effect, largest for German-Chinese co-inventions and smallest for fully Chinese inventions. Column (4) assesses whether there is a shift in

Table 2.4: Accounting for Quality and Growth of MLP Targeted Disciplines

	Inventions in CN		Citations to CN articles	
	(1)	(2)	(3)	(4)
MLP (quality)=1 × China=1 × Post=1	-0.0233		-0.0636	
	[-0.119,0.073]		[-0.141,0.014]	
MLP (growth)=1 × China=1 × Post=1		-0.0074		-0.0336
		[-0.104,0.089]		[-0.111,0.044]
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
China x Year FE	Yes	Yes	Yes	Yes
MLP x Year FE	Yes	Yes	Yes	Yes
Observations	15,617	15,617	15,617	15,617
Unique Firms	959	959	959	959
Adj. R-squared	0.68	0.68	0.69	0.68

Source: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, Microdatabase Direct Investment (MiDi), 1999-2016, Web of Science, PATSTAT, own calculations.
Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the firm level. Confidence intervals at the 95% level are reported in square brackets. This table presents results from a linear triple differences estimation that uses alternative 'MLP' measures, which account for the pre-period quality (columns 1 and 3) and the actual post-period growth of the MLP-targeted scientific fields (columns 2 and 4).

Table 2.5: Alternative Dependent Variables

	DE-CN co-invention (1)	fully CN invention (2)	CN-other co-invention (3)	% PAT invented in CN (4)	% CIT to CN articles (5)
MLP=1 × China=1 × Post=1	-0.0217	-0.0009	-0.0053	-0.0117	-0.0095
	[-0.080,0.036]	[-0.086,0.085]	[-0.095,0.084]	[-0.027,0.003]	[-0.068,0.049]
Firm FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
China x Year FE	Yes	Yes	Yes	Yes	Yes
MLP x Year FE	Yes	Yes	Yes	Yes	Yes
Observations	15,617	15,617	15,617	9,466	2,213
Unique Firms	959	959	959	804	214
Adj. R-squared	0.60	0.66	0.67	0.34	0.11

Source: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, Microdatabase Direct Investment (MiDi), 1999-2016, Web of Science, PATSTAT, own calculations.
Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the firm level. Confidence intervals at the 95% level are reported in square brackets. This table presents results from a linear triple differences estimation. The dependent variables in columns (1) – (3) are the log number of German-Chinese co-invented patents, of fully Chinese invented patents, and of co-inventions between Chinese and other foreign inventors, respectively. The dependent variable in column (4) is the share of patents that are invented in China as % of total patents by the MNE. Column (5) uses the share of non-patent references to Chinese articles as % of total number of non-patent references made to scientific publications by the MNE.

the relative importance of Chinese inventions among the patents of the multinational firm, following China's MLP policy. The dependent variable is the number of patents (fully or partly) invented in China as a share of total patents. Similarly, column (5) explores if there is a shift towards Chinese scientific articles among all NPL articles that the MNE references in its patents. I find no evidence for this.

Simple Difference-in-Differences Results. As a sense-check, I also run a difference-in-differences estimation among the subsample of firms that were already in China in 2006, with 'MLP' and 'Post' defined as in the baseline. I compare within-firm changes in inventions and knowledge sourcing of MNEs with high exposure to MLP targeted subjects, relative to MNEs with low MLP exposure. That is, for a moment, I disregard the difference in the differential impact compared to MNEs outside China (i.e. the third difference in the DDD approach).

$$lny_{it} = \beta \left[MLP_i^{High} \times Post2006 \right] + \alpha_i + \lambda_t + \epsilon_{it} \tag{2.4}$$

The coefficient β on the interaction term is the standard difference-in-differences estimator, α_i and λ_t are firm fixed-effects and year fixed-effects. The results in Table B-4 and Figure B-5 in the Appendix suggest that the MLP does not have a significant differential effect on the MNEs in China with a technology profile closely corresponding to the MLP targeted scientific fields, relative to other MNEs in China with low MLP exposure.

The substantial differences in size and inventive activities between multinational firms with subsidiaries in China and those without (see Section 2.3.2) render an equivalent difference-in-differences approach using 'China' as treatment variable inappropriate. As an alternative check, I review the dependent variables separately for the MNEs with and without FDI in China in Figure B-6 in the Appendix. The comparison shows that MNEs with FDI in China both have consistently more patents that originate from Chinese inventors and rely more on Chinese scientific knowledge. This is a simple comparison of the raw data, yet it shows that the multinational firms in the sample do tap into local knowledge by setting up subsidiaries in China. The results of this study are not at odds with the literature that MNEs use FDI for technology sourcing purposes (e.g., Griffith et al., 2006; Branstetter, 2006).

2.6 Conclusion

This paper used China's major science and technology policy in 2006 to causally identify the relationship between China's rise in science and technology and German multinationals' innovation and knowledge sourcing behavior. The results presented in this paper suggest that the MLP policy did not have a significant effect on the multinationals in China. Using a triple differences estimation approach, I find no evidence for MNEs that have a technology-science profile closely related to the MLP policy to offshore more of their R&D activities to China, relative to their low MLP exposure peers. I also find no evidence that they cite Chinese science relatively more often. Compared to MNEs without subsidiaries in China, I observe that MNEs with FDI in China do have a stronger growth in Chinese inventions and citations to Chinese science. That is, China's science and technology boom can generate spillover benefits for foreign firms that are in the country. However, among the MNEs with FDI in China, there seems no causal link between the MLP targeted scientific fields and the growth in innovation and knowledge sourcing from China.

The results may be surprising given that the MLP was at the heart of China's new science and technology strategy and related to the country's overall plan to renew the economy and ensure continuous growth in the long-run. It is beyond the scope of this paper to pin down the exact reasons why the MLP does not seem to have a significant effect on the MNEs in China. However, a few accounts by other scholars studying the Chinese innovation system give a cue about possible reasons and directions for further research.

Liu et al. (2017) discuss China's "top-down and plan-driven" approach to the implementation of the MLP, where the manifold R&D programs are often separately managed by different agencies without much coordination or national standard for quality control. Programs typically focus on easily quantifiable results, with the number of publications in international journals being the main indicator for the evaluation of universities, scientific programmes and individual scientists. This leads Chinese academics to excessively focus on the number of publications, in times of rapid expansion of government-funded research with yet underdeveloped evaluation processes (Fu et al., 2013).[32] At the same time, the incentive structure for researchers does not necessarily encourage collaboration among scientists or research-industry partnerships: credentials are typically only given to the first author on a paper, promotions require fast output that lead to short-term focus, furthermore the need to secure additional funding through industry collaborations is reduced through extensive government research funding (Fu et al., 2013).

Another aspect that may explain the absence of growth in co-inventions relates to China's government procurement and its emphasis on "indigenous innovation". Springut et al. (2011) raise the concern that China's government procurement rules give preference to domestic "indigenously innovated" products, which would even put products developed in China, but by a foreign owned enterprise, at a disadvantage. Interviews with R&D personnel of multinational firms in China conducted by Branstetter et al. (2014) underscore this narrative, highlighting the concern that Chinese government policy is likely to become less welcoming to international co-invention. Lastly, Chinese government's ability to identify target industries and technology as well as research institutions as recipients of funds (i.e. a "picking-the-winners" strategy) may simply have limitations.

[32] The robustness analysis in Section 2.5.2 aims to account for the quality of the scientific articles.

There are a number of avenues for further research. Using the available data, casting a closer look at the most patent intensive firms could provide interesting insights. The results from the heterogeneity analysis suggest that firms with larger patent stocks and prior co-inventions with China are more likely to benefit from the MLP induced boom in scientific output. In an extension of this paper, I would focus on these patent intensive firms and perform a within firm analysis, examining if firms shift their R&D portfolio towards more MLP affected technology classes. Similarly, a more in-depth study of industries like telecommunication or high-speed trains, where China managed to become internationally competitive, may provide enriching insights. Furthermore, one could investigate the complementarity or substitutability between the patents a MNE invents in China versus patents that are invented at the German headquarter. From the interviews by Branstetter et al. (2014) we know that "re-engineering products for the Chinese marke" plays a non-negligible role in the number of co-inventions. In general, more detailed data on Chinese-German collaborations and inventor networks would be useful to better understand the type of interaction between R&D personnel. This is left for future research.

3

Openness as Platform Strategy

Evidence from a Quasi-Experiment in Crowdfunding

3.1 Introduction

There is large consent on the existence of a funding gap for R&D and innovation, especially for young firms and startups (Hall and Lerner, 2010). Due to the inherent uncertainty over the risk and return of their planned ventures, entrepreneurs face considerable difficulties in attracting external finance (Cassar, 2004; Cosh et al., 2009). Recently emerged microfinance systems, in particular crowdfunding, are seen to narrow this gap and contribute to early stage business development, while showing less geographical and demographical bias than established types of financing (Agrawal et al., 2014; Belleflamme and Schwienbacher, 2014; Mollick, 2013).

The success story of Internet-based crowdfunding platforms, which helped capital seekers raise more than US$ 6 billion in 2019[1] from individual funders, is intriguing given that entrepreneurial finance markets usually rely on specialized professionals. These professionals, such as business angels and venture capitalists, stand out by their expertise and accumulated experience to evaluate investments into business ideas despite the problems arising from asymmetric information and moral hazard (Lerner, 2002; Ueda, 2004). In contrast, individual funders on crowdfunding platforms make their investment decisions without the involvement of experts, even though they also face considerable information asymmetry concerning the quality of projects and the capabilities of project creators (Ahlers et al., 2015; Belleflamme and Schwienbacher, 2014; Courtney et al., 2017). Indeed, crowdfunding is characterized by a high default rate among financed projects (Mollick and Kuppuswamy, 2014).

To prevent market failure, crowdfunding platforms may hence establish screening processes for project applications with the goal to maximize the chances that funded projects will be of high quality and thus reward their funders (Belleflamme et al., 2015). While the platform literature has identified certain advantages of access control to increase platform value (Boudreau and Hagiu, 2009; Hagiu, 2011; Evans, 2012), preventing entrepreneurs from offering their projects on the platform stands at odds with crowdfunding's core objective to "democratize access to capital" (Sorenson et al., 2016). This study provides empirical insights into the inherent quantity vs. quality trade-off crowdfunding platforms face when deciding on their degree of openness.

The crowdfunding literature has focused primarily on quality signals, provision points and crowd dynamics as the main mechanisms to prevent market failure (e.g., Agrawal et al., 2015; Burtch et al., 2015; Courtney et al., 2017; Kuppuswamy and Bayus, 2017; Zhang and Liu, 2012). Much less attention has been paid to the role of platform operators themselves, even though they usually take a more active role than simply providing the infrastructure. A platform's decision to open up or to restrict access to its marketplace determines the number and type of agents participating on both sides of the market. Building on the premise of positive

This chapter is based on joint work with Fabian Gaessler.

[1]For more information on the global crowdfunding market, see e.g., https://www.statista.com/outlook/335/100/crowdfunding/worldwide#market-revenue, last accessed on March 23, 2021.

network effects, scholars typically assume that users place a higher value on platforms with a larger user base on either side (Cennamo and Santalo, 2013), and that market thickness can be a source of competitive advantage (Armstrong, 2006; Eisenmann, 2006; Eisenmann et al., 2011; Boudreau, 2012). Crowdfunding platforms, which act as intermediaries between project creators and funders, generate revenue through maximizing the number of successful transactions. An increase in product quantity and variety should promote successful matches between project creators and (potential) funders – which directly translate into higher platform performance.

However, openness may come at the cost of unrestricted entry of low quality projects. If funders fail to distinguish high quality from low quality, this may erode the demand side's trust in the offered projects. As a consequence, (potential) funders and high quality fund seekers alike may turn their back to the platform, leading to a downward spiral of quality. While preventing "lemon" projects from entering the platform could reduce market failure, a thorough due diligence of projects is costly due to the low scalability of manual evaluations and the limits of automatic screenings. Furthermore, it is ex-ante unclear whether access control also leads to situations where the platforms screen out the "wrong" (i.e. high quality) projects or deter them from applying in the first place. This may have potential implications not only for the quantity but also for the composition of entrepreneurial ventures offered on the platform in terms of other dimensions, such as variety and novelty. Crowdfunding platforms show varying degrees of quality and fraud controls (Cumming and Zhang, 2018). However, a causal relationship between platform openness, crowdfunding success and investment performance has yet to be identified.

We empirically investigate the effects of platform openness in the reward-based crowdfunding market. Reward-based crowdfunding platforms link fundraising project creators to individual backers, who are promised a tangible but non-financial reward for their investment, typically the product resulting out of the project (Belleflamme et al., 2015). The reward-based crowdfunding market is dominated by two platforms, Kickstarter (KS) and Indiegogo (IGG), which will be the primary and secondary subjects of our empirical analysis.

We exploit a strategic decision at KS, to switch from access control to de facto openness, as a quasi-experiment to disentangle the diverse effects of platform openness on the two sides of the market. Crowdfunding platforms set up rules and procedures to govern the economic activities on their platform. Since the platform's launch in 2009, 16 million funders, so-called "backers", pledged about US$ 4.5 billion, successfully financing more than 160,000 projects.[2] At KS, each project had to undergo a manual review to obtain approval before listing on the platform. The pre-screening shall ensure that projects abide to the platform rules and limit the number of poor quality and non-eligible projects. This stood in contrast to the open strategy of its main

[2]Current statistics on Kickstarter can be found at https://www.kickstarter.com/help/stats, last accessed on March 23, 2021.

competitor, IGG. In June 2014, however, KS decided to abandon its access control and adopted an open platform policy, enabling projects to launch without prior approval (Kickstarter, 2014; Wessel et al., 2017).

We use the synthetic control method (Abadie et al., 2010) to construct an adequate control group for KS, which is based on the weighted average of IGG regions and which accounts for the different geographical coverage of the two platforms as well as dynamic trends. We combine this with a difference-in-differences approach to disentangle the causal effects of platform openness on KS's market thickness and successful matches. We assess the effects of openness on quality using novel text analysis methods and evaluate reward delivery and demand-side feedback of projects that reached the funding stage.

First, we find that the policy change had indeed an instant effect on market thickness; both quantity and variety of projects increased. Notably, we observe an increase in KS's market share relative to IGG in those countries where both compete with each other. Second, we find that the number of market matches, i.e. projects that got financed, increased in absolute but not in relative terms – the funding success rate of projects dropped by almost a third. Third, we compare the quality of funded projects launched in a narrow time window around the policy change and find that those projects without pre-screening are on average of lower quality. They more frequently fail to deliver their rewards and garner more complaints from their backers (an increase by around 18% and 13%, respectively).

Our overall results suggest that the change to openness led in fact to more market matches and higher revenues for KS in the short-run. However, in the medium to long-run, it is unclear whether KS's decision ultimately generated value for the platform ecosystem. Increased competition on the supply side and a lower funding success rate decrease the value for project creators. Similarly, backers face a higher rate of project default, which may dampen their platform experience. These findings suggest that there are limits to the "wisdom of the crowd" in screening out low quality projects on their own. Effective platform control hence facilitates the successful matching of project creators and backers and helps reduce information-related market failure in the financing of entrepreneurial ventures.

We contribute to the existing literature in the following ways. First and foremost, we add to the recent stream of crowdfunding studies. These studies predominantly focus on the characteristics of projects and funding dynamics as determinants for fundraising success and disregard the role of platform rules and procedures within the crowdfunding ecosystem (Dushnitsky and Fitza, 2018). Our results help understand how platform design affects crowdfunding decisions, and whether regulation determines funding success. Furthermore, instead of analyzing the phenomenon in an isolated fashion, we use panel data on two crowdfunding platforms and establish that platform design affects not only the focal platform's ecosystem, but also the competing platform's project supply. This highlights the need to understand the crowdfunding space as a whole and to consider the interdependencies between active platforms.

Second, our study relates to the overall literature on entrepreneurial finance, in particular the importance of due diligence for investors, and the survival of entrepreneurial firms (Amit et al., 1998). If platform design determines which projects enter the platform and reach their crowdfunding goal, the frequently cited “signalling value” of funding success for subsequent financing rounds becomes platform-specific (Dushnitsky and Zunino, 2019).

Finally, this paper provides insights into the strategic value of quality as a competitive advantage for platforms. So far, we know little about the conditions that enhance or mitigate the competitive advantage of platform quality (McIntyre and Srinivasan, 2017). Whereas the relationship between quality and platform performance has been studied inter alia in the videogame industry (Boudreau and Hagiu, 2009; Zhu and Iansiti, 2012) and Facebook applications (Claussen et al., 2013), crowdfunding platforms – and intermediaries in general – have largely escaped a thorough empirical examination.

3.2 Theoretical Framework

Platform markets are characterized by the coordinated interaction of two distinct groups and the presence of indirect network effects that each group exerts on the other (Rochet and Tirole, 2003; Rysman, 2009). Platforms create value by providing the infrastructure and a set of rules to facilitate the interactions between participants from both sides. The platform may benefit either indirectly from coordinating the supply of complementary goods that increase the value of their own good (e.g. technology platforms), or directly from intermediating between supply and demand side for a transaction-based fee (e.g. crowdfunding platforms). In the crowdfunding context, the platform helps agents on one side of the market (project creators) to get in contact with and to receive funding from the other side of the market (project backers).

3.2.1 Market Thickness and Demand Heterogeneity

Particularly for exchange platforms, a large number of participants on both sides is paramount, as it directly translates into market thickness. Thicker markets increase the probability that participants will find a “match” and complete a transaction. Where preferences and products are heterogeneous, consumers benefit from a large number of suppliers offering distinct products, just as suppliers benefit from a large demand side (Halaburda et al., 2017). The more suppliers are active on the platform the more diverse the goods typically are (Boudreau, 2012). Yet, the value distribution of these products can be highly skewed, with very few superstar products attracting most consumers (Bresnahan et al., 2015). If platforms have only incomplete information on the consumers’ preferences and/or the products’ value, they may decide to offer a broad range of products and rely on the demand side to narrow it down (Einav et al., 2016).

If a crowdfunding platform chooses openness, few (if any) project ideas are screened out. We therefore assume that the lack of screening leads to more projects on the platform and consequently an increase in market thickness on the supply side. Furthermore, crowdfunding is characterized by a large product space in which crowdfunding platforms seek to offer a large variety of ventures (Belleflamme et al., 2015). This considerable product heterogeneity is a key difference to many other two-sided markets and links market thickness to product variety. We therefore propose:

Hypothesis 1 (H1): Market thickness: An open crowdfunding platform is associated with a larger quantity and variety of projects.

Another supply-side aspect concerns the ambiguous effect of openness on the novelty of crowdfunding projects. On the one hand, openness may attract – and even promote the creation of – highly novel projects that would otherwise be deemed too risky or speculative to benefit the platform (cf. Hagiu and Wright, 2020). On the other hand, openness may also attract copycats and generic project ideas. As it is hard to assume which of these two mechanisms prevails, we refrain from proposing a directional hypothesis.

3.2.2 Market Matches and Platform Congestion

Despite the positive effect of platform openness on market thickness, a very large number of participants, in particular on the supply side, may congest the platform and lower the chance for successful market matches. First and foremost, it may increase search costs for the demand side. Even if consumers prefer a diverse set of products, finding the optimal one may become difficult in light of an overly large choice set (Arnosti et al., 2019; Li and Netessine, 2020). If consumers also benefit from direct network effects, where the value of a product increases with coordinated consumption, too many products and hence dispersed consumption may have a negative effect on overall platform value (Casadesus-Masanell and Halaburda, 2014; Halaburda et al., 2017). To address this challenge, platforms largely employ digital technology to ensure scalability without loss of efficiency (Goldfarb and Tucker, 2019).[3]

Second, not only consumers but also suppliers may experience negative intra-group externalities if there are too many peers on the platform (Belleflamme and Toulemonde, 2009). Due to increased competition suppliers face a lower likelihood of interacting and transacting with the demand side. The competition effect is considered most salient if the offered products are close substitutes and consumer coordination is important. In case of high product heterogeneity, suppliers are more likely to find their niche.

We expect that thicker crowdfunding markets are associated with a higher number of market

[3] Particular platform design choices, such as advanced search functions, rankings and algorithm-based recommendations for "trending products" can be implemented to mitigate search frictions and to ease user coordination (Fradkin, 2017).

matches. Individuals on the demand side, the backers, display a high degree of taste heterogeneity. Hence, a larger number of projects on reward-based crowdfunding platforms tends to increase the probability of finding a project that better fits the taste of a backer (Belleflamme et al., 2015). Certainly, greater market thickness increases the search costs for backers. However, the typical online platform design choices prevent congestion problems by employing, for instance, matching algorithms, where users can search through particular categories and subcategories, rely on rankings etc. This facilitates a good fit between backers' tastes and project characteristics (Belleflamme et al., 2015). We therefore propose that openness leads to an increase of market transactions – i.e., successfully funded projects – on the crowdfunding platform.

Hypothesis 2a (H2a): Market matches: An open crowdfunding platform is associated with a larger number of funded projects in absolute terms.

One particular aspect of crowdfunding is the strong need for backer coordination. Crowdfunding projects seek small contributions from a large group of individuals. If a crowdfunding project fails to raise sufficient money, it will likely not be able to strive and deliver the promised rewards. Hence, the contribution of one individual, a so-called pledge to the crowdfunding project, will only lead to a successful transaction if others contribute as well. In fact, many crowdfunding platforms only complete a transaction once a certain threshold, the funding goal, has been reached (Agrawal et al., 2014). With an increasing number of projects listed on the platform, the need for coordination increases as well as the complexity thereof (Wash and Solomon, 2014). Considering heterogeneous backer preferences and product characteristics, and assuming imperfect coordination among backers, we expect that market thickness leads to an absolute but not a relative increase of market transactions.

Hypothesis 2b (H2b): Market matches: An open crowdfunding platform is associated with a smaller number of funded projects in relative terms.

3.2.3 Project Quality

Lastly and arguably most importantly, platform openness may lead to the entrance of undesirable participants onto the platform.[4] A potential decline in average quality is arguably the most important trade-off to an increase in quantity. If consumers are unlikely to gain complete information about the supply side's offer, the door stands wide open for opportunistic behavior. Suppliers may misrepresent the product's true quality or even conduct pure fraud.[5] This is in line with the notion that the direction of network effects depends on the level of opportunistic behavior, reputation and perception of trust (Afuah, 2013). Asymmetric information

[4]As argued in Hagiu (2011), quality issues in two-sided markets typically originate from the supply side.
[5]We hence consider a relatively broad notion of "quality" of products.

and the resulting uncertainty may lead to market failure (Boudreau and Hagiu, 2009; Evans, 2012). The extent of this "lemons problem" depends on how much the demand side suffers from information asymmetry. If consumers can identify low quality products, the platform's value should remain unscathed for consumers and high-quality suppliers.

In the crowdfunding context, potential funders face unique information challenges. First, there is considerable asymmetric information concerning the quality of projects and the capabilities of project creators (Ahlers et al., 2015; Belleflamme and Schwienbacher, 2014; Courtney et al., 2017). A crowdfunding project's quality primarily depends on the technical feasibility and commercial viability of the product offered. As these products are usually at a conceptual or development stage, crowdfunding projects – as technology entrepreneurship in general – are highly uncertain. Failed or delayed delivery of the promised reward is a common event in crowdfunding (Mollick and Kuppuswamy, 2014). Second, crowdfunding gives room to moral hazard arising from the limited (or non-existent) liability towards the funders. Especially in reward-based crowdfunding, with nearly perfect separation of investment and project execution, project backers have few means to monitor the actions of creators and to detect fraud.

There are several mechanisms through which crowdfunding platforms can reduce information asymmetries (Agrawal et al., 2014; Belleflamme et al., 2015). First, they may put in place monitoring systems ensuring that the project creators follow basic liability rules. This prevents or at least mitigates fraud. Second, the platform may set up an infrastructure that maximizes transparency and coordination. This increases the effect of reputation signals concerning the identity and track record of project creators and facilitates coordination among (potential) investors. Nonetheless, an open crowdfunding platform shifts the responsibility and effort to separate low from high quality projects to the demand side. However, it seems questionable that quality assessment by a group of atomistic and typically non-professional backers can fully compensate for centralized platform screening – given the limited project information, ineffective coordination and high costs relative to each funder's marginal monetary contribution. If backers can screen out low quality projects only imperfectly, an overall decrease in the average quality of listed projects cascades to a lower average quality among successfully funded projects.[6] In contrast, denying low quality projects access to the platform curtails the left tail of the quality distribution and increases the average project quality (Cumming and Zhang, 2018). Furthermore, this mitigates information asymmetries as project creators have the incentive to invest in detailed and sound project descriptions in order to pass the screening process. Consequently, we propose the following relationship between platform openness and project quality:

[6] There are convincing arguments why funders are subject to constraint investment behavior. On individual level, search and evaluation costs quickly become excessively high relative to the size of the investment. Second, funders lack experience and training as investors. In fact, there is evidence that funders predominantly follow the behavior of peers as can be inferred from intra-campaign timing of pledges (Burtch et al., 2015; Kuppuswamy and Bayus, 2017) and herding behavior around reputable investors (Kim and Viswanathan, 2018).

Hypothesis 3 (H3): Project quality: An open crowdfunding platform is associated with a lower average quality of funded projects.

In summary, we argue that the relationship between openness and performance is ambiguous in the crowdfunding context. First, openness is expected to increase project quantity and variety, whereas the effect on novelty remains ex ante ambiguous. Second, a thicker supply-side should result in more market matches in absolute terms, granted that search frictions on the demand side are mitigated by platform design. However, it is unclear whether the relative frequency of market matches is higher as well. Project creators may suffer from stronger competition effects such that only a smaller percentage of projects can be matched with enough interested funders. Third, platform openness should lead to a decline in average quality on the supply side, as projects that would have previously been screened out can now enter the platform. If funders on the demand side cannot fully compensate for the lack of pre-screening, this should be reflected in lower consumer experience. Hence, even if the platform benefits from a thicker market and potentially more market matches in the short-run, the long-run effects on platform value are unclear and depend on the degree of project competition and the extent of market failure due to lower quality. We empirically investigate these different mechanisms in the following.

3.3 Empirical Setting: Reward-Based Crowdfunding

3.3.1 Kickstarter and Indiegogo

We study the two largest reward-based crowdfunding platforms, Kickstarter (KS) and Indiegogo (IGG), to examine the effect of openness on platform performance. KS and IGG are both reward-based crowdfunding platforms that link fundraisers with potential backers to solicit funds for a project or idea. Instead of a monetary return, backers are offered some form of non-monetary reward for their investment, typically the good the entrepreneur intends to develop (cf. Belleflamme et al., 2015).[7]

IGG was launched in January 2008. KS entered the crowdfunding market later in April 2009, but quickly developed its reputation to become larger than IGG. Although both platforms operate under the reward-based funding model and are considered the closest competitors to each other, they differ in their level of openness. IGG has historically been the more open platform with a hands-off strategy. From the beginning, it opened its platform for a wide array of projects and did not impose any specific limitations on the type of projects it would

[7]Crowdfunding platforms are typically defined by the existence and type of reward for the funders of a project (cf. Belleflamme et al., 2015). Equity-based crowdfunding comes with a stake in the venture in return for a monetary investment and hence resembles established equity markets. Reward-based crowdfunding, however, offers funders non-monetary returns. Funders usually receive the project's output (newly developed products or services) in return for a financial contribution. So, similar to equity-based crowdfunding, the funder's return on investment ultimately depends on the project's fate.

support. In contrast, KS maintained a narrower focus of its projects with a number of specific requirements and restrictions to define what qualifies as a KS project.[8]

It is central to our study that KS and IGG used to employ different strategies concerning the control over projects that wished to launch on their websites. While IGG did not screen the projects entering its platform and allowed them to launch instantaneously, KS conducted a rigorous review process: each project had to be manually screened by a KS member of staff before the project could go live. KS's human approval process was not to be understood as an endorsement by the platform based on taste or the likelihood of a project's funding. Rather, the vetting process checked for a project's feasibility and content standards, ensuring that each project adhered to the rules of the platform. This included whether the project fell under one of the eligible categories at KS, whether it was a discrete project with a beginning, middle and end as well as of a "creative" nature. KS used its approval process as a screening device for detecting poor quality and non-eligible projects. Therefore, KS was much more curated than IGG and was considered to be more difficult for a project to gain access to.

3.3.2 Policy Change at Kickstarter

On June 3, 2014 Kickstarter announced via a blog entry on its own website that it would make two major changes regarding its approval process. It would introduce the "Launch Now" feature that allows project creators to bypass the previously mandatory manual approval process by a KS member of staff and simplify the rules to open up the platform to new kinds of projects.[9]

> "Today we're excited to announce two important changes that make Kickstarter easier to use than before (...) We want creators to have the support and freedom they need when building their projects. That's why we're introducing a feature called Launch Now. It gives creators a simple choice: go ahead and launch your project whenever you're ready, or get feedback from one of our Community Managers first. (...) We're also introducing dramatically simplified rules for Kickstarter projects. (...) we boiled them down to three basic principles."[10]

Following the change, instead of employing a human approval process, there is now a simply algorithmic check. The algorithm looks at keywords in the project description, the creator's track record on KS and checks if rewards, funding goal and other major information fields are filled out on the project website. If a project passes the algorithmic check, it can launch immediately.[11] KS relaxed its regulations and review process in an effort to open up and scale the platform. This policy change moved KS from an access controlled to a de facto open

[8]This included the outright ban on cosmetics, eyewear, and health, medical and safety products.

[9]Kickstarter website, blog entry "Introducing Launch Now and Simplified Rules" from June 3, 2014, at https://www.kickstarter.com/blog/introducing-launch-now-and-simplified-rules-0, last accessed on March 23, 2021.

[10]The revised guidelines now comprise only the following three basic principles: Projects have to (1) create something to share with others, (2) be honest and clearly presented and (3) cannot fundraise for charity, offer financial incentives, or involve prohibited items.

[11]For more details on the screening process before and after the policy change, see Wessel et al. (2017).

platform. This move was thought to help KS compete against its biggest rival IGG, whose "open-door" policy had proven to be a successful formula for the platform. Kickstarter had lost lucrative projects to IGG due to its strict policies in the past.[12]

Besides size and profitability of a project, crowdfunding platforms benefit from projects that can generate wide media attention, expanding the existing network of backers and increasing the general popularity of the platform (Agrawal et al., 2014). KS CEO Yancey Strickler said that "for a brand or a community to have definition, there have to be rules", but now that KS has entered the "mature era" where backers understand the risks and rewards of KS projects, it looks to open up the platform.[13] The policy change raised concerns among members of the crowdfunding community, though, that it could lead to a decline in quality, damaging the brand and marketplace of KS.[14]

3.4 Empirical Strategy

We use Kickstarter's decision to abandon its manual upfront screening as a quasi-experiment to assess how openness impacts platform performance. In particular, we study the effect of this policy change on market thickness, market matches, and ultimately project quality.

We conduct our analysis in two steps. First, we estimate how the change in the platform strategy affected the number and composition of projects launching on the platform and projects funded at the overall platform level. We hereby draw on a control group that should capture time-variant effects that may otherwise confound the estimated effect. Second, we investigate how the change affected project quality. We hereby compare project level quality measures of funded projects that launched on the platform right before and after the policy change. Focusing on a narrow time window around the policy change, we analyse if there is an immediate and discontinuous effect on project quality following the abandonment of the previously mandatory manual screening by Kickstarter. We use project delivery and backer satisfaction as performance measures to assess the quality of a project.

3.4.1 Synthetic Control Method for Control Group Construction

To evaluate the impact of the policy change on platform-level outcome variables, we need to identify a control group for comparison. Indiegogo, KS's sole competitor of similar size, represents a natural candidate for a control group. However, a direct comparison of KS and IGG

[12] See e.g., article on WIRED from November 12, 2012, at https://www.wired.com/2012/12/kickstarter-rejects/, last accessed on March 23, 2021.

[13] Article on The Verge from June 3, 2014, at http://www.theverge.com/2014/6/3/5775548/kickstarter-s-next-campaign , last accessed on March 23, 2021.

[14] See, for instance, Liam Collins from NESTA: "*Kickstarter (will) start picking up the sort of projects it rejected in the past. (...) But I'd expect this to lead to an increased failed-project rate and potentially more issues with late delivery of rewards,*" as quoted by BBC on June 4, 2014, at http://www.bbc.com/news/technology-27699267, last accessed on March 23, 2021.

may be a naïve one. As both crowdfunding platforms compete for projects in several countries, we are concerned the policy change at KS may interfere with IGG. Under the reasonable assumption that project creators may conduct platform shopping, a change at KS could also affect the number of projects at IGG, limiting its appropriateness as control group. To address possible market interaction and competition effects, we make use of the fact that KS and IGG differ in geographical coverage.

IGG allows project creators worldwide to launch a project on their site and is currently available to 235 countries and regions.[15] By contrast, KS is only available to fundraisers residing in certain countries, although individuals from around the world are allowed to contribute funds to the projects on KS. Initially, KS was available only to project creators in the US, but expanded its activities over the course of 2013 to include the UK, Canada, Australia and New Zealand.[16] We can therefore distinguish between IGG projects that are prone to substitution effects and IGG projects that do not experience any indirect treatment, since they are located in countries non-eligible to KS.[17] To the best of our knowledge, we are the first to exploit the differences in geographical coverage of KS and IGG for the empirical strategy.

KS's set of eligible countries is a perfect subset of IGG's; we speak of those projects not eligible at KS as the "IGG exclusive" control group. Figure C-1 in the Appendix maps the limited number of countries where KS was available to project creators at the time of the policy change. Projects in countries that were eligible to both platforms (US, Canada, UK, New Zealand and Australia) are termed "IGG inclusive". These IGG projects are excluded from the control group, so to avoid double counting of potential substitution effects between the two platforms and hence overestimation of the impact of the policy change. However, IGG exclusive countries as a whole may still be an inadequate control group to KS, given potential differences in time-trends, popularity and institutional set-up between the regions where crowdfunding activity takes place.

To address this concern, we employ the synthetic control method (Abadie and Gardeazabal, 2003; Abadie et al., 2010) to estimate the causal effects of the policy change at KS. The method builds on the difference-in-differences approach but constructs an arguably better comparison group for the estimation. The basic idea of the synthetic control method is that a linear combination of potential control units often provides a much better comparison than any single control unit alone, especially when there are only few potential control units available and the parallel trend assumption of the difference-in-differences approach may not hold. The synthetic control method provides a systematic way to construct a "synthetic control" group as the weighted average of all potential control units (in our case, IGG exclusive countries) that

[15] Indiegogo website, at https://www.indiegogo.com/about/our-story, last accessed on March 23, 2021.

[16] This list is constantly expanding and now further includes the Netherlands, Denmark, Ireland, Norway, Sweden, Germany, France, Spain, Italy, Austria, Belgium, Switzerland, Luxembourg, Hong Kong, Singapore, Mexico and Japan. Kickstarter website, at https://www.kickstarter.com/help/faq/creator+questions, last accessed March 23, 2021.

[17] We use platform-specific criteria regarding the project creator's residence to identify the two groups.

best resembles the treated unit (KS countries) before treatment. The synthetic control method has the advantage that it does not extrapolate outside of the support of the data (Abadie et al., 2015) and that it allows for unobserved confounders whose effect may vary in time (Abadie et al., 2010).[18] By contrast, the difference-in-differences model accounts for unobserved confounders but restricts the fixed-effects to be constant in time.

In our context this means that we take a weighted average of IGG exclusive countries (more precisely, IGG exclusive regions) as a synthetic control group to evaluate how KS's policy change affected the activities on the crowdfunding platform. We consider all IGG projects that were launched in IGG exclusive countries[19] during the period of study for the so-called "donor pool" of potential control units. We assign them to one of the following eight geographical regions based on the project creator's country of residence: Africa, Asia, Central & Eastern Europe, Northern Europe, Southern Europe, Western Europe, South & Middle America and "Other".[20]

The post-treatment outcome for the synthetic control serves as counterfactual for Kickstarter in the absence of the policy change. We use the synthetic control group to calculate a DiD estimate, comparing the average difference between Kickstarter and its synthetic counterpart pre-treatment with the difference post-treatment. We combine the synthetic control method with the difference-in-differences approach here, in order to account for seasonal effects in the post-treatment period.

$$DD_{KS,SCM} = (Outcome^{KS}_{post} - Outcome^{SCM}_{post}) - (Outcome^{KS}_{pre} - Outcome^{SCM}_{pre}).$$

In a first step, we evaluate the effects the policy change has had on the number of projects launching on Kickstarter as set out above. By nature of the policy change, the effects of abandoning the due diligence are most immediately reflected in the number of projects (the demand side). The effects of the changes on the demand side may then trickle down to the backers and their funding behavior (the supply side), with implications for the overall funding success rate on the platform. For the evaluation of the impact of the policy change on pledging behavior and funding outcomes we employ weighted difference-in-differences estimation, where we use the weights generated by the synthetic control model in the first step.

For the construction of the synthetic control, we include all pre-intervention values of the outcome variable as predictors, save the last two weeks before intervention, which we use as a validation period. Abadie et al. (2010) point out that matching on pre-intervention outcome

[18]We refer to Abadie et al. (2010, 2015) for a detailed description of the method.

[19]That is, all countries worldwide except for the US, UK, Canada, Australia, New Zealand, and the Netherlands. The Netherlands were added to KS in April 2014, which falls into the time window of our study and would confound the estimated effects of the policy change. We hence exclude the Netherlands from our analysis.

[20]Countries need to be aggregated at a regional level since many countries have only very few projects per week, making matching infeasible. Countries that have equal or less than 10 projects over the period studied are assigned to 'Other'.

values helps accounting for unobserved factors and their heterogeneous effect on the outcome variable. We divide the pre-intervention period into a training period and a validation period, as suggested by Abadie et al. (2010), to minimize the out-of-sample prediction errors. In addition to lagged outcome values, we include the number of pledges and Google Index data in the list of predictor variables. Google Index data for the search words “Crowdfunding”, “Kickstarter” and “Indiegogo” serve as a proxy for the popularity of crowdfunding in general and the popularity of the two respective platforms. We match the predictor variables on the trends in the outcome variable of technology projects launched each week, rather than on the project level.[21]

3.4.2 Pre- versus Post-Policy Evaluation

To test the hypothesis that KS compromised on the quality of its projects by abandoning the manual upfront screening, we evaluate project quality in a quasi-experimental pre- versus post-policy design.[22] Conditional on getting funded, we compare the post-funding performance of KS projects that launched immediately before to projects that launched immediately after the policy change. The validity of the quasi-experimental comparison strongly depends on the assumption that project creators were not able to manipulate their assignment, i.e., whether their launch was before or after the policy change. Given the lack of any prior announcements or hints in the crowdfunding community, we see no evidence that (potential) project creators could anticipate the abandonment of due diligence and consider this aspect fulfilled. We restrict the sample of our analysis to a narrow time window around the policy change, to mitigate against time-variant confounding factors.

Merely looking at the funding of a project falls short in providing a sufficient assessment of the project quality. A project reaching its funding goal only implies that the project is deemed worth funding and that a match between fund seekers and fund providers occurred. However, tangible measures of quality – whether a project is executed and delivered as planned and if to the satisfaction of backers – only become apparent after the funding stage. We therefore use reward delivery and backer satisfaction as post-funding performance measures to assess the quality of a project. Our primary dependent variable to assess changes in quality is the propensity of reward delivery. As an alternative measure of project quality, we use backer complaints in the commentary section of a project’s webpage. Negative feedback to the project reflects instances where projects do not meet the expectations of backers.

Each project on KS has its own webpage, which stays online even after the funding period. It is the main forum for interaction between the project creator and backers: project creators

[21] Kickstarter is much larger in its size, as measured by the number of projects, than any of the Indiegogo regions as potential control units. As the synthetic control method by Abadie et al. (2010) restricts the weights to sum to one, it would neither be feasible nor sensible to match on the level of projects.

[22] In contrast to the first part of our analysis, where we look at platform level variables, we undertake the quality assessment at the individual project level. Hence, we compare KS projects pre- to post-policy change, rather than to IGG projects, for reasons of better comparability and data availability.

keep their supporters abreast via regular project updates, and crowdfunders can ask about the progress via comments. We collect information from updates, comments and general internet research to determine if a project delivered the promised reward to its backers. Moreover, we run a sentiment analysis on the rich data contained in the comments to identify complaints.[23]

3.5 Data & Descriptive Analysis

We use a comprehensive dataset of projects on KS, our focal platform of interest, and IGG, for the control group.[24] It includes information on all initiated crowdfunding projects at project-level and additional micro-level information, such as comments and updates, for the subset of successfully financed projects. We restrict our sample to projects at both platforms whose launch day falls within the period of 20 weeks before and 20 weeks after the policy change on 3 June 2014.[25] We further limit our sample to projects that are part of the "Technology" category. These projects are most likely to involve high technology ventures. Furthermore, backers are more likely to expect a tangible reward for their investment that is directly linked to our definition of project success.[26] In our empirical analysis, we focus on the following variables.

Projects. The number of projects in the category "Technology" launched in a given week; our primary measure of supply-side market thickness.

Pledge goal. The ex-ante determined amount of money a project creator seeks to raise. To standardize the pledge goal as well as other financial variables, we convert monetary values to Euro with the currency exchange rate on the day of project launch.

Money pledged. The amount of money pledged to a given project.

Funding. If a technology project reaches or surpasses its ex-ante set funding goal, we denote a project as funded.

Projects funded. The number of successfully funded "Technology" projects launched in a given week. This represents our measure of market matches.

Project novelty. We measure the novelty of a project by calculating its text-based similarity to the stock of all previous KS projects (up to 20 weeks before the policy change). We argue that the description of novel projects relies more likely on new or at least uncommon words

[23]We accomplish this with an automatic sentiment analysis tool: SentiStrength (cf. Thelwall et al., 2010). We extend the pre-defined dictionary with several key words and idioms specific to the crowdfunding community. The list can be accessed at: http://tinyurl.com/oapsefaqeic2018, last accessed on March 23, 2021.

[24]The dataset has been collected via a webscraping approach between late 2015 and 2016.

[25]In week 21 after the policy change, the platform opened up to several other countries (see Table C-1 in the Appendix).

[26]The simplified rules for KS projects also meant that previously banned products, such as sunglasses, bath and beauty products can now receive funding. In order to ensure that our analysis only captures the effects of abandoning access control, but not the enlarged set of eligible project types, we limit our analysis to projects of the "Technology" category. Table C-2 in the Appendix lists all press releases by KS during 2014; there were no other changes at the platform level that affected the "Technology" category.

instead of frequent word combinations.[27] We therefore process the descriptive project titles as text input. After deleting punctuation and stopwords, we calculate *Jaccard* similarity scores (with 0 depicting completely different – and 1 practically identical titles) for each project pair.[28] Novelty is then defined as 1 minus the focal project's average similarity score with the ten most similar prior projects.

Reward delivery. We distinguish funded projects by the delivery of rewards to the backers. As KS does not publicly label projects by their status of delivery, we manually identified projects where rewards were delivered to backers based on comments, updates and general internet research. Project creators may offer a variety of rewards ranging from public acknowledgment to final products, depending on the size of the contribution. As we are most interested in rewards that are linked to the project itself, we further distinguish between material and immaterial rewards.

Backer complaint. We quantify negative feedback of project backers from the comments section of the respective crowdfunding project page.[29] To ease comparison between projects, we analyze only comments within 12 months after the expected delivery date and identify complaints with a list of general and context specific key words that signal discontent with the status or the result of the project.

Platform and crowdfunding popularity. To account for time-varying popularity and general awareness of KS (respectively IGG) and crowdfunding as such, we draw on Google trends indices. A Google trends index represents a normalized weekly measure of the relative frequency a particular term was queried on Google's search engine.[30]

Time variables. To control for seasonal and platform specific time trends, we use in our panel regressions additional time-variant variables, such as platform age and month dummies.

We compare platform activities at KS before and after the policy change, using crowdfunding data from January 13, 2014 to October 19, 2014, which covers 20 weeks before and after the policy change on June 3, 2014. We set the data as panel data with crowdfunding activities aggregated at the week level over time and by platform. Figure 3.1 shows the absolute number of technology projects launched by week and by platform. While the number of new launches on KS (IGG) ranged around 50-70 (120-150) projects each week, that number experienced a discrete increase at KS following the removal of the access control. In contrast, the average number of projects launched each week on IGG shows no such development around the policy change. The overall positive trend for both platforms may be explained by the general rising

[27] Similar methods have been used, for instance, to measure patent similarity (Arts et al., 2018).

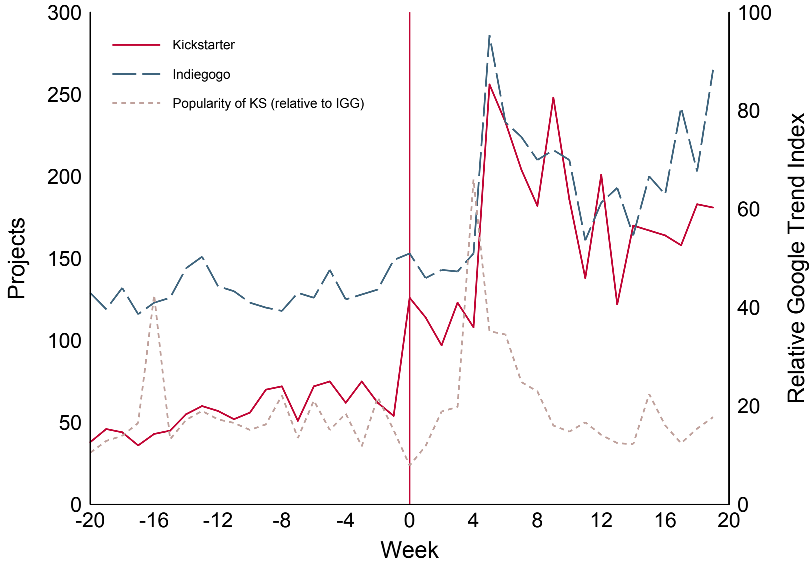
[28] To this end, we draw on the STATA package "matchit," see https://econpapers.repec.org/software/bocbocode/s457992.htm, last accessed on March 23, 2021.

[29] This method has been used in similar manner by other recent studies (e.g., Courtney et al., 2017; Madsen and McMullin, 2019).

[30] There is a growing literature that employs Google trends data to predict demand for as well as sales of various goods (e.g., Choi and Varian, 2012; Von Graevenitz et al., 2016).

popularity and awareness of crowdfunding as a means of raising project finance.

Figure 3.1: Technology Projects by Week at Kickstarter and Indiegogo



Notes: The chart shows the number of technology projects launching on Kickstarter and Indiegogo (all countries) by week (left axis), and the popularity of Kickstarter relative to Indiegogo as measured by Google trends indices (right axis). The vertical line at zero marks the time of the policy change.

There are two additional phenomena that help explain the general increase in crowdfunding projects on both platforms from week 5 onwards. First, our analysis shows that there is an annual peak of projects in the week following the 4th July (week 5 after the policy change) at both KS and IGG due to Independence Day in the United States. Second, crowdfunding in general and KS in particular, received a lot of media traction following the overly successful “potato salad” project.[31] The red dotted line in Figure 3.1 shows the popularity of KS relative to IGG, as measured by the difference in the Google trends indices for the two platforms. It shows how the increased public interest and awareness of KS preceded the spike in projects. Following the “potato salad” project, many creators called for more intervention by KS to prevent these kinds of projects. The fact that the potato salad project was allowed to launch on KS is evidence that the entry bars were lowered substantially (to non-existent) by abandoning the manual screening procedure.

Table 3.1 provides descriptive statistics for KS before and after the policy change. A naïve comparison of the means suggests a significant increase in the number of technology projects launched at KS. However, at the same time, we observe that KS as a platform and crowdfunding as a whole gained in popularity, which hints at a time trend and the need for a control group to identify the actual effect of the policy change on the number of projects. The increase in

[31] For more information on that project, see for instance http://money.cnn.com/2014/08/03/technology/social/potato-salad-crowdfunding/index.html, last accessed on March 23, 2021.

projects is not associated with an influx of small value projects. If at all, the pledge goal of projects increased following the policy change, albeit not at a statistically significant level. A comparison of the word count of project descriptions suggests a drop in the level of effort and preparedness by project creators. Project level variables related to financing suggest a decline both in terms of total money pledged and the overall funding rate. Looking at quality-related variables of the projects in the post-funding period, we see an overall decrease in the reward delivery rate and a strong surge in backer comments.

Table 3.1: Summary Statistics (Kickstarter)

	Pre-policy mean/sd	Post-policy mean/sd	Difference b/se
A. Platform variables (± 20 weeks):			
Projects	56.25	168.05	111.80***
	[12.19]	[45.97]	(10.63)
Platform popularity	36.77	42.90	6.13*
	[7.62]	[13.75]	(3.51)
Crowdfunding popularity	57.46	64.97	7.51***
	[2.85]	[6.25]	(1.54)
Observations (week level)	20	20	40
B1. Project variables (± 20 weeks):			
Pledge goal (in 1,000 EUR)	45.11	87.95	42.84
	[208.84]	[1,758.09]	(52.54)
Project description (in 1,000 words)	6.54	4.07	–2.48***
	[4.79]	[4.26]	(0.15)
B2. Project variables related to funding (± 20 weeks):			
Total pledged (in 1,000 EUR)	23.82	9.72	–14.10***
	[111.34]	[58.98]	(2.60)
Funding rate	0.30	0.16	–0.14***
	[0.46]	[0.36]	(0.01)
Observations (project level)	1,125	3,361	4,486
B3. Project variables related to quality (± 10 weeks):			
Reward delivered	0.75	0.61	–0.14***
	[0.44]	[0.49]	(0.05)
Project updates	6.34	6.32	–0.02
	[4.92]	[5.46]	(0.51)
Backer comments	261.79	342.63	80.83
	[754.64]	[1,395.67]	(114.82)
Observations (project level, conditional on funding)	177	262	439

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard deviations in square brackets; robust standard errors presented in parentheses. Tier A presents summary statistics for platform-specific variables at week level. Tier B presents summary statistics for project-specific variables at project level. The sample underlying the statistics in B.3 is restricted to projects that received funding and were launched within 10 weeks before and after the policy change.

3.6 Results

In the following, we analyze the effect of Kickstarter's change to an open platform on market thickness, the frequency of market matches, and project quality.

3.6.1 Market Thickness

We apply the synthetic control method to construct a synthetic control platform that best matches the trend in projects launching on KS in the pre-treatment period.[32] The synthetic control puts a large weight on Western Europe and Southern Europe. Other regions that receive positive, albeit smaller, weights are Central & Eastern Europe, Northern Europe and South America (see Table 3.2).

Table 3.2: Weights of the Synthetic Control Platform

Region	Africa	Asia	Central & Eastern Europe	Northern Europe	South America	Southern Europe	Western Europe	Other
Weight	-	-	0.12	0.07	0.06	0.33	0.42	-

Notes: This table lists the weights of regions that form the synthetic control platform ("KS-synth"). Each regions comprises all countries not eligible for Kickstarter projects ("IGG exclusive"). Regions without a weight are not part of the synthetic control.

Project Quantity. Figure 3.2a shows the number of technology projects for KS (treatment) and its synthetic counterpart (control) during the period 20 weeks before and after the policy shock. The vertical line marks KS's decision to abandon the manual screening in the first week of June 2014. The number of projects at KS and the synthetic control were about the same in the pre-policy change period, but they diverge noticeably following the decision for platform openness, with the number of KS projects doubling in the first week.

The estimated effect of KS's policy change on the number of projects launching on the platform is visualized in Figure 3.2b, which plots the gap between KS and the synthetic control. Our results from the synthetic control approach suggest that on average, over the 20 weeks following the policy change, the number of technology projects launched on KS increased by about 80 per week. This is an increase by 142% relative to the average level in the 20 weeks preceding the intervention at 56 projects.

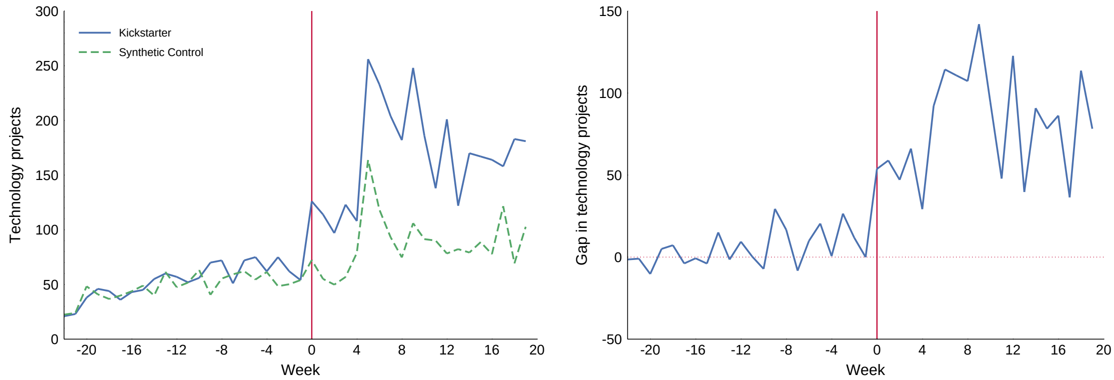
The trajectory of the gap is not smooth; while the gap widens to 50-60 projects per week in the first month following the policy change, it comes down in week four post-treatment. We see a marked increase of project creators launching projects in the week of Independence Day on both KS and IGG (and hence the synthetic version of KS). Nonetheless, the gap between KS and its synthetic counterpart is consistently above 30 projects, that is, 53% above the average level in the pre-policy change period. This suggests that the policy change led to a

[32]We hereby draw on the STATA package "synth_runner" by Galiani et al. (2017).

substantial increase in the weekly numbers of projects launching on KS.[33]

Figure 3.2: Treatment vs. Synthetic Control

(a) Trends in Number of Projects **(b)** Gap between Treatment and Control



Notes: Figure (a) on the left illustrates the trends of technology projects launching on Kickstarter and the synthetic control platform in absolute numbers by week. Figure (b) on the right depicts the difference in technology projects between Kickstarter and the synthetic control platform. A positive gap in technology projects indicates a larger number of projects launching on Kickstarter. The vertical line at zero marks the time of the policy change.

One valid concern in the context of the synthetic control method is that there are unobserved characteristics or effects after the policy change that are not captured by the synthetic version of KS.[34] We therefore estimate a difference-in-differences using the synthetic control group, comparing the average difference between KS and its synthetic counterpart before the policy change with the difference afterwards. As a cross-check, we also estimate a traditional difference-in-differences model across KS and all "IGG exclusive" countries as the control group, without weighting.

Table 3.3 shows the difference-in-differences results. We find that projects launching on KS each week increased by 51% after the policy change with high significance ($p<0.01$). The result holds regardless of whether we use the synthetic counterpart (column 2) or IGG exclusive countries (column 4) as the control group.

Project Variety and Novelty. Figure 3.3a shows the distribution of "Technology" projects across the 16 available subcategories during the 20 weeks before and 20 weeks after the policy change. In general, we see a shift from hardware based projects ("Hardware", "Technology (general)") to more software based projects ("Apps", "Web"). We calculate the Herfindahl-Hirschman-Index (HHI) as a measure for project variety before and after Kickstarter's policy

[33] We graph standardized significance levels for each lead period, based on the comparison of the estimated main effect for KS to the exact distribution of the placebo effects (see Figure C-3 in the Appendix).

[34] As the synthetic control is constructed based on (average) pre-intervention characteristics only, seasonal effects related to the 4th July or the "potato salad" project are not captured by the synthetic control.

Table 3.3: Number of Launched Projects

DV: Technology projects	(1) KS vs. KS-synth	(2) KS vs. KS-synth	(3) KS vs. IGG-excl	(4) KS vs. IGG-excl
Policy change (d)	1.047***	0.512***	1.039***	0.559***
	(0.092)	(0.091)	(0.086)	(0.070)
Time-varying controls	No	Yes	No	Yes
Month effects	No	Yes	No	Yes
Degrees of freedom	1	18	1	18
Log likelihood	–358.250	–302.226	–353.324	–292.076
Chi-squared	130.047	834.840	144.804	866.092
Observations	80	80	80	80

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Robust standard errors presented in parentheses. Columns (1) and (2) show the results of a difference-in-differences estimation with the synthetic control. Columns (3) and (4) show results of traditional difference-in-differences estimation. The coefficients are estimated with negative binomial models and can be interpreted as semi-elasticities. KS: Kickstarter. KS-synth: synthetic control. IGG-excl: IGG countries not eligible for KS projects. Time-varying controls include Google trends indices (current and lagged by 1 and 2 weeks) and platform age and time.

change.[35] We find that the HHI decreases from 0.23 to 0.13 in the post-policy period, indicating a higher level of variety on the platform following the removal of access control.

Figure 3.3b shows the distributions of novelty scores for "Technology" projects during the 20 weeks before and 20 weeks after the policy change. Overall, projects after the policy change appear to be neither more nor less novel than those projects beforehand. If at all, the post-policy distribution seems slightly more dispersed – with more projects of particularly low or high novelty. This, however, remains indicative. Applying the two-sample Kolmogorov-Smirnov test, we cannot reject the null-hypothesis that the pre- and post-policy distributions are the same (p=0.769).

Source of Projects. We argue that the increase in market thickness on the platform is primarily due to the entrance of low quality project that previously would have been rejected. However, there are additional channels (advertisement, backlog, and business stealing) that we briefly discuss. First, the increase in potential entrants may be due to a jump in marketing efforts or public awareness parallel to the policy change (week zero in Figure 3.2). We found no evidence that KS has increased its advertisement efforts during the time of observation.[36]

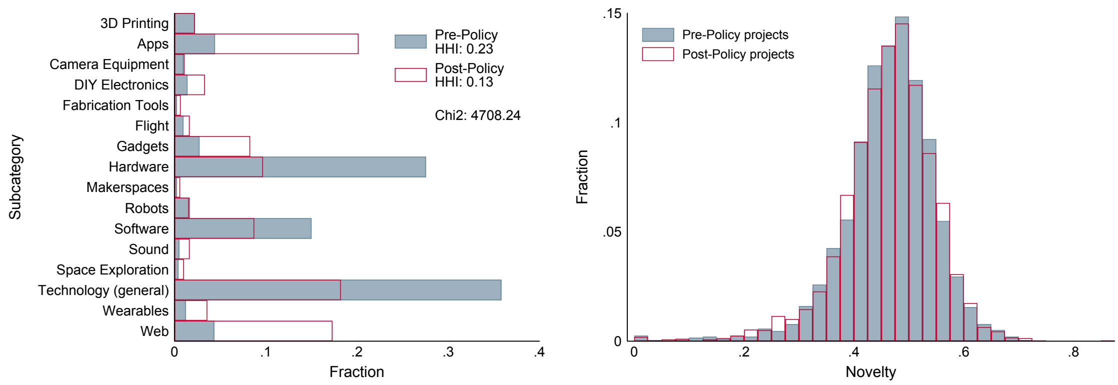
Second, the increase could originate from releasing pending projects from the backlog that the

[35]That is HHI $= \sum_{i=1}^{N} s_i^2$ where s^i is the share of "Technology" projects launching in subcategory i, and $N(= 16)$ is the number of subcategories.

[36]The announcement of the policy change was published at KS's developer blog. Media coverage was limited (we are aware of only two blogs mentioning it). Lastly, site visits and the Google trends index show no increase directly following the policy change.

Figure 3.3: Project Variety and Novelty

(a) Distribution of Projects by Subcategories **(b)** Distribution of Projects by Novelty



Notes: Figure (a) on the left shows the distribution of projects by subcategories. Sample includes all "Technology" projects launched within 20 weeks before and after the policy change at Kickstarter. Applying the Chi2-equality of distribution test for categorical variables, we can reject the null-hypothesis that the pre- and post-policy distributions are the same with $p<0.001$ and $chi2>4700$. Figure (b) on the right shows the distributions of projects by novelty score before and after the policy change at Kickstarter. Applying the two-sample Kolmogorov-Smirnov test, we cannot reject the null-hypothesis that the pre- and post-policy distributions are the same ($p=0.769$).

manual access control caused. The review process, now done algorithmically, can take as few as five minutes as opposed to several days or even weeks before. There was a steady increase in the "launch time" up until the first half of 2014, with a marked drop when KS decided to switch to an open platform.[37] This suggests there was indeed a backlog of projects waiting for approval prior to the policy change, creating capacity constraints for the platform to expand its marketplace.

Lastly, given the strong competition between the two leading crowdfunding platforms, it seems likely that IGG experienced some substitution effects in the "inclusive countries", as KS started to welcome projects that would have been rejected under the old stricter access control regime at KS and gone to IGG instead. We test for potential substitution effects between KS and IGG, using the indirect treatment group "IGG inclusive" in Table 3.4. The model in column (2) estimates the policy effect for KS relative to IGG in those countries where they compete. The estimated effect is positive and considerably larger relative to the main specification, which suggests that there is indeed a substitution effect when KS and IGG operate in the same country. The policy change led to a general increase of projects on KS, but also to a relative decrease of IGG projects in these contested countries, resulting in an overall larger estimated effect.

The estimation in column (4) further supports this result. It estimates a placebo effect of KS's policy change on the number of IGG projects in those countries where IGG and KS are competing ("IGG inclusive" projects), relative to the number of IGG projects where KS is not

[37]We use information on the date a project first registered on the platform and the date it went live to calculate the "launch time" for each project. Figure C-4 in the Appendix plots the average launch time for KS projects from 2011 to 2015.

present and the two platforms do not compete ("IGG exclusive" projects). We indeed find that the number of IGG projects in contested countries, where IGG was indirectly exposed to KS's decision to abandon access control, declined by 17% ($p<0.05$) – which translates to about 16 projects per week. With the strategic decision to open up, Kickstarter captured projects from IGG and hereby enlarged its market share. Furthermore, this suggests that, prior to the policy change, a considerable number of projects rejected or deterred by KS used to undertake platform shopping and go to IGG instead. Overall, this highlights that openness affects not only the platform's focal ecosystem but the crowdfunding market at large through shifting projects from one platform to another.

Above results on the number of launched projects and on competition are based on specifications that include time-varying controls and month fixed-effects. We show robustness of our results with week dummies in Table C-3 in the Appendix.

Table 3.4: Competition Effects

DV: Technology projects	(1) KS vs. IGG-incl	(2) KS vs. IGG-incl	(3) IGG-incl vs. IGG-excl	(4) IGG-incl vs. IGG-excl
Policy change (d)	1.031***	0.735***	0.347***	–0.168**
	(0.079)	(0.073)	(0.075)	(0.066)
Time-varying controls	No	Yes	No	Yes
Month effects	No	Yes	No	Yes
Degrees of freedom	1	18	1	17
Log likelihood	–358.107	–311.060	–338.280	–286.412
Chi-squared	171.709	570.072	21.530	306.697
Observations	80	80	80	80

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Robust standard errors presented in parentheses. Columns (1)-(4) show results of traditional difference-in-differences estimation. The coefficients are estimated with negative binomial models and can be interpreted as semi-elasticities. KS: Kickstarter. IGG-excl: IGG countries not eligible for KS projects. IGG-incl: IGG countries eligible for KS projects. Time-varying controls include Google trends indices (current and lagged by 1 and 2 weeks) and platform age and time.

3.6.2 Market Matches

We now turn to the effect of KS's decision to open up on the number of matches between project creators and backers. We define market matches as projects that reached their funding goal due to the contributions from backers. For this, we estimate weighted difference-in-differences models on the number of funded projects and the funding rate as the dependent variable.

Table 3.5 shows the results of the weighted difference-in-differences estimations. The result in column 1 suggests that the policy change is not associated with a clear change in project value, as the pledge goal shows a small and very noisy effect in the opposite direction (if at all). At the same time, there is a significant ($p<0.01$) increase in the absolute number of successfully funded projects by about 9 projects (column 2). The results suggest that the increased number

Table 3.5: Project Funding

Dependent variable:	(1) Pledge goal (in 1,000 EUR)	(2) Funded projects	(3) Funded projects (in 1,000 EUR)	(4) Funding rate
Policy change (d)	–34.746	8.560***	239.276	–0.085***
	(88.139)	(1.985)	(327.236)	(0.028)
Time-varying controls	Yes	Yes	Yes	Yes
Month effects	Yes	Yes	Yes	Yes
Degrees of freedom	17	17	17	17
Log likelihood	–358.107	–311.060	–338.280	–286.412
Observations	80	80	80	80

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Robust standard errors presented in parentheses. Columns (1)-(4) show the results of a difference-in-differences estimation with the synthetic control. The coefficients are estimated with linear models and can be interpreted as absolute changes. KS: Kickstarter. KS-synth: synthetic control. IGG-excl: IGG countries not eligible for KS projects. Fixed effects linear model. Time-varying controls include Google trends indices (current and lagged by 1 and 2 weeks) and platform age and time.

of projects coming onto the platform attracts more funds, which implies that KS increased its turnover thanks to the policy change. There is some evidence for that in column 3, even though the standard errors are very large, suggesting that turnover at week level is a highly volatile variable. We also find that the funding rate drops by 9 percentage points ($p<0.01$) following the policy change (column 4). Given that the average funding rate used to be around 30%, this change amounts to a significant reduction by almost a third.[38] This in turn means that the proportion of projects reaching their funding goal decreased considerably, making the platform less attractive to potential entrants on the supply side.

3.6.3 Project Quality

We finally turn to analyzing whether the abandonment of manual due diligence affected the backers' return on investment. For this, we focus at the subsample of successfully financed projects in the weeks before and after the policy change.

Reward Delivery. Backers contribute to a specific project in the expectation to receive a promised reward. As low quality projects should be more likely to fail in their endeavor, we assess the effect of the policy change on reward delivery. We group all successfully funded technology projects by their launch date, and compare the delivery performance before and after Kickstarter removed its upfront screening. Table 3.6 shows the results of multivariate analyses with increasingly narrower windows of observations around the cut-off point and project subcategory fixed effects.[39] We find a robust, largely significant drop in the reward de-

[38] There is some general downward trend of the funding rate to be observed before the policy change, which may be due to the fact that funding occurs weeks after project launch (see Figure C-2 in the Online Appendix).

[39] As elaborated above, we observe a change in project composition on the platform.

livery rate of around −10 to −15 percentage points. This removes the concern that our finding of the lower funding rate is driven by a particular subcategory. Given that the estimated effect remains fairly constant in magnitude when narrowing down the window of observation strongly suggests a discontinuous change due to the new open strategy.[40]

Table 3.6: Reward Delivery

	(1)	(2)	(3)	(4)
Sample definition:	± 10 weeks	± 7 weeks	± 5 weeks	± 3 weeks
DV: Reward delivered				
Policy change (d)	–0.110**	–0.150***	–0.139**	–0.095
	(0.047)	(0.056)	(0.064)	(0.082)
Subcategory effects	Yes	Yes	Yes	Yes
Degrees of freedom	15	15	15	15
Log likelihood	–272.506	–176.843	–124.664	–86.254
Observations	439	300	222	150

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Robust standard errors presented in parentheses. Columns (1)-(4) show the results of linear regression estimations. The coefficients can be interpreted as absolute changes. Sample includes all funded projects launched on Kickstarter within the respective time window before or after the policy change. The unit of observation is the funded project.

Backer Satisfaction. One may argue that the event of reward delivery is a rather crude proxy for quality. For instance, a lower propensity of reward delivery may imply a higher degree of risk, which could be compensated by the prospect of more attractive rewards. Second, the binary dependent variable of reward delivery neglects that sent out rewards may be of underwhelming quality or that only a subset of backers actually receives the reward.

The backer' appraisal of a project's quality can be most directly inferred through a sentiment analysis of comments on the project's website. We first denote each comment as either negative (a complaint) or neutral/positive depending on its score in our sentiment analysis. Comments are a heavily used way to communicate with other backers and the project creators. Even when we limit our window of observation to 5 weeks before and after the policy change, we have almost 23,000 comments for 158 different projects.

Table 3.7 shows that the policy change has led to a robust increase of complaints by around 4–6 percentage points (columns 1 to 4). If we focus on the relatively small set of comments by very experienced backers (so-called "superbackers" at KS), we find even larger coefficients: the significant effect is up to twice as strong for these "superbackers" with an increase in complaints by 7–12 percentage points (columns 5 to 8).

[40]We further test if observable pre-funding project characteristics are balanced to support the validity of our quasi-experimental pre- versus post-policy design. We find that all project level variables set by the project creator, like pledge goal, project duration, number of rewards offered and product (reward) value, are not significantly affected by the policy change (Table C-4 in the Appendix).

Table 3.7: Backer Complaints

	(1)	(2)	(3)	(4)
Sample definition:	± 10 weeks	± 7 weeks	± 5 weeks	± 3 weeks
DV: Backer complaints (share)				
Policy change (d)	0.058***	0.066***	0.036	0.046
	(0.018)	(0.022)	(0.025)	(0.032)
Subcategory effects	Yes	Yes	Yes	Yes
Degrees of freedom	15	15	15	13
Log likelihood	315.455	205.584	141.070	93.027
Observations	331	214	158	110

	(5)	(6)	(7)	(8)
Sample definition:	± 10 weeks	± 7 weeks	± 5 weeks	± 3 weeks
DV: "Superbacker" complaints (share)				
Policy change (d)	0.084***	0.119***	0.086**	0.071
	(0.029)	(0.031)	(0.033)	(0.048)
Subcategory effects	Yes	Yes	Yes	Yes
Degrees of freedom	11	11	10	10
Log likelihood	96.952	68.570	51.478	35.174
Observations	218	151	111	80

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Robust standard errors presented in parentheses. Columns (1)-(8) show the results of linear regression estimations. The coefficients can be interpreted as absolute changes. Sample includes all funded projects launched on Kickstarter within the respective time window before or after the policy change. The unit of observation is the funded project. Observations are weighted by the number of comments. Funded projects without ("superbacker") comment excluded. "Superbacker" is an official Kickstarter label given to backers with considerable prior backing activities.

3.6.4 Robustness

Alternative Specifications of the Synthetic Control Method. We select the specification where the resulting synthetic control has the best match quality, for which we report the results in the main text. Excluding number of pledges and Google Index data as predictor variables and matching based on lagged outcome variables only does not change the results qualitatively and barely quantitatively. Varying the pre-treatment periods to be included for the predictor variables or the length of the training period does not affect the overall results. The estimated effect of the policy change on the number of project is positive and significant across all tested alternative specifications.

Excluding Period after 4th July and "Potato Salad". We discussed that the peak of projects followed the coincidence of the 4th July and the "potato salad" project. To ensure that our results are not driven by these events, we run an alternative difference-in-differences specification, where we narrow the post-treatment window to four weeks – all coefficients of interest remain robust.

Placebo Test for 2013. We run a placebo test with KS data from 2013 over the same calendar time window to test if the estimated effects of the policy change are driven by seasonal effects. The placebo difference-in-differences analysis returns a negative coefficient, as opposed to a positive coefficient in our baseline estimation, for the effect of the policy change on the number of projects launching on the platform. There is no effect on either the funding rate or the absolute number of funded projects.

Alternative Samples for Reward Delivery. As the economic impact of investment failure depends on its monetary size, we weight projects by the total amount of money they collected and exclude small size projects with funding below 1,000 EUR. Additionally, we exclude projects where the return on investment was immaterial (e.g. a thank you note or an honorable mention). The negative effect of the policy change on the delivery rate remains throughout robust (see Table C-5 in the Appendix).

3.7 Conclusion

In this study, we analyzed the strategic decision of the leading reward-based crowdfunding platform, Kickstarter, to switch from access control to de facto openness and the resulting effects on crowdfunding activities. First, removing the manual pre-entry screening had an instant effect on market thickness, both in terms of higher quantity and variety of projects. However, openness had no significant effect on the projects' novelty. Second, the funding rate of projects went down, even though the absolute number of these market matches increased. Third, projects that reached their funding goal were of poorer quality. They more frequently failed to deliver their rewards and were subject to more complaints. These findings suggest that the demand side fell short in distinguishing projects of different qualities, and that openness seems to have lowered platform value for its ecosystem.

The results of our study raise the question about the rationale for KS to change its platform openness. Interestingly, KS chose to change this non-price parameter to modulate its platform ecosystem in the face of dynamic competition. KS never officially disclosed the reasoning for this change in platform design, but its course of action may be explained by considerations related to the cost and effectiveness of the screening procedure.

First of all, manual screening is costly and not scalable. As our analysis showed, it created a backlog of projects waiting for approval. Notably, KS was only available to project creators in English speaking countries prior to the policy change, but expanded to a number of new countries over the course of the following year. This suggests that removing the screening procedure also meant removing language barriers to international expansion. Second, there are limits to the ability of platforms to accurately evaluate projects by their quality. In case of ineffective screening, with many projects wrongfully excluded, this may even exacerbate the problem of low quality projects. Nevertheless, it seems that the platform has overestimated the ability of its users to judge the quality of entrepreneurial ventures.

We contribute to the crowdfunding and platform literature by studying openness as a key parameter of platform governance. We provide insights into the relationship between quality and platform openness, and show how an open strategy can affect the value of a platform in the context of platform competition. Using a quasi-experimental setting, we are able to disentangle the complicated effects of changes in openness on the economic activities in the crowdfunding market. We show that, in a thus far growing market environment, market thickness might trump project quality, and that openness can be an appropriate instrument to achieve this goal. However, as markets mature, quality is likely to become more relevant. Since we show that openness may come at the expense of project quality and funder experience, it is crucial for crowdfunding platforms to find the right balance between the two. In particular, strategic choices on platform governance should be considered in the context of the competitive environment a platform finds itself in.

We conclude with some limitations of our study. We note that our empirical analysis estimates a lower bound effect size. Even after the introduction of "Launch Now" project creators still had the option to voluntarily undergo a manual screening to receive feedback from a KS member of staff.[41] We expect these projects to be on average of higher quality than those that launch instantly without screening. Hence, our results can be interpreted as a lower bound. Moreover, our analysis focuses on technology projects. It would be interesting to see how platform openness impacts an ecosystem with projects of predominantly artistic nature, for which quality is less clearly defined. Finally, a causal inference on the long-term effects is beyond the scope of this paper, due to multiple subsequent changes at KS. However, we note that the policy change was reversed in early 2017 – Kickstarter re-instated the mandatory manual review for "Technology" and "Design" projects.[42] Platform openness did not seem to have generated the desired outcome for KS at least in some categories, and suggests heterogeneous effects across different types of projects. We leave a more detailed comparison between these project categories to future research.

[41]This implies that the new policy was not associated with perfect enforcement. Unfortunately, we cannot identify those projects that have undergone a voluntary screening. But since projects cannot signal to backers whether or not they did choose voluntary screening, there is no incentive to do it just for signaling purposes but for actually improving the project.

[42]This happened nearly clandestinely without any formal announcement. During the course of writing this paper, the authors of this paper noticed that KS changed the launch requirements for "Technology" and "Design" projects in the FAQ section of its website at https://www.kickstarter.com/help/faq/creator+questions?ref=faq_livesearch, last accessed on March 23, 2021.

A

Appendix to Chapter 1

The Transmission of Sectoral Shocks Across the Innovation Network

A.1 Appendix: Comparing First Stage Results to Bloom, Draca, van Reenen (2016)

How can we reconcile the negative coefficient of our first stage IV estimation with the positive coefficient found in BDVR? There are three key differences that can explain this. First, BDVR use a panel of textile firms while in this paper the main dataset is a panel of non-textile firms. In particular, as a reminder, the regressor in each observation is an average of patenting changes of all 2,380 textile firms, weighted by the technological distance to the non-textile firm. In turn, this leads to the second and third key differences: BDVR run an unweighted regression in which all textile firms are given the same weight. Contrary to their approach, the relative importance of our textile firms depends on their relative importance in the technology network and on the firm's recent patent stock; textile firms with a higher degree of network centrality (i.e. understood as closer technological ties to non-textile firms) and also firms with a higher patenting propensity are more likely to drive the results.

Table A.1: Direct Effects of Quotas Removal on Textile Firms

DV: ΔlnPat of TXT firms	Base	Weighted by		Sample split	
	(1)	(2) tech-connect.	(3) tech-connect. & pat.stock	(4) tech-connect. <p50	(5) tech-connect. >p50
Toughness of quotas in 2000	0.0401***	−0.0166	−0.1344**	0.0977***	−0.0248**
	(0.007)	(0.020)	(0.049)	(0.008)	(0.010)
Country FE	Yes	Yes	Yes	Yes	Yes
Industry(SIC-4D) FE	Yes	Yes	Yes	Yes	Yes
No. of clusters	13	13	13	13	13
Observations	11,900	11,900	11,900	5,905	5,995
Unique Firms	2,380	2,380	2,380	1,181	1,199

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the SIC4 industry level. This table presents the direct effects results for the panel of 2,380 textile firms in our sample. The dependent variable is the change in the log of patents by textile firms. The regressor of interest is the quota toughness prior to Chinas accession to the WTO. Column (1) is an unweighted regression on the full sample of textile firms. Column (2) weights each textile firm by its average technological proximity to the pool of non-textile firms, Column (3) additionally weights each textile firm by its patent stock. Columns (4) and (5) split the textile firms by the median technological proximity to non-textile firms.

Table A.1 illustrates these differences and how to reconcile the two approaches. Column (1) transforms the data to a panel of textile firms, as in BDVR. We replicate their results as we also obtain a positive estimated coefficient in this unweighted estimation. In Column (2), we weight each textile firm by its average technological proximity to the pool of non-textile firms. The coefficient turns negative (albeit not yet statistically significant), implying that the positive coefficient on Column (1) tends to be driven by textile firms with a low technological proximity to non-textile firms. In Column (3) we additionally weight by each firm's patent stock and now the estimated coefficient is both negative and statistically significant. In summary, while Columns (1)-(3) are based on exactly the same panel of textile firms, accounting for technological distance and firm size explains why we obtain a negative coefficient in our first stage. In Columns (4)-(5) we split the sample of textile firms by their average technological distance to non-textile firms. For firms with distant technological ties to non-textile firms we again replicate the positive and significant coefficient obtained in BDVR. For firms with close technological ties to non-textile firms we rather find a negative and significant coefficient, consistent with our first stage IV estimation results.

A.2 Appendix: Tables

Table A-1: Baseline Results
Conditional on Financial Data Availability for Sample Firms

DepVar: dlnPAT of NTXT firms	(1)	(2)	(3)	(4)
Method	IV 1st stage	IV 2nd stage	Reduced Form	OLS
Toughness of quotas in 2000	-0.5334^{***}		-0.0183^{***}	
	(0.005)		(0.005)	
dlnPAT of TXT firms		0.0343^{***}		0.0065^{*}
		(0.008)		(0.004)
Country-Year FE	Yes	Yes	Yes	Yes
Industry(SIC-4D)-Year FE	Yes	Yes	Yes	Yes
Underidentification test		41.1		
Weak identification test		13,962.0		
No. of clusters	394	394	394	394
Observations	121,710	121,710	121,710	121,710
Unique Firms	24,342	24,342	24,342	24,342

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the SIC4 industry level. This table estimates the same models as in the baseline IV, Reduced Form and OLS specifications, but restricted to the sample of firms for which we have non-missing financial data in ORBIS. Columns (1) and (2) correspond to Table 1.4 Columns (5) and (6). Column (3) corresponds to Table 1.3 Column (3), and Column (4) to Table 1.2 Column (3). All specifications include country-year fixed effects as well as a full set of year dummies interacted with a full set of industry 4-digit dummies.

Table A-2: Heterogeneity by Firm Size

DV: ΔlnPat of NTXT firms	(1) all (baseline)	(2) all w/ revenue info	(3) 1st quartile	(4) 2nd quartile	(5) 3rd quartile	(6) 4th quartile
ΔlnPat of TXT firms	0.0329^{***}	0.0334^{***}	0.0188^{**}	0.0194^{**}	0.0321^{**}	0.0636^{**}
	(0.006)	(0.007)	(0.007)	(0.010)	(0.013)	(0.026)
Country-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry(SIC-4D)-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Underidentification test	51.9	50.0	38.3	38.1	45.5	56.0
Weak identification test	15,609.7	14,509.7	12,323.9	9,007.0	9,237.6	11,591.7
No. of clusters	471	391	245	228	247	300
Observations	225,060	146,425	36,340	36,225	36,375	36,325
Unique Firms	45,012	29,285	7,268	7,245	7,275	7,265

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the SIC4 industry level. This table shows heterogeneous effects by non-textile firm size in an IV framework, where the second stage is presented. Firm size is determined by revenue, where firms in the 1st quartile are smallest. The dependent variable and regressor of interest are as in Table 1.2. Column (1) repeats the baseline result of Table 1.4 Column (6). Column (2) conditions on non-missing revenue info in ORBIS. The remaining columns (3)-(6) split the non-textile firms by size into quartiles. Column (3) restricts the sample to the smallest non-textile firms and Column (6) does the same for the largest non-textile firms. All specifications include country-year fixed effects as well as a full set of year dummies interacted with a full set of industry 4-digit dummies.

Table A-3: Accounting for Patent Quality

DepVar: Reduction in PAT of NTXT firms	(1) All patents	(2) EPO patents	(3) EPO cit.-weight
Reduction in PAT of TXT firms	0.0329***	0.0191***	0.0135**
	(0.006)	(0.006)	(0.005)
Country-year FE	Yes	Yes	Yes
Industry(SIC-4D)-year FE	Yes	Yes	Yes
Underidentification test	51.9	51.9	51.9
Weak identification test	15,609.7	15,609.7	15,609.7
No. of clusters	471	471	471
Observations	225,060	225,060	225,060
Unique Firms	45,012	45,012	45,012

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the SIC4 industry level. In this table we address patent quality by non-textile firms in an IV framework where the second stage is presented. Column (1) is our baseline result (as in Table 1.4 Column (6)), which includes patents filed with any patent authority. Column (2) restricts the patent count to those filed with the European Patent Office (EPO); as citations across patent authorities cannot be directly compared, we believe it is sensible to focus only on the most important patent authority for our European manufacturing firms. Finally, Column (3) weights the change in the log of patents by the number of citations received in the first five years post-grant. All specifications include country-year fixed effects as well as a full set of year dummies interacted with a full set of industry 4-digit dummies.

Table A-4: Alternative Lag Specification

DepVar: Reduction in PAT of NTXT firms	(1)	(2)	(3)
Reduction in PAT of TXT firms	0.0329***		
	(0.006)		
L.Reduction in PAT of TXT firms		0.0381***	
		(0.008)	
L2.Reduction in PAT of TXT firms			0.0588***
			(0.014)
Country-year FE	Yes	Yes	Yes
Industry(SIC-4D)-year FE	Yes	Yes	Yes
Underidentification test	51.9	52.4	52.4
Weak identification test	15,609.7	4,829.5	2,373.0
No. of clusters	471	471	471
Observations	225,060	180,048	135,036
Unique Firms	45,012	45,012	45,012

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the SIC4 industry level. Column (1) repeats our baseline IV result. Columns (2) and (3) use one-year and two-year lags of the regressor, respectively. All specifications include country-year fixed effects as well as a full set of year dummies interacted with a full set of industry 4-digit dummies.

Table A-5: Alternative Technological Proximity Measure

DV: ΔlnPat of NTXT firms	(1)	(2)	(3)	(4)	(5)	(6)
Method	1st stage	2nd stage	1st stage	2nd stage	1st stage	2nd stage
Toughness of quotas in 2000	-0.2128^{***}		-0.2139^{***}		-0.2156^{***}	
	(0.002)		(0.002)		(0.001)	
ΔlnPat of TXT		0.0094^{***}		0.0093^{***}		0.0078^{***}
		(0.002)		(0.002)		(0.002)
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry(SIC-2D)-Year FE	No	No	Yes	Yes	No	No
Industry(SIC-4D)-Year FE	No	No	No	No	Yes	Yes
Underidentification test		55.0		50.3		48.9
Weak identification test		9,005.7		18,000.9		24,951.9
No. of clusters	455	455	450	450	399	399
Observations	123,035	123,035	123,010	123,010	122,755	122,755
Unique Firms	24,607	24,607	24,602	24,602	24,551	24,551

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the SIC4 industry level. This table presents the instrumental variable estimation results for the full panel of non-textile firms, using an alternative measure of $tecih_{ij}$ that is based on the overlap in citation behavior between a firm-pair. The dependent variable is the change in the log of patents by non-textile firms. The instrument is the weighted change in the quota toughness prior to China's accession to the WTO for each textile firm. The endogenous regression is the weighted change in the log of patents by textile firms. Columns (1), (3), and (5) present first stage results, while columns (2), (4), and (6) present second stage results. Equations (1) and (2) are the baseline specification that control for country-year fixed effects. Equations (3) and (4) additionally include a full set of year dummies interacted with a full set of industry 2-digit dummies. Equations (5) and (6) instead include a full set of year dummies interacted with a full set of industry 4-digit dummies. The table reports test statistics for underidentification (Kleiberger-Paap rk LM statistic) and weak identification (Kleibergen-Paap rk Wald statistic).

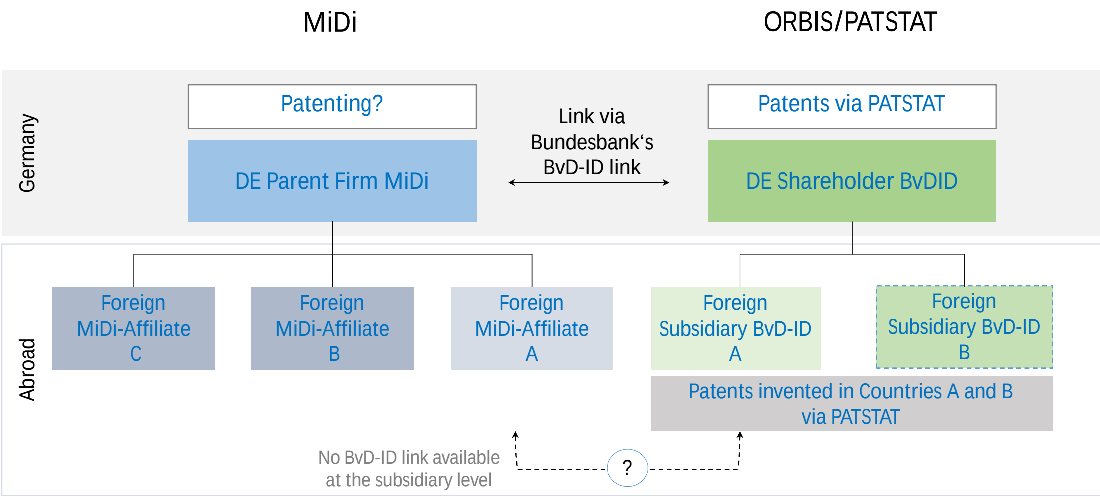
B

Appendix to Chapter 2

Multinational Innovation and China's Science and Technology Boom

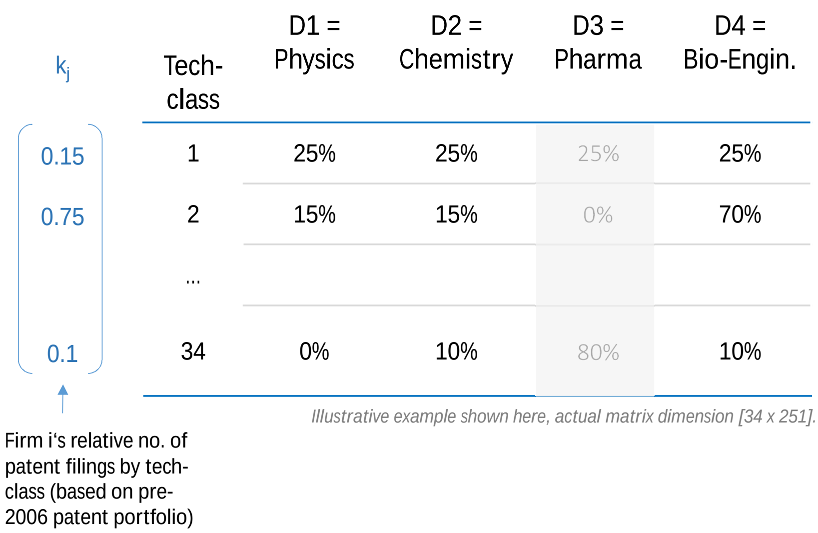
B.1 Appendix: Figures

Figure B-1: Linking MiDi and Patent Data (via BvD's ORBIS)



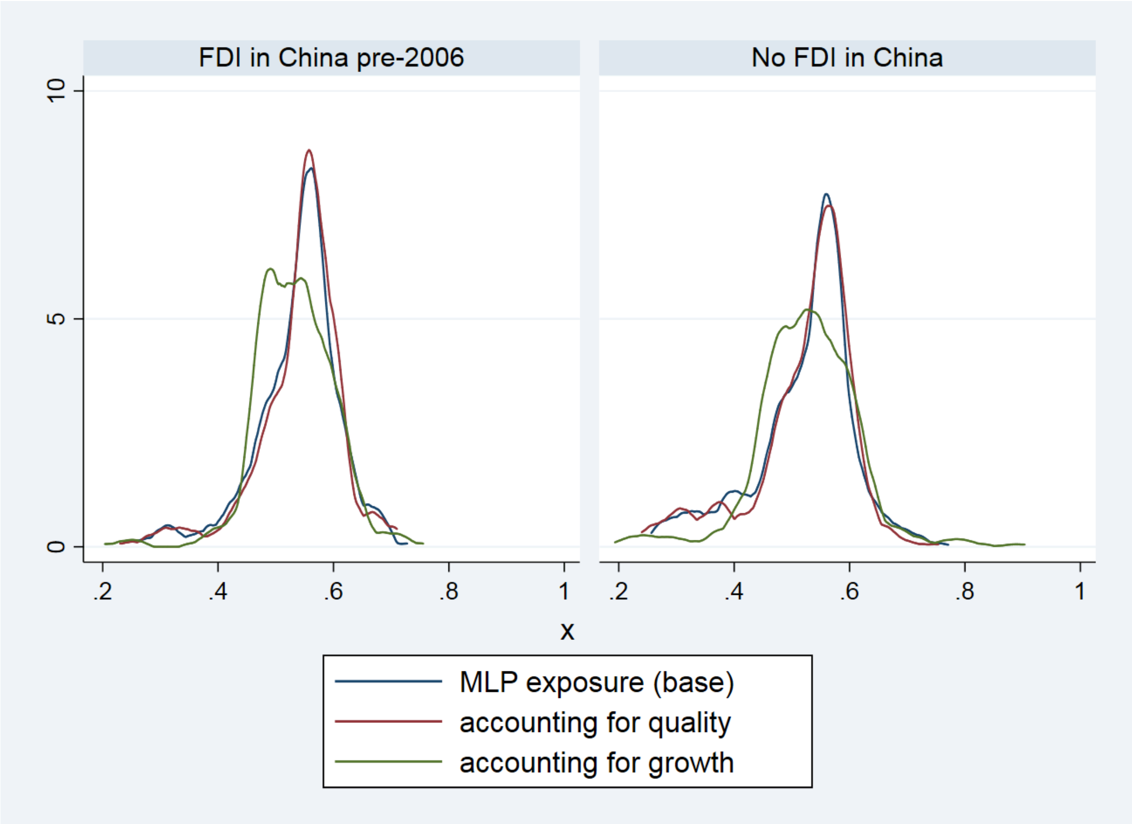
Source: Own illustration.

Figure B-2: Illustration of the Technology-to-Science Concordance Matrix



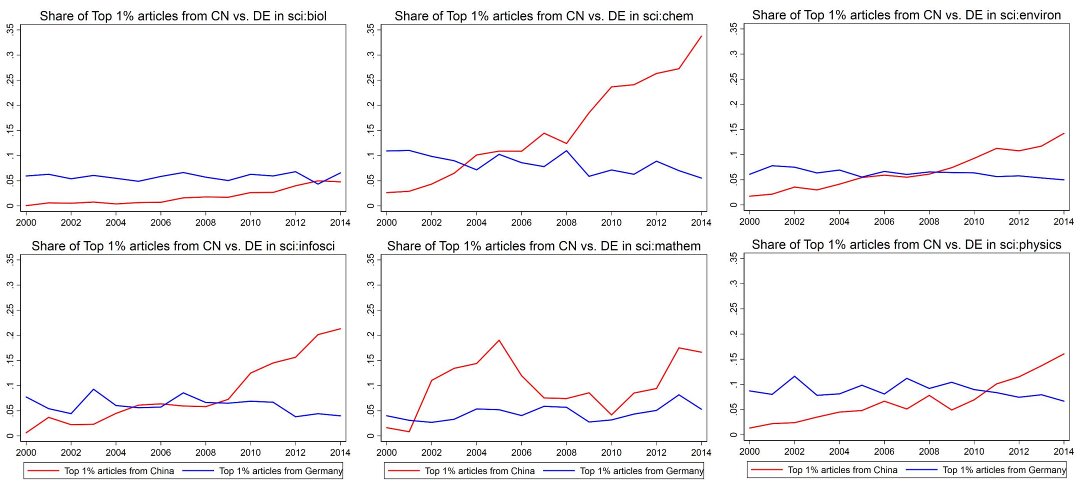
Source: Own illustration.

Figure B-3: MLP Measures - Kernel Density



Source: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, Microdatabase Direct Investment (MiDi), 1999-2016, Web of Science, PATSTAT, own calculations.

Figure B-4: Top 1% Cited Articles from China vs. from Germany

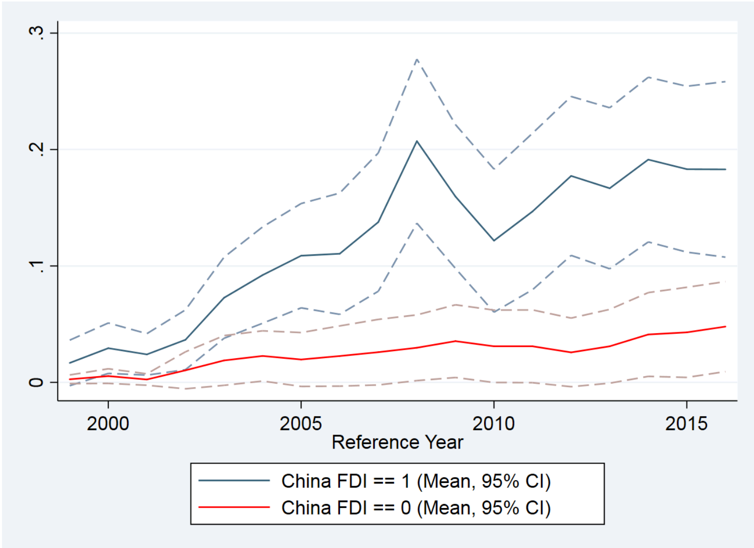


Source: Web of Science, own calculations.
Notes: This figure shows the share of top 1% cited scientific articles that comes from China versus the share of articles that comes from Germany, for various scientific fields (from top left, clockwise): biology, chemistry, environmental science, information science, mathematics, physics. Note that the scientific fields are shown here at a more aggregate level than is actually used in the analysis.

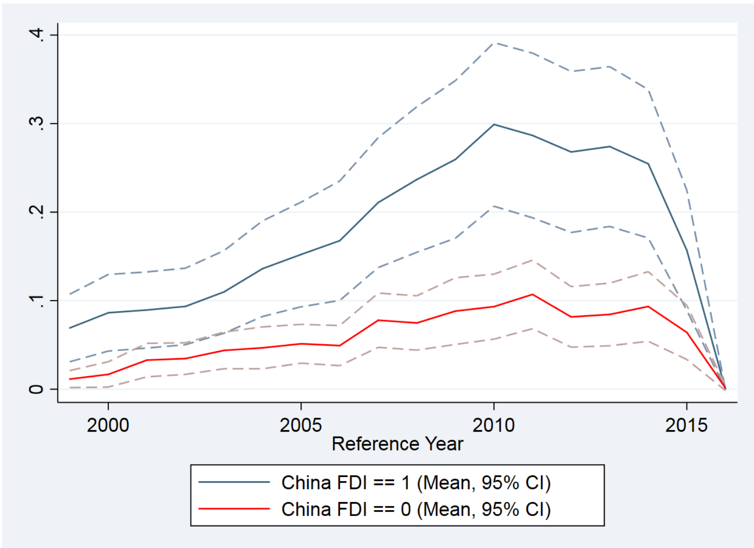
Figure B-5: Difference-in-Differences Estimation: Annual Betas

Source: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, Microdatabase Direct Investment (MiDi), 1999-2016, Web of Science, PATSTAT, own calculations.

Figure B-6: Comparing Outcome Variables for MNEs with vs. without FDI in China



(a) Inventions in China



(b) Citations to Chinese articles

Source: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, Microdatabase Direct Investment (MiDi), 1999-2016, PATSTAT, own calculations.
Notes: Figure (a) shows the log number of patents invented in China for MNEs with FDI in China and for MNEs without FDI in China. Figure (b) shows the log number of patent citations to Chinese scientific articles for the two types of MNEs. The dashed lines give the 95% confidence intervals.

B.2 Appendix: Tables

Table B-1a: Summary Statistics for MNEs in China: MLP High vs. Low

Pre-2006 average	Low MLP Exposure mean/sd	High MLP Exposure mean/sd	Difference b/se
No. of Employees of DE mother firm	1,833.09	2,402.92	569.83
	[5,112.8]	[9,786.1]	(845.23)
Turnover of DE mother firm (in m€)	534.51	802.30	267.80
	[1,571.8]	[4,073.2]	(334.87)
Balance sheet total of DE mother firm (in m€)	1,206.50	1,149.45	-57.05
	[3,603.8]	[4,002.4]	(410.08)
No. of FDI countries	9.43	6.92	-2.51**
	[11.1]	[7.6]	(1.01)
No. of Subsidiaries	14.39	11.60	-2.79
	[21.4]	[19.2]	(2.17)
FDI, primary + secondary (in m€)	400.22	336.92	-63.31
	[1,336.7]	[1,110.5]	(131.33)
No. of Employees of the foreign subsidiaries	2,927.13	4,171.06	1,243.94
	[6,781.4]	[14,436.5]	(1,215.42)
Turnover of the foreign subsidiaries (in m€)	924.59	1,053.80	129.21
	[2,640.6]	[3,948.8]	(361.06)
Balance sheet total of foreign subsidiaries (in m€)	960.32	996.57	36.26
	[3,210.5]	[3,646.4]	(368.33)
Patent stock as of 2005	1,877.82	1,387.53	-490.30
	[9,217.3]	[7,245.1]	(885.57)
Year of 1st patent	1,962.27	1,967.37	5.09
	[31.8]	[29.9]	(3.31)
Patent inventions per year	57.67	69.29	11.62
	[207.0]	[333.7]	(29.87)
Patents invented in CN	0.18	0.14	-0.05
	[1.2]	[0.6]	(0.10)
Share of patents fully or partly invented in CN	0.00	0.01	0.01*
	[0.0]	[0.1]	(0.01)
NPL citations to scientific articles	69.21	7.61	-61.60*
	[416.2]	[53.5]	(31.46)
NPL citations to CN articles	0.76	0.12	-0.64**
	[4.0]	[0.7]	(0.31)
Share of NPL citations to CN articles	0.03	0.07	0.03
	[0.1]	[0.2]	(0.03)
Observations	171	178	349

Source: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, Microdatabase Direct Investment (MiDi), 1999-2016, Web of Science, PATSTAT, own calculations.
Notes: Sample includes 349 German patenting MNEs in the MiDi with FDI in China as of 2006; distinction between 171 MNEs with a low and 178 MNEs with a high exposure to MLP targeted scientific fields. Figures show pre-period averages (1999-2006).

Table B-1b: Summary Statistics: MNEs with vs. without FDI in China

Pre-2006 average	No FDI in China mean/sd	FDI in China mean/sd	Difference b/se
No. of Employees of DE mother firm	974.77	2,124.59	1,149.82**
	[6,640.7]	[7,852.7]	(482.91)
Turnover of DE mother firm (in m€)	388.33	671.50	283.17*
	[2,062.8]	[3,112.0]	(170.01)
Balance sheet total of DE mother firm (in m€)	784.48	1,177.31	392.83
	[5,169.7]	[3,807.5]	(319.57)
No. of FDI countries	2.85	8.15	5.30***
	[3.1]	[9.5]	(0.42)
No. of Subsidiaries	3.68	12.97	9.29***
	[4.7]	[20.3]	(0.86)
FDI, primary + secondary (in m€)	68.13	367.94	299.81***
	[254.5]	[1,225.2]	(51.42)
No. of Employees of the foreign subsidiaries	830.30	3,561.57	2,731.27***
	[2,942.5]	[11,351.5]	(485.69)
Turnover of the foreign subsidiaries (in m€)	188.70	990.49	801.80***
	[707.0]	[3,367.7]	(141.46)
Balance sheet total of foreign subsidiaries (in m€)	312.08	978.81	666.73***
	[2,055.3]	[3,434.9]	(177.30)
Patent stock as of 2005	367.84	1,627.76	1,259.92***
	[2,393.0]	[8,262.0]	(358.09)
Year of 1st patent	1,972.41	1,964.87	-7.54***
	[26.7]	[30.9]	(1.90)
Patent inventions per year	15.79	63.60	47.81***
	[108.5]	[278.6]	(12.68)
Patents invented in CN	0.31	0.16	-0.15
	[7.5]	[1.0]	(0.40)
Share of patents fully or partly invented in CN	0.00	0.01	0.01**
	[0.0]	[0.1]	(0.00)
NPL citations to scientific articles	7.34	37.79	30.45**
	[51.3]	[295.0]	(12.25)
NPL citations to CN articles	0.08	0.43	0.35***
	[0.6]	[2.9]	(0.12)
Share of NPL citations to CN articles	0.05	0.05	-0.00
	[0.2]	[0.1]	(0.02)
Observations	610	349	959

Source: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, Microdatabase Direct Investment (MiDi), 1999-2016, Web of Science, PATSTAT, own calculations.
Notes: Sample includes 959 German patenting MNEs in the MiDi, 610 without and 349 with FDI in China. Figures show pre-period averages (1999-2006).

Table B-2: Citations from Patent Tech-Classes to Scientific Disciplines (Examples)

Tech Area	Scientific Subject	Share of Citations to Subject	Cumulative Sum
Semiconductor	Physics, Applied	26%	26%
	Engineering, Electrical & Electronic	20%	46%
	Materials Science, Multidisciplinary	9%	56%
	...	...	...
Medical Technology	Cardiovascular System & Cardiology	9%	9%
	Surgery	8%	17%
	Engineering, Biomedical	7%	24%
	Radiology, Nuclear Medicine & Med. Imaging	5%	29%
	Clinical Neurology	4%	33%
	...	...	...

Source: Web of Science, PATSTAT, own calculations.
Notes: This table illustrates that technology areas vary in the degree they rely on knowledge from scientific disciplines, and in the diversity of fields they cite. The five technology classes with the most and the least diverse set NPL references to scientific fields are:

Bottom 5 (most concentrated)

1. Basic Comm. Process
2. Telecom
3. Digital Comm.
4. Optics
5. Semiconductors

Top 5 (most diverse)

1. Other Machines
2. Mech. Elements
3. Engines/Pumps/Turbines
4. Medical Techn.
5. Textiles/Paper Machines

Table B-3: Mapping the MLP Policy into Scientific Disciplines

MLP Section	MLP Item	WoS Code	Science Field (WoS Name)
Main Areas	Clean, efficient coal development and utilization, coal liquefaction, and gasification-based co-generation	zq id	Mining & Mineral Processing Energy & Fuels
Basic Research	New Principles and methodologies for materials design and fabrication	pm ub dw	Materials Science, Multidisciplinary Physics, Applied Chemistry, Applied
Basic Research	Major mechanical issues in aeronautics aeronautics and space science	ai dt	Engineering, Aerospace Thermodynamics
Frontier Technologies	Stem cell based human tissue engineering technology	km	Genetics & Heredity

Source: MLP 2006, Web of Science, own analysis.
Notes: This table illustrates the mapping of the subjects identified as focal in the MLP into scientific disciplines. A MLP subject can be associated with one or multiple scientific disciplines, where applicable. Classification scheme of scientific disciplines as per Web of Science.

Table B-4: Difference-in-Differences Estimation

	Inventions in CN			Citations to CN articles		
	(1)	(2)	(3)	(4)	(5)	(6)
MLP=1 × Post=1	−0.020	−0.020	−0.017	−0.074 **	−0.072 **	−0.069 *
	(0.046)	(0.045)	(0.046)	(0.036)	(0.036)	(0.036)
MLP=1			−0.010			−0.113 **
			(0.026)			(0.045)
Post=1		0.115 ***			0.149 ***	
		(0.033)			(0.030)	
Firm FE	Yes	Yes	No	Yes	Yes	No
Year FE	Yes	No	Yes	Yes	No	Yes
Observations	5,882	5,882	5,882	5,882	5,882	5,882
Unique Firms	349	349	349	349	349	349
Adj. R-squared	0.63	0.63	0.01	0.72	0.70	0.03

Source: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, Microdatabase Direct Investment (MiDi), 1999-2016, Web of Science, PATSTAT, own calculations.
Notes: ***/**/* indicate significance at the 1%/5%/10% level. Standard errors are clustered at the firm level. This table presents results from a linear difference-in-differences estimation using firms' different exposure to the MLP policy as treatment variable. The years after 2006 are indicated by the 'Post' dummy. The sample consists of 349 patenting German MNEs in the MiDi with FDI in China as of 2006.

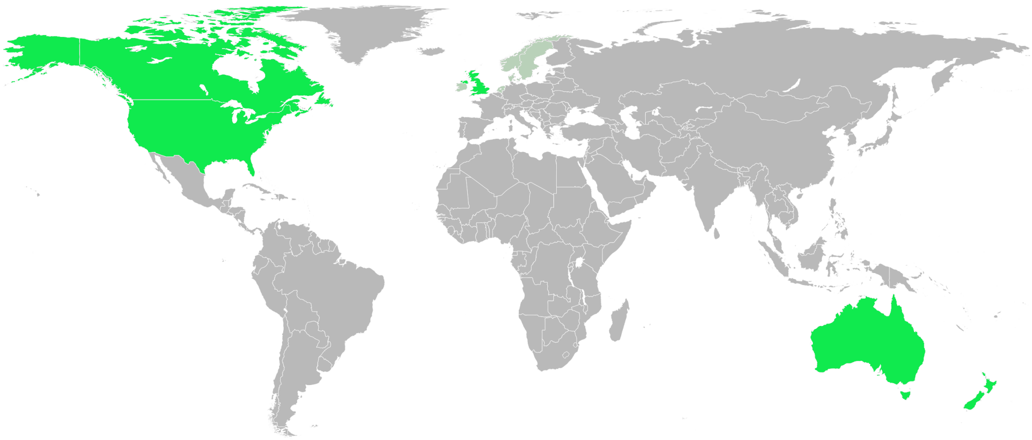
C

Appendix to Chapter 3

Openness as Platform Strategy: Evidence from a Quasi-Experiment in Crowdfunding

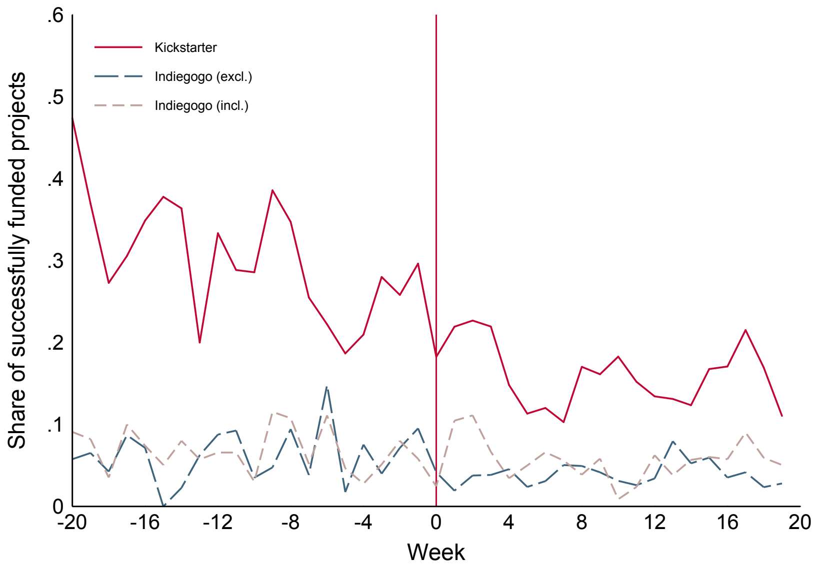
C.1 Appendix: Figures

Figure C-1: Geographical Coverage for Kickstarter Projects



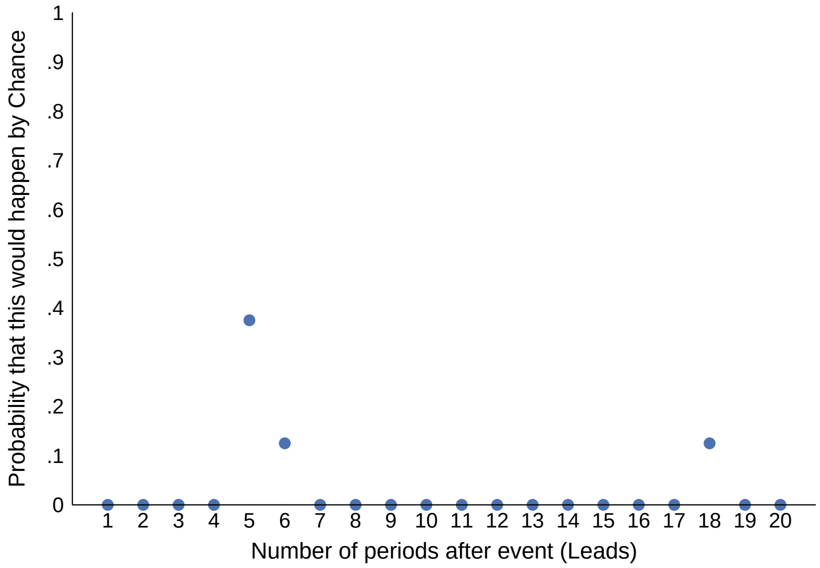
Notes: As of beginning of 2014: USA; GBR, CAN, AUS, NZL. Introduced during 2014: NLD (April), DNK, IRL, NOR, SWE (all October). Non-eligible countries in 2014.

Figure C-2: Funding Rate (per Week) by Platform



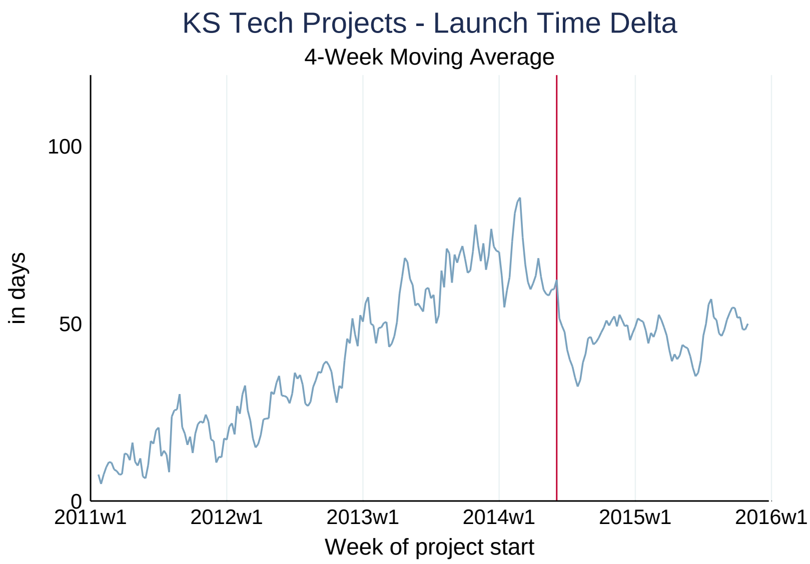
Notes: The figure shows the average funding rate of technology projects launching on Kickstarter, Indiegogo (exclusive) and Indiegogo (inclusive) by week. The vertical line at zero marks the time of the policy change.

Figure C-3: Lead Specific Significance Levels



Notes: The figure plots the p-values per-period for the months after the policy change for standardized effects (see Stata's $synth_runner$ package for details), and shows that they are signficant in all but one period.

Figure C-4: "Launch Time Delta" Trend 2011-2015 (4-Week Moving Averages)



Notes: The figure shows the time trend of the launch time delta, which is the time lag between first registration of the project and launch of the campaign on Kickstarter. Launch time delta is depicted as a 4-week moving average. The vertical line at zero marks the time of the policy change.

C.2 Appendix: Tables

Table C-1: Eligible Countries for Project Creators at Kickstarter

Country	Introduction date	Launch date
United States	22/04/2009	22/04/2009
United Kingdom	10/10/2012	31/10/2012
Canada	05/08/2013	09/09/2013
Australia	15/10/2013	13/11/2013
New Zealand	15/10/2013	13/11/2013
Netherlands	07/04/2014	29/04/2014
Denmark	15/09/2014	21/10/2014
Ireland	15/09/2014	21/10/2014
Norway	15/09/2014	21/10/2014
Sweden	15/09/2014	21/10/2014

Source: https://www.kickstarter.com/blog.
Notes: Only countries introduced in 2009-2014 listed.

Table C-2: Events at Kickstarter and Indiegogo

Date	Platform	Event
19/12/2013	KS	New feature: Advanced Search
11/06/2014	KS	New categories: Journalism and Crafts
11/12/2014	IGG	New feature: Android App

Source: https://www.kickstarter.com/blog,
and https://en.go.indiegogo.com/blog.

Table C-3: Number of Launched Projects (Week Fixed Effects)

DV: Technology projects	(1) KS vs. KS-synth	(2) KS vs. IGG-excl	(3) KS vs. IGG-incl	(4) IGG-incl vs. IGG-excl
Policy change (d)	0.543***	0.599***	0.746***	–0.151***
	(0.052)	(0.052)	(0.049)	(0.052)
Week effects	Yes	Yes	Yes	Yes
Degrees of freedom	40	40	40	40
Log likelihood	–271.699	–267.019	–284.875	–255.183
Chi-squared	1637.916	1545.003	1442.160	474.834
Observations	80	80	80	80

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Robust standard errors presented in parentheses. Columns (1) and (2) mirror the specifications in Table 3.3, columns (2) and (4), with week instead of month fixed effects. Likewise, columns (3) and (4) mirror the specifications in Table 3.4, columns (2) and (4), with week instead of month fixed effects. KS: Kickstarter. KS-synth: synthetic control. IGG-excl: IGG countries not eligible for Kickstarter projects. IGG-incl: IGG countries eligible for Kickstarter projects. Time-varying controls include Google trends indices (current and lagged by 1 and 2 weeks) and platform age and time.

Table C-4: Reward Delivery (Robustness Checks I)

DV:	(1) Pledge goal (in 1,000 EUR)	(2) Project duration	(3) Number of project rewards	(4) Average reward value
Policy change (d)	3.501	–1.713	0.618	0.157
	(5.297)	(1.249)	(0.593)	(0.353)
Subcategory effects	Yes	Yes	Yes	Yes
Degrees of freedom	15	15	15	15
Log likelihood	–1101.031	–785.606	–620.103	–504.992
Observations	222	222	222	222

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Robust standard errors presented in parentheses. Columns (1)-(4) show the results of linear regression estimations. The coefficients can be interpreted as absolute changes. Sample includes all funded projects launched on Kickstarter within 5 weeks before or after the policy change. The unit of observation is the funded project.

Table C-5: Reward Delivery (Robustness Checks II)

Sample definition: DV: Reward delivered	(1) Projects weighted by total pledged	(2) Total pledged $\geq$ 1000 EUR	(3) Only material rewards	(4) All of previous
Policy change (d)	–0.126**	–0.168**	–0.157**	–0.135**
	(0.053)	(0.065)	(0.067)	(0.059)
Subcategory effects	Yes	Yes	Yes	Yes
Degrees of freedom	15	15	14	14
Log likelihood	–75.550	–108.256	–104.258	–61.385
Observations	222	204	191	180

Notes: ***/**/* indicate significance at the 1%/5%/10% level. Robust standard errors presented in parentheses. Columns (1)-(4) show the results of linear regression estimations. The coefficients can be interpreted as absolute changes. Sample includes all funded projects launched on Kickstarter within 5 weeks before or after the policy change. The unit of observation is the funded project. In column (1), funded projects are weighted by the total money pledged (in 1,000 EUR). The sample in column (2) includes all funded projects with at least 1,000 EUR in total pledged. The sample in column (3) includes all funded projects with material rewards (and immaterial rewards excluded). The sample in column (4) combines all previous sample criteria. It includes all funded projects launched on Kickstarter within 5 weeks before or after the policy change with at least 1,000 EUR in total pledged and material rewards, and observations being weighted by the total money pledged (in 1,000 EUR).

List of Figures

List of Tables

Bibliography

Abadie, A., A. Diamond, and J. Hainmueller (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association 105*(490), 493–505.

Abadie, A., A. Diamond, and J. Hainmueller (2015). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science 59*(2), 495–510.

Abadie, A. and J. Gardeazabal (2003). The Economic Costs of Conflict: A Case Study of the Basque Country. *The American Economic Review 93*(1), 113–132.

Acemoglu, D., U. Akcigit, and W. R. Kerr (2016). Innovation Network. *Proceedings of the National Academy of Sciences of the United States of America 113*(41), 11483–11488.

Adams, J. D., G. C. Black, J. R. Clemmons, and P. E. Stephan (2005). Scientific Teams and Institutional Collaborations: Evidence from U.S. Universities, 1981-1999. *Research Policy 34*(3), 259–285.

Afuah, A. (2013). Are Network Effects Really all about Size? The Role of Structure and Conduct. *Strategic Management Journal 34*(3), 257–273.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005). Competition and Innovation: An Inverted-U Relationship. *The Quarterly Journal of Economics 120*(2), 701–728.

Agrawal, A. K., C. Catalini, and A. Goldfarb (2014). Some Simple Economics of Crowdfunding. *Innovation Policy and the Economy 14*(1), 63–97.

Agrawal, A. K., C. Catalini, and A. Goldfarb (2015). Crowdfunding Geography, Social Networks, and the Timing of Investment Decisions. *Journal of Economics & Management Strategy 24*(2), 253–274.

Ahlers, G. K., D. Cumming, C. Günther, and D. Schweizer (2015). Signaling in Equity Crowdfunding. *Entrepreneurship Theory & Practice 39*(4), 955–980.

Ahmadpoor, M. and B. F. Jones (2017). The Dual Frontier: Patented Inventions and Prior Scientific Advance. *Science 357*(6351), 583–587.

Aitken, B. and A. Harrison (1999). Do Domestic Firms Benefit from Foreign Direct Investment? Evidence from Panel Data. *The American Economic Review 89*(3), 605–618.

Alcácer, J. and W. Chung (2007). Location Strategies and Knowledge Spillovers. *Management Science 53*(5), 760–776.

Alcácer, J., M. Gittelman, and B. Sampat (2009). Applicant and Examiner Citations in U.S. Patents: An Overview and Analysis. *Research Policy 38*(2), 415–427.

Almeida, P. (1996). Knowledge Sourcing by Foreign Multinationals: Patent Citation Analysis in the U.S. Semiconductor Industry. *Strategic Management Journal 17*(S2), 155–165.

Almeida, P. and B. Kogut (1999). Localization of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science 45*(7), 905–917.

Amit, R., J. Brander, and C. Zott (1998). Why do Venture Capital Firms Exist? Theory and Canadian Evidence. *Journal of Business Venturing 13*(6), 441–466.

Amiti, B. M. and J. Konings (2007). Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia. *American Economic Review 97*(5), 1611–1638.

Armstrong, M. (2006). Competition in Two-sided Markets. *The RAND Journal of Economics 37*(3), 668–691.

Arnosti, N., R. Johari, and Y. Kanoria (2019). Managing Congestion in Matching Markets. *Manufacturing and Service Operations Management*, forthcoming.

Arts, S., B. Cassiman, and J. C. Gomez (2018). Text Matching to Measure Patent Similarity. *Strategic Management Journal 39*(1), 62–84.

Audretsch, D. B. and M. P. Feldman (1996). R&D Spillovers and the Geography of Innovation and Production. *American Economic Review 86*(3), 630–640.

Autor, D., D. Dorn, G. H. Hanson, P. Shu, and G. Pisano (2020). Foreign Competition and Domestic Innovation: Evidence from U.S. Patents. *American Economic Review: Insights 2*(3), 357–374.

Aw, B., M. J. Roberts, and D. Y. Xu (2011). R&D Investment, Exporting, and Productivity Dynamics. *American Economic Review 101*(4), 1312–1344.

Bacchiocchi, E. and F. Montobbio (2010). International Knowledge Diffusion and Home-Bias Effect: Do USPTO and EPO Patent Citations Tell the Same Story? *Scandinavian Journal of Economics 112*(3), 441–470.

Belleflamme, P., N. Omrani, and M. Peitz (2015). The Economics of Crowdfunding Platforms. *Information Economics and Policy 33*, 11–28.

Belleflamme, P. and E. Toulemonde (2009). Negative Intra-group Externalities in Two-sided Markets. *International Economic Review 50*(1), 245–272.

Belleflamme, Paul, T. L. and A. Schwienbacher (2014). Crowdfunding: Tapping the Right Crowd. *Journal of Business Venturing 29*(5), 585–609.

Bernard, A. B., A. Moxnes, and Y. U. Saito (2019). Production Networks, Geography, and Firm Performance. *Journal of Political Economy 127*(2), 639–688.

Bernstein, J. I. and M. I. Nadiri (1989). Research and Development and Intra-Industry Spillovers: An Empirical Application of Dynamic Duality. *The Review of Economic Studies 56*(2), 249.

Berry, H. and A. Kaul (2015). Global Sourcing and Foreign Knowledge Seeking. *Management Science 61*(5), 1052–1071.

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

Blit, J. (2017). Learning Remotely: R&D Satellites, Intra-Firm Linkages, and Knowledge Sourcing. *Journal of Economics and Management Strategy 26*(4), 757–781.

Bloom, N., M. Draca, and J. Van Reenen (2016). Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity. *Review of Economic Studies 83*(1), 87–117.

Bloom, N., M. Schankerman, and J. Van Reenen (2013). Identifying Technology Spillovers and Product Market Rivalry. *Econometrica 81*(4), 1347–1393.

Boeing, P., E. Mueller, and P. Sandner (2016). China's R&D Explosion - Analyzing Productivity Effects Across Ownership Types and Over Time. *Research Policy 45*(1), 159–176.

Boudreau, K. and A. Hagiu (2009). *Platforms, Markets and Innovation*, Chapter 7 Platform Rules: Multi-sided Platforms as Regulators, pp. 163–191. Edward Elgar.

Boudreau, K. J. (2012). Let a Thousand Flowers Bloom? An Early Look at Large Numbers of Software App Developers and Patterns of Innovation. *Organization Science 23*(5), 1409–1427.

Branstetter, L. (2006). Is Foreign Direct Investment a Channel of Knowledge Spillovers? Evidence from Japan's FDI in the United States. *Journal of International Economics 68*(2), 325–344.

Branstetter, L., G. Li, and F. Veloso (2014). The Rise of International Co-Invention. In A. B. Jaffe and B. F. Jones (Eds.), *The changing frontier: Rethinking science and innovation policy*, pp. 135–168. University of Chicago Press.

Branstetter, L. G. (2001). Are Knowledge Spillovers International or Intranational in Scope? Microeconometric Evidence from the U.S. and Japan. *Journal of International Economics 53*(1), 53–79.

Breschi, S. and F. Lissoni (2005). "Cross-Firm" Inventors and Social Networks: Localized Knowledge Spillovers Revisited. *Annales d'Économie et de Statistique* (79/80), 189–209.

Bresnahan, T., J. Orsini, and P.-L. Yin (2015). Demand Heterogeneity, Inframarginal Multihoming, and Platform Market Stability: Mobile Apps. Mobile Innovation Group Working Paper.

Burtch, G., A. Ghose, and S. Wattal (2015). The Hidden Cost of Accommodating Crowdfunder Privacy Preferences: a Randomized Field Experiment. *Management Science 61*(5), 949–962.

Bustos, P. (2011). Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms. *American Economic Review 101*(1), 304–340.

Cai, J., A. M. Santacreu, and N. Li (2017). Knowledge Diffusion, Trade and Innovation Across Countries and Sectors. FRB St. Louis Working Paper No. 2017-29.

Callaert, J., B. Van Looy, A. Verbeek, K. Debackere, and B. Thijs (2006). Traces of Prior Art: An Analysis of Non-Patent References Found in Patent Documents. *Scientometrics 69*(1), 3–20.

Cantwell, J. (1989). *Technological Innovation and Multinational Corporations*. B. Blackwell.

Casadesus-Masanell, R. and H. Halaburda (2014). When Does a Platform Create value by Limiting Choice? *Journal of Economics & Management Strategy 23*(2), 259–293.

Cassar, G. (2004). The Financing of Business Start-ups. *Journal of Business Venturing 19*(2), 261–283.

Cattani, G. (2006). Technological Pre-Adaptation, Speciation, and Emergence of New Technologies: How Corning Invented and Developed Fiber Optics. *Industrial and Corporate Change 15*(2), 285–318.

Cennamo, C. and J. Santalo (2013). Platform Competition: Strategic Trade-offs in Platform Markets. *Strategic Management Journal 34*(11), 1331–1350.

Choi, H. and H. Varian (2012). Predicting the Present with Google Trends. *Economic Record 88*, 2–9.

Chung, W. and J. Alcácer (2002). Knowledge Seeking and Location Choice of Foreign Direct Investment in the United States. *Management Science 48*(12), 1534–1554.

Claussen, J., T. Kretschmer, and P. Mayrhofer (2013). The Effects of Rewarding User Engagement: The Case of Facebook Apps. *Information Systems Research 24*(1), 186–200.

Cohen, W. M. . and D. A. Levinthal (1989). Innovation and Learning: The Two Faces of R&D. *The Economic Journal 99*(397), 569–596.

Cosh, A., D. Cumming, and A. Hughes (2009). Outside Entrepreneurial Capital. *The Economic Journal 119*(540), 1494–1533.

Courtney, C., S. Dutta, and Y. Li (2017). Resolving Information Asymmetry: Signaling, Endorsement, and Crowdfunding Success. *Entrepreneurship Theory & Practice 41*(2), 265–290.

Criscuolo, C., J. E. Haskel, and M. J. Slaughter (2010). Global Engagement and the Innovation Activities of Firms. *International Journal of Industrial Organization 28*(2), 191–202.

Cumming, D. and Y. Zhang (2018). Are Crowdfunding Platforms Active and Effective Intermediaries? Working Paper.

De Rassenfosse, G. and E. Raiteri (2017). Technology Protectionism and the Patent System: Strategic Technologies in China. SSRN No. 2803379.

Deutsche Bundesbank (2019). Bestandserhebung über Direktinvestitionen. *Statistische Sonderveröffentlichung 10, April 2019*.

Dorner, M. and D. Harhoff (2018). A Novel Technology-Industry Concordance Table Based on Linked Inventor-Establishment Data. *Research Policy 47*(4), 768–781.

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

Dushnitsky, G. and M. Fitza (2018). Are We Missing the Platforms for the Crowd? Comparing Investment Drivers Across Multiple Crowdfunding Platforms. *Journal of Business Venturing Insights 10*.

Dushnitsky, G. and D. Zunino (2019). The Role of Crowdfunding in Entrepreneurial Finance. In H. Landström, A. Parhankangas, and C. Mason (Eds.), *Handbook of Research on Crowdfunding*, Chapter 3, pp. 46–92. Elgar.

Egger, P., M. Fahn, V. Merlo, and G. Wamser (2014). On the Genesis of Multinational Foreign Affiliate Networks. *European Economic Review 65*, 136–163.

Einav, L., C. Farronato, and J. Levin (2016). Peer-to-Peer Markets. *Annual Review of Economics 8*, 615–635.

Eisenmann, T. R. (2006). Internet Companies' Growth Strategies: Determinants of Investment Intensity and Long-term Performance. *Strategic Management Journal 27*(12), 1183–1204.

Eisenmann, T. R., G. Parker, and M. W. Van Alstyne (2011). Platform Envelopment. *Strategic Management Journal 32*(12), 1270–1285.

European Commission (2016). Internationalisation of Business Investments in R&D and Analysis of Their Economic Impact.

Evans, D. S. (2012). Governing Bad Behavior by Users of Multi-Sided Platforms. *Berkeley Technology Law Journal 27*(2), 1201–1250.

Feldman, M. P. and D. B. Audretsch (1999). Innovation in Cities: Science-Based Diversity, Specialization and Localized Competition. *European Economic Review 43*(2), 409–429.

Fradkin, A. (2017). Search, Matching, and the Role of Digital Marketplace Design in Enabling Trade: Evidence from Airbnb. Working Paper.

Freeman, R. B. and W. Huang (2015). China's Great Leap Forward in Science and Engineering. In A. Geuna (Ed.), *Global Mobility of Research Scientists. The Economics of who goes where and why*, pp. 155–175. Academic Press.

Fu, J., R. Frietsch, and U. Tagscherer (2013). Publication Activity in the Science Citation Index Expanded (SCIE) Database in the Context of Chinese Science and Technology Policy from 1977 to 2012. Fraunhofer ISI Discussion Papers Innovation Systems and Policy Analysis.

Gaessler, F., B. H. Hall, and D. Harhoff (2018). Should There Be Lower Taxes on Patent Income? NBER Working Paper No. w24843.

Galiani, S., P. J. McEwan, and B. Quistorff (2017). External and Internal Validity of a Geographic Quasi-Experiment Embedded in a Cluster-Randomized Experiment. In M. D. Cattaneo and J. C. Escanciano (Eds.), *Regression discontinuity designs*, Volume 38 of *Advances in Econometrics*, pp. 195–236. United Kingdom: Emerald Publishing.

Goldberg, P. K., A. K. Khandelwal, N. Pavcnik, and P. Topalova (2010). Imported Intermediate Inputs and Domestic Product Growth: Evidence from India. *Quarterly Journal of Economics 125*(4), 1727–1767.

Goldfarb, A. and C. Tucker (2019). Digital Economics. *Journal of Economic Literature 57*(1), 3–43.

Griffith, B. R., R. Harrison, and J. V. A. N. Reenen (2006). How Special Is the Special Relationship ? Using the Impact of U.S. R & D Spillovers on U.K . Firms as a Test of Technology Sourcing. *The American Economic Review 96*(5), 1859–1875.

Griffith, R., S. Lee, and B. Straathof (2017). Recombinant Innovation and the Boundaries of the Firm. *International Journal of Industrial Organization 50*, 34–56.

Griffith, R., S. Redding, and J. Van Reenen (2004). Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries. *Review of Economics and Statistics 86*(4), 883–895.

Gruber, M., D. Harhoff, and K. Hoisl (2013). Boundaries : Scientists vs . Engineers. *Management Science 59*(4), 837–851.

Gu, S., B. Å. Lundvall, L. Ju, F. Malerba, and S. S. Serger (2009). China's System and Vision of Innovation: An Analysis in Relation to the Strategic Adjustment and the Medium- to Long-Term S&T Development Plan (2006-20). *Industry and Innovation 16*(4-5), 369–388.

Hagiu, A. (2011). Quantity Vs. Quality: Exclusion by Platforms with Networks Effects. Working Paper.

Hagiu, A. and J. Wright (2020). Platforms and the Exploration of New Products. *Management Science 66*(4), 1527–1543.

Halaburda, H., M. Jan Piskorski, and P. Yıldırım (2017). Competing by Restricting Choice: The Case of Matching Platforms. *Management Science 64*(8), 3574–3594.

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

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

Harhoff, D., F. Narin, F. M. Scherer, and K. Vopel (1999). Citation Frequency and the Value of Patented Inventions. *Review of Economics and Statistics 81*(3), 511–515.

Haskel, J. E., S. C. Pereira, and M. J. Slaughter (2007). Does Inward Foreign Direct Investment Boost the Productivity of Domestic Firms? *Review of Economics and Statistics 89*(3), 482–496.

Howell, S. T. (2018). Joint Ventures and Technology Adoption: A Chinese Industrial Policy that Backfired. *Research Policy 47*(8), 1448–1462.

Imbens, G. and J. M. Wooldridge (2007). Difference in Difference Estimation. NBER Lecture Note.

Jaffe, A. B. (1986). Technological Opportunity and Spillovers of R&D : Evidence from Firms' Patents, Profits, and Market Value. *American Economic Review 76*(5), 984–1001.

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993). Geographic Localization of Knowledge. *The Quarterly Journal of Economics 108*(3), 577–598.

Javorcik, B. S. (2004). Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers Through Backward Linkages. *American Economic Review 94*(3), 605–627.

Jia, N., K. G. Huang, and C. Man Zhang (2019). Public Governance, Corporate Governance, and Firm Innovation: An Examination of State-Owned Enterprises. *Academy of Management Journal 62*(1), 220–247.

Jiang, K., W. Keller, L. D. Qiu, and W. Ridley (2019). International Joint Ventures and Internal vs. External Technology Transfer: Evidence from China. NBER Working Paper No. w24455.

Jones, B. F. (2009). The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder? *Review of Economic Studies 76*(1), 283–317.

Jones, B. F., S. Wuchty, and B. Uzzi (2008). Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science. *Science 322*(5905), 1259–1262.

Jones, C. I. and P. M. Romer (2010, January). The New Kaldor Facts: Ideas, Institutions, Population, and Human Capital. *American Economic Journal: Macroeconomics 2*(1), 224–45.

Kalemli-Ozcan, S., B. Sørensen, C. Villegas-Sanchez, V. Volosovych, and S. Yesiltas (2015). How to Construct Nationally Representative Firm Level Data from the ORBIS Global Database. NBER Working Paper No. w21558.

Kantor, S. and A. Whalley (2019). Research Proximity and Productivity : Long-Term Evidence from Agriculture. *Journal of Political Economy 127*(2), 819–854.

Keller, W. (2002). Geographic Localization of International Technology Diffusion. *The American Economic Review 92*(1), 120–142.

Keller, W. (2010). *International Trade, Foreign Direct Investment, and Technology Spillovers*, Volume 2. Elsevier B.V.

Keller, W. and S. R. Yeaple (2009). Multinational Enterprises, International Trade, and Productivity Growth: Firm Level Evidence from the United States. *The Review of Economics and Statistics 91*(4), 821–831.

Kickstarter (2014). Introducing Launch Now and Simplified Rules. Corporate blog (June 3). Authored by Yancey Strickler.

Kim, K. and S. Viswanathan (2018). The 'Experts' in the Crowd: The Role of Experienced Investors in a Crowdfunding Market. *MIS Quarterly 43*(2), 347–372.

Krugman, P. R. (1979). Increasing Returns, Monopolistic Competition, and International Trade. *Journal of International Economics 9*(4), 469–479.

Kuppuswamy, V. and B. L. Bayus (2017). Does My Contribution to Your Crowdfunding Project Matter? *Journal of Business Venturing 32*(1), 72–89.

Lerner, J. (2002). When Bureaucrats Meet Entrepreneurs: the Design of Effective 'Public Venture Capital' Programmes. *The Economic Journal 112*(477), F73–F84.

Li, C. (2017). Locations of Foreign Direct Investment: The Role of Immigrants, Bilateral Investment Treaties and Patents. Ph.D. thesis, LMU Munich.

Li, J. and S. Netessine (2020). Higher Market Thickness Reduces Matching Rate in Online Platforms: Evidence from a Quasiexperiment. *Management Science 66*(1), 271–289.

Liu, E. (2019). Industrial Policies in Production Networks. *The Quarterly Journal of Economics 134*(4), 1883–1948.

Liu, F. C., D. F. Simon, Y. T. Sun, and C. Cao (2011). China's Innovation Policies: Evolution, Institutional Structure, and Trajectory. *Research Policy 40*(7), 917–931.

Liu, X., S. S. Serger, U. Tagscherer, and A. Y. Chang (2017). Beyond Catch-Up - Can a New Innovation Policy Help China Overcome the Middle Income Trap? *Science and Public Policy 44*(5), 656–669.

Lychagin, S., J. Pinkse, M. E. Slade, and J. V. Reenen (2016). Spillovers in Space: Does Geography Matter? *Journal of Industrial Economics 64*(2), 295–335.

Madsen, J. and J. L. McMullin (2019). Economic Consequences of Risk Disclosures: Evidence from Crowdfunding. *The Accounting Review 95*(4), 331–363.

McIntyre, D. P. and A. Srinivasan (2017). Networks, Platforms, and Strategy: Emerging Views and Next Steps. *Strategic Management Journal 38*(1), 141–160.

Mokyr, J. (2002). *The Gifts of Athena: Historical Origins of the Knowledge Economy*. Princeton University Press.

Mollick, E. (2013). Swept Away by the Crowd. Working Paper.

Mollick, E. and V. Kuppuswamy (2014). After the Campaign: Outcomes of Crowdfunding. Working Paper.

Neuhäusler, P., R. Frietsch, and H. Kroll (2019). Probabilistic Concordance Schemes for the Re-Assignment of Patents to Economic Sectors and Scientific Publications to Technology Fields. Fraunhofer ISI Discussion Papers Innovation Systems and Policy Analysis.

NSF (2018). Science & Engineering Indicators 2018. Technical Report.

OECD (2017). Science, Technology and Industry Scoreboard 2017. Technical Report.

Pavcnik, N. (2002). Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants. *Review of Economic Studies 69*(1), 245–276.

Poege, F., D. Harhoff, F. Gaessler, and S. Baruffaldi (2019). Science Quality and the Value of Inventions. *Science Advances 5*(12), 1–9.

Rochet, J.-C. and J. Tirole (2003). Platform competition in two-sided markets. *Journal of the European Economic Association 1*(4), 990–1029.

Rysman, M. (2009). The Economics of Two-Sided Markets. *Journal of Economic Perspectives 23*(3), 125–143.

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

Schmoch, U. (2008). Concept of a Technology Classification for Country Comparisons. Final Report to the World Intellectual Property Organisation (WIPO).

Shu, P. and C. Steinwender (2019). The Impact of Trade Liberalization on Firm Productivity and Innovation. *Innovation Policy and the Economy 19*(1), 39–68.

Singh, J. (2005). Collaborative Networks as Determinants of Knowledge Diffusion Patterns. *Management Science 51*(5), 756–770.

Sorenson, O., V. Assenova, G.-C. Li, J. Boada, and L. Fleming (2016). Expand Innovation Finance via Crowdfunding. *Science 354*(6319), 1526–1528.

Springut, M., S. Schlaikjer, and D. Chen (2011). China's Program for Science and Technology Modernization: Implications for American Competitiveness. Report prepared for the US-China Economic and Security Review Commission.

Thelwall, M., K. Buckley, G. Paltoglou, D. Cai, and A. Kappas (2010). Sentiment Strength Detection in Short Informal Text. *Journal of the American Society for Information Science and Technology 61*(12), 2544–2558.

Ueda, M. (2004). Banks versus Venture Capital: Project Evaluation, Screening, and Expropriation. *The Journal of Finance 59*(2), 601–621.

UNCTAD (2008). World Investment Report 2008: Transnational Corporations and the Infrastructure Challenge. Technical Report.

Uzzi, B., S. Mukherjee, M. Stringer, and B. Jones (2013). Atypical Combinations and Scientific Impact. *Science 342*(6157), 468–472.

Von Graevenitz, G., C. Helmers, V. Millot, and O. Turnbull (2016). Does Online Search Predict Sales? Evidence from Big Data for Car Markets in Germany and the UK. Working Paper.

Wash, R. and J. Solomon (2014). Coordinating Donors on Crowdfunding Websites. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pp. 38–48. ACM.

Watzinger, M. and M. Schnitzer (2019). Standing on the Shoulders of Science. CRC TRR 190 Discussion Paper Series No. 215.

Weitzman, M. L. (1998). Recombinant Growth. *The Quarterly Journal of Economics 113*(2), 331–360.

Wessel, M., F. Thies, and A. Benlian (2017). Opening the Floodgates: The Implications of Increasing Platform Openness in Crowdfunding. *Journal of Information Technology 32*(4), 344–360.

Xu, C. (2011). The Fundamental Institutions of China's Reforms and Development. *Journal of Economic Literature 49*(4), 1076–1151.

Zacchia, P. (2019). Knowledge Spillovers through Networks of Scientists. *The Review of Economic Studies 87*(4), 1989–2018.

Zhang, J. and P. Liu (2012). Rational Herding in Microloan Markets. *Management Science 58*(5), 892–912.

Zhao, M. (2006). Conducting R&D in Countries with Weak Intellectual Property Rights Protection. *Management Science 52*(8), 1185–1199.

Zhu, F. and M. Iansiti (2012). Entry into Platform-based Markets. *Strategic Management Journal 33*(1), 88–106.